

LEVERAGING SATELLITE DATA AND MACHINE LEARNING TO ENHANCE
PAVEMENT CONDITION ASSESSMENT

by

MOHAMMAD ZOBAIR IBNE BASHAR

M.S., Louisiana State University, 2018

B.S., Islamic University of Technology, 2014

A thesis submitted to the Faculty of the Graduate School
of the University of Colorado Boulder
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

Department of Civil, Environmental, and Architectural Engineering
2022

Committee Members:

Cristina Torres-Machi

Keith Molenaar

Rajagopalan Balaji

Paul Goodrum

Yichang James Tsai

Bashar, Mohammad Zobair Ibne (Ph.D. in Civil Engineering, Department of Civil, Environmental, and Architectural Engineering)

Leveraging Satellite Data and Machine Learning to Enhance Pavement Condition Assessment

Dissertation directed by Assistant Professor Cristina Torres-Machi

Abstract: This research aims at establishing the feasibility of applying satellite data and machine learning (ML) to pavement applications as a way to envision the future of infrastructure asset management. The motivation for this study stems from several knowledge gaps including the assumption of accurate observations in pavement maintenance decision-making, the exclusion of low priority roads from annual distress surveys due to cost constraints, and the lack of quantitative evidence supporting the effectiveness of machine learning algorithms in modeling pavement performance. To address these gaps, the following specific objectives were defined: (i) quantify the value of uncertain optical satellite imagery in the pavement maintenance decision-making, (ii) evaluate the performance of machine learning algorithms in predicting pavement condition, and (iii) apply deep learning to publicly available satellite data to cost-effectively estimate pavement condition. Two satellite data types (i.e., high-resolution multispectral and Synthetic Aperture Radar (SAR) imagery) and different machine learning techniques including partially observable Markov decision process (POMDP) and deep learning were investigated in the context of evaluating pavement condition and making maintenance decisions. The optimal maintenance policies simulated using POMDP models show that satellite observations result in up to 6.5% reductions in cost over the pavement life cycle. The meta-analysis of existing literature indicated that machine learning algorithms can capture on average 15.6% more variability in International Roughness Index (IRI) than traditional techniques. Artificial Neural Network model is

recommended to model IRI because of its consistent performance over a significant number of studies with varying sample sizes and data sources. The framework introduced to model IRI using SAR data was found to be highly effective in addressing the challenges of removing traffic noises from pavements, suppressing speckles without comprising the road features, and eliminating the effects of terrain on SAR backscatters. The resulting deep learning model resulted in accurate IRI predictions with mean absolute errors ranging from 13.9 to 14.6 inches/mile. The associated prediction intervals were found to capture 81% of the actual IRI values within their upper and lower limits. The developed framework was packaged as a software with a graphical user interface to facilitate its implementation by transportation agencies.

ACKNOWLEDGEMENTS

My first and earnest gratitude goes to my advisor Dr. Cristina Torres-Machi for her excellent mentorship throughout the course of my Ph.D. I have been extremely lucky to have an advisor who cared so much not only about my work, but also about my well-being. She always went the extra mile to accommodate my needs, for which I will forever be indebted. Without her assistance and dedicated involvement in every step throughout the process, this work would have never been accomplished.

I also wish to thank my Ph.D. committee Prof. Keith Molenaar, Prof. Rajagopalan Balaji, Prof. Paul Goodrum, and Prof. Yichang James Tsai for their effort in improving my work along the process. I would also like to acknowledge the University of Colorado Boulder and Minnesota Department of Transportation for the funding and facilities to support my research.

Bou, thanks for always being by my side. I am truly grateful for every little thing you ever did to turn my Ph.D. journey into a pleasure.

To my family: Abbu, Ammu, Khosh, and Suba, I would not be here without you. To the rest of my friends, family, and support system, thank you for helping me along the journey: Bahauddin vaiya, Mouree apu, Maria, Bill, Mel, Vignan, Mark, Francisco, Eli, El, Briar, Evan, Irtiza, Shykat, Zihan, Tanvir, Ahnaf, Abid vai, Ratul, Mahdi vai and anyone else I may have missed.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	1
1.1 Problem Statement	1
1.2 Background	2
1.3 Points of Departure.....	12
1.4 Research Objectives	15
1.5 Research Methods	15
1.6 Dissertation Organization.....	17
CHAPTER 2. QUANTIFYING THE VALUE OF SATELLITE-BASED PAVEMENT MONITORING IN PARTIALLY OBSERVABLE STOCHASTIC ENVIRONMENTS....	19
2.1 Abstract	19
2.2 Introduction	20
2.3 Background	22
2.4 Methodology	29
2.5 Results	43
2.6 Summary and Conclusions.....	55
CHAPTER 3. PERFORMANCE OF MACHINE LEARNING ALGORITHMS IN PREDICTING THE PAVEMENT INTERNATIONAL ROUGHNESS INDEX	59
3.1 Abstract	59
3.2 Introduction	60
3.3 Overview of the ML Techniques used in Performance Modeling.....	62
3.4 Methodology	64
3.5 Results	70

3.6	Conclusions and Recommendations.....	81
CHAPTER 4. A DEEP LEARNING FRAMEWORK TO ESTIMATE PAVEMENT ROUGHNESS USING SYNTHETIC APERTURE RADAR DATA		
4.1	Abstract	85
4.2	Introduction	85
4.3	Challenges in Using SAR to Monitor Pavements	87
4.4	Proposed Framework.....	90
4.5	Case Study	99
4.6	Results	104
4.7	Conclusions and Recommendations.....	115
CHAPTER 5. CONCLUSIONS, CONTRIBUTIONS, AND LIMITATIONS.....		
5.1	Conclusions and Recommendations.....	118
5.2	Contributions	121
5.3	Limitations.....	124
5.4	Future Research.....	125
REFERENCES		127

LIST OF TABLES

Table 2-1 Pavement classification criteria based on roughness condition.....	30
Table 2-2 Costs associated with different actions in 1000\$ per lane mile.....	38
Table 2-3 Total expected life-cycle cost estimates with 95% confidence intervals based on 10,000 simulations for a 25-year service life.....	53
Table 3-1 Summary of the selected studies for meta-analysis.....	71
Table 3-2 Summary of the optimum hyperparameters for the RF models.....	79
Table 4-1 RQI performance categories.....	102
Table 4-2 Description of the acquired SAR data	103

LIST OF FIGURES

Figure 1.1 Electromagnetic spectrum showing SAR bands on top and multispectral wavelengths on bottom (adapted from Earthdata 2020).....	6
Figure 1.2 SAR signal penetration behavior for different bands (Earthdata 2020).	7
Figure 2.1 Sequential decision-making within an MDP framework.....	25
Figure 2.2 Sequential decision process in a POMDP model	27
Figure 2.3 Panchromatic satellite images (top row) showing differences in pixel brightness for pavements in (a) good, (b) fair, and (c) poor condition. Ground-based digital images for the same sections are shown in the bottom row for comparison	31
Figure 2.4 Histograms resulting from satellite observations are used to describe pavement states based on different levels of pixel brightness values	36
Figure 2.5 Average DN values across different bands for pavements, when they are classified based on (a) IRI, and (b) rutting	43
Figure 2.6 (a) First and (b) Second-order texture metrics for pavements in good, fair, and poor roughness condition.....	44
Figure 2.7 DN histograms for pavements in good, fair, and poor conditions and their corresponding levels of satellite observations (very low, low, medium, high, and very high)	45
Figure 2.8 Policy realization for (a) <i>Settingperm</i> , and (b) <i>Settingopt</i> with current satellite accuracy..	47
Figure 2.9 Policy realization for <i>Settingopt</i> with 80% accurate satellite inspections	48
Figure 2.10 Policy as a function of belief about pavement condition for the optional inspection action setting with (a) actual, and (b) 80% satellite inspection accuracy, where D = Do nothing, M = Minor repair, A =Annual distress survey, and S = Satellite-based inspection	49

Figure 2.11 Policy realization for (a) <i>Settingblind</i> , and (b) <i>Settingsat</i> with actual satellite observation accuracy	50
Figure 2.12 Optimal value functions for (a) case 1, and (b) case 2, and their corresponding values of satellite-based monitoring at different levels of satellite inspection accuracy, where the shaded region indicates the current range of accuracy	52
Figure 2.13 Expected value of satellite-based monitoring with 95% confidence intervals based on 10,000 simulations for a 25-year service life.....	55
Figure 3.1 Forest plot illustrating the performance of traditional approaches.....	73
Figure 3.2 Forest plot illustrating the performance of ANN algorithms.....	74
Figure 3.3 Forest plot illustrating the performance of RF algorithms.	75
Figure 3.4 Forest plot illustrating the performance of SVM algorithms.....	75
Figure 3.5 Percentage of ANN models with (a) number of hidden layers, (b) number of neurons in input layer, and (c) number of neurons in hidden layer as a function of input features.....	77
Figure 3.6 Predictive performance of SVM algorithms with different kernels.	80
Figure 3.7 Sample size correlations ML algorithms' predictive performance.....	81
Figure 4.1 SAR backscatters depend on the surface roughness. (a) Smooth surfaces will have lower backscattering coefficients than (b) intermediate, and (c) rough surfaces. Image adapted from (Meyer 2019)	88
Figure 4.2 SAR backscattering in the (a) presence, and (b) absence of traffic. Image adapted from (Karimzadeh and Matsuoka 2020).....	89
Figure 4.3 Proposed framework to estimate pavement condition using SAR imagery	91
Figure 4.4: Road reference points overlaid on top of (a) satellite, and (b) processed SAR image	96

Figure 4.5 US-169 pavement surface showing locations with (a) low, and (b) high IRI values.	100
Figure 4.6 Distribution of (a) IRI, and (b) γ_0 values in the final dataset.....	104
Figure 4.7 Backscatters in (a) VV, and (b) VH polarization for pavements in different condition	105
Figure 4.8 Performance of filters in suppressing speckles in pavement pixels.....	106
Figure 4.9 (a) Original image as compared to (b) IDAN, (c) Lee, (d) Refined Lee filtered image	106
Figure 4.10 Processed SAR image (a) without, and (b) with radiometric terrain correction.....	107
Figure 4.11 Seasonal variations in VV backscatter values from pavements over the study period.....	109
Figure 4.12 (a) Satellite image, (b) an individual SAR image, and (c) the minimum intensity projection image generated from a stack.	110
Figure 4.13 Performance of the model during (a) training, and (b) testing.	111
Figure 4.14 (a) Residual plot, and (b) normal Q-Q plot showing the distribution of residuals.	111
Figure 4.15 Performance of (a) simple linear regression model based on γ_0 , and (b) multiple linear regression based on all the features.	112
Figure 4.16 Prediction intervals associated with point estimations in comparison to actual IRI values ..	113
Figure 4.17 Classification accuracy of the model for different RQI classes.....	114
Figure 4.18 SAR-C user interface.....	115

CHAPTER 1. INTRODUCTION

1.1 Problem Statement

Pavement condition is typically characterized by roughness, surface friction, and distresses such as cracking, rutting, patching, raveling, and potholes. State Departments of Transportation (DOTs) primarily use automated distress surveys to record and keep track of these distresses over the years. The process involves a van equipped with cameras, sensors, and radars that runs along the roads and records georeferenced condition data (Herold et al. 2008). A 2019 synthesis study reported the average cost of automated condition surveys to range from \$34 to \$199/mi based on the number of distresses being recorded and the length of the pavement network (Pierce and Weitzel 2019). Despite being an automated process, ground-based pavement monitoring remains an expensive and time-consuming form of data collection due to the amount of driving and post processing required at a network level. As a result, DOTs often limit the pavement condition assessment to major highways, as mandated by federal regulations (FHWA 2016) and only a limited information is available on the condition of ancillary roads such as local roads and ramps.

Recent advances in satellite remote sensing and the availability of high-resolution aerial imagery offer opportunities to cost-effectively assess pavement condition (Pan et al. 2016). Satellite data can be collected monthly at a network level, as compared to the automated approach, where high priority roads may be evaluated every year and the data collection frequency can go up to 2-3 years for low density rural roads (Fagrhi and Ozden 2015). Furthermore, satellite data can also be used for a wide variety of purposes by agencies, which will significantly lower the overall cost of data acquisition. Satellite imagery provides a wide spatial coverage and spectrum, thus can be used to evaluate pavement condition at a network level in a more cost-effective manner. Although many

studies have researched the use of satellite imagery in pavement monitoring, they are largely focused on evaluating the feasibility of this data type in detecting and classifying road surface types (Fagrhi and Ozden 2015; Hoppe et al. 2014; Mohammadi 2012; Pan et al. 2016). Only two studies so far have attempted to estimate pavement condition using satellite data (Meyer et al. 2020; Suanpaga and Yoshikazu 2010). These models, however, do not fully exploit the capabilities of advanced satellite data processing techniques, as detailed in the subsequent chapters, and are also limited in terms of addressing the issues of traffic on road. No standard approach or guideline in estimating pavement condition leveraging this source of big data exists till date.

Integration of Machine Learning (ML) techniques with this approach has the potential to fill up this gap. Several research efforts in the last decade have established machine learning as a superior alternative to current approaches in pavement management (Koch et al. 2015; Salehi and Burgueño 2018; Spencer et al. 2019). Transportation agencies, however, show reluctance in adopting these techniques due to their perception of machine learning being a black box, and the high variability in the performance of these models as reported by previous studies (Abduljabbar et al. 2019; Flintsch and Chen 2004; Pell et al. 2015).

1.2 Background

This section briefly introduces necessary background information on the current practices in monitoring pavements, the concepts of satellite remote sensing, and ML in the context of managing pavements. This leads to the research gaps this dissertation seeks to address and provides an overview of the subsequent chapters of this dissertation.

1.2.1 Current Practice in Pavement Condition Assessment

Pavement condition assessment is an essential part of pavement management, since this is used to define current network condition, predict future condition of the pavements (through performance models), and determine the appropriate time and optimum treatment strategies to repair and rehabilitate the pavements in need (Pierce and Weitzel 2019). The pavement condition surveys typically record the type, severity, and extent of distresses on the pavement surface. Automated distress surveys are conducted to record images and pavement profile using a sophisticated van that collects data at posted speeds. Line scan-technology with a highly focused narrow laser beam is used to scan the pavement surface laterally. This technique is not affected by lighting condition and therefore produces high quality 3D elevation pavement image (Tsai and Li 2012). Advanced 3D systems identify stripes, cracks, and aggregates by capturing intensity of reflected light and measures cracks, spalls, and potholes using 3D elevation data (Wang and Smadi 2011). This data has been used to successfully detect cracks of widths up to 0.04 in (Wang et al. 2015). A significant number of programs and algorithms are available to automatically detect rutting, faulting, raveling, and potholes using 3D laser data. Most of these programs are proprietary and the outcomes are slightly different from one another (Pierce and Weitzel 2019).

While these surveys result in highly accurate and detailed pavement condition data, the coverage and frequency of these surveys are often limited due to budgetary constraints and the efforts required to acquire and process data of such volume. This is a major limitation of the current practice, since early detection of pavement damage can help significantly reduce the progressive development of the distresses if preventive actions are taken on time. Haider et al. (2011) evaluated the impact of pavement monitoring frequency on maintenance decision making and observed that

longer intervals typically underpredict pavement deterioration and overpredict expected life of the pavements. More frequent condition assessment will not only increase the service life of the pavements, but also reduce the life-cycle maintenance costs, and ensure road safety (Li et al. 2017). Computer vision and deep learning based approaches offer the possibility to inspect pavements more frequently using non-conventional sources of data such as: street level imagery, airborne imagery, and satellite data. Koch et al. (2015) reported that a complete automation in computer vision-based defect detection has been achieved for both asphalt and concrete pavements using digital imagery, although defect properties retrieval and estimating pavement condition rating have remained a challenge.

The current automated pavement condition assessment system and the proposed spaceborne evaluation techniques are two different forms of remote sensing. Therefore, to better distinguish between these two, the current approach will be referred to as a ‘ground-based’ technique throughout this dissertation, whereas the spaceborne remote sensing techniques which will be referred to as satellite remote sensing.

1.2.2 Satellite Remote Sensing

Satellite remote sensing refers to the process of acquiring distant information using a spaceborne device. Physical characteristics of an area is detected and monitored by measuring the emitted and reflected radiation from a distance. To effectively acquire imagery at a global scale, the earth observation satellites typically orbit the earth at a relatively low altitude of about 400 to 500 mi from the surface of the earth (ICEYE 2021). The efforts in using images collected from these satellites to monitor pavements are mostly concentrated around optical imagery and synthetic

aperture radar data (Murdzek et al. 2018; Schnebele et al. 2015). These two methods, however, are fundamentally different and as a result, they provide certain edges over one another.

Optical Imagery

Optical satellite imaging is a passive mode of remote sensing since these sensors do not have a source of radiation of their own. These sensors can only image parts of the world that reflect sunlight. Another major limitation of optical imagery is that it cannot penetrate clouds. It is also affected by a wide variety of other factors such as atmospheric visibility, angle of sun radiation, and view angle of the sensor. These images are acquired at wavelengths within visible light (400 – 700 *nm*), and near infrared bands (700 – 1100*nm*) as shown in Figure 1.1. These images can be used to view ground features such as pavements in the same way a human eye does. This allows for easy interpretation of the surface or ground feature being analyzed. Despite of having certain limitations as pointed out earlier, spaceborne optical sensors (e.g., Landsat, MODIS, ASTER, IKONOS, Quick Bird, Orbview, GeoEye, Worldview) have been proven to be highly effective for a wide variety of applications ranging from agriculture to security monitoring. Emergence of very of high resolution (~30cm) optical imagery have received the attention of researchers to estimate road surface condition using this technique.

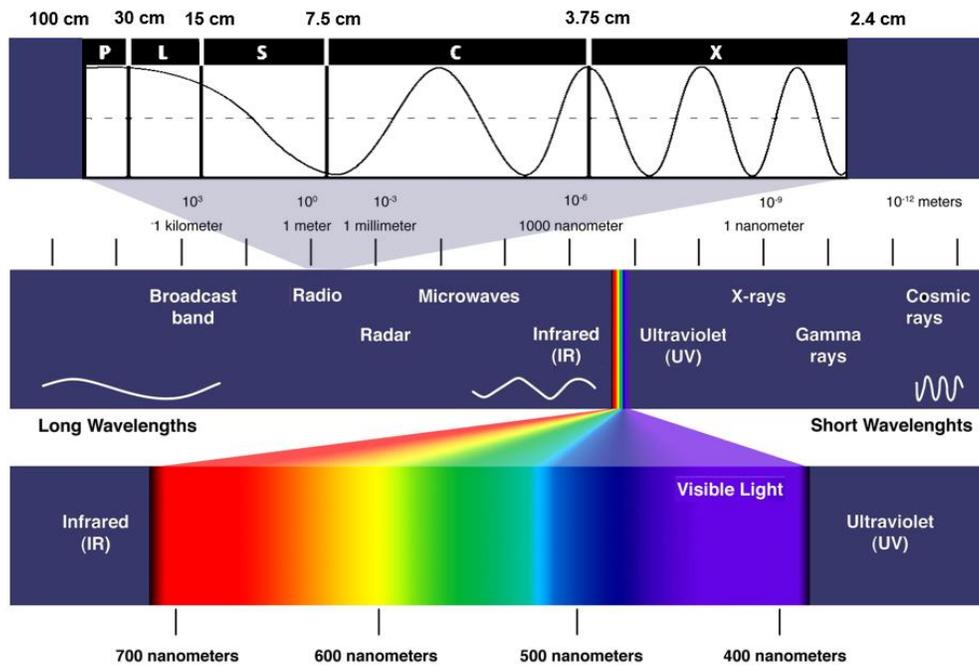


Figure 1.1 Electromagnetic spectrum showing SAR bands on top and multispectral wavelengths on bottom (adapted from Earthdata 2020).

Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a satellite imaging system that uses a small antenna to simulate a long synthetic antenna by using the motion of a satellite (Fagrhi and Ozden 2015). These sensors have their own source of radiation and as a result, they can monitor the earth day and night. Cloud penetrating ability makes this technique highly versatile as compared to the optical sensors. They can send and receive signals in specific polarizations, which is very useful in characterizing features on the ground. Interferometry is another powerful feature of SAR, which exploits the temporal phase information recorded by the sensors to detect changes in land surface topography. SAR sensors operate within microwave bands with wavelengths ranging from 2.4 to 100cm. Figure 1.1 illustrates the bands: X, C, S, L and P which are typically found in SAR sensors. X-band sensors result in a high-resolution SAR image. This is predominantly used in urban

infrastructure monitoring since these signals scatter close to surface, while C and L-band signals penetrate deeper into the surface as shown in Figure 1.2. C-band data is referred to as SAR workhorse as this is widely used in global mapping, change detection, monitoring areas with low vegetation, ice, ocean, and maritime navigation (Meyer 2019). S-band data is mostly limited to earth observation and agricultural monitoring. L-band images are considered as medium resolution and are typically used in interferometric applications. Spaceborne P-band SAR is still in experimental phase and is intended to study biomass. SEASAT was the first satellite with an L-band SAR imaging system which was launched in 1978. Success of this satellite led to the launch of C-band satellites ERS-1, ERS-2, ENVISAT, Radarsat-1, Radarsat-2, Sentinel-1 and X-band satellites TerraSAR-X, TanDEM-X, COSMO-SkyMed, and PAZ SAR. While all the X-band satellites are commercial, Sentinel-1 data is available for public use at no cost.

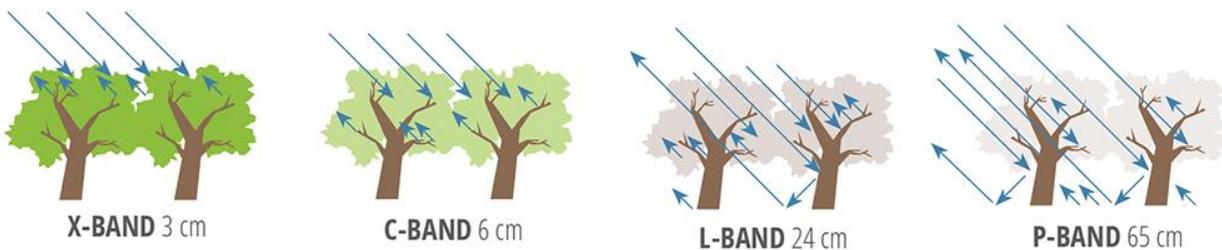


Figure 1.2 SAR signal penetration behavior for different bands (Earthdata 2020).

1.2.3 Machine Learning

ML is the science of making computers learn and act intelligently and improving their learning by feeding them data in the form of observations and real-world interactions (Faggella 2020). While ML techniques has been around for more than 60 years, it has only recently showed its true potential with the advent of big data, and the advancement in processing power and data storage technology. Pavement management, however, is still mainly based on traditional computational

techniques. The most common approaches used in predicting pavement condition include regression analysis, Markov chains, mechanistic-empirical, survivor curves, semi-Markov, and Bayesian models (Flintsch and Chen 2004; Li 2018). While the sophisticated ground-based monitoring systems are resulting in a high volume of data with a great number of variables, these traditional approaches are often ineffective in capturing complex relationships in big and noisy data (Salehi and Burgueño 2018; Spencer et al. 2019).

Although the presence of machine learning in current pavement management practices are insignificant, ML-based techniques have been researched extensively in relation to different aspects of the pavement management process. Due to the exponential growth and success of the ML algorithms, the civil infrastructure community has witnessed substantial efforts in applying ML to a wide variety of processes related to pavements. Before identifying the research needs in this area, we would like to briefly introduce the ML algorithms relevant to the following chapters in this dissertation. We identify four major types of ML applications in relation to the pavement management process: (1) Regression, (2) Classification, (3) Clustering, and (4) Deep Learning.

Regression

Regression algorithms are used to predict future events and identify patterns in data. It is a form of supervised learning where labelled data is used to train the models to investigate the relationship between a dependent and one or more independent variables. This form of predictive modeling usually results in an output that is continuous in form. This is useful for a wide variety of applications in pavement management including estimating pavement condition measures such as the International Roughness Index (IRI). Commonly used regression algorithms include:

- Artificial Neural Network (ANN): Used for very large data sets and when high degree of accuracy is desired at the cost of greater processing time and computational power.
- Support Vector Regression (SVR): Used for smaller but complex dataset.
- Ridge regression: Used to analyze multiple regression data that suffers from multicollinearity.
- Lasso Regression: Suitable for simple models with fewer features (i.e., sparse models) showing high levels of multicollinearity and where the data values shrink towards a central point.
- ElasticNet Regression: Emerged as a solution to the limitations of both lasso and ridge regression as the variable selection in these two algorithms are sometimes too much dependent on data.
- Bayesian Regression: Used for modeling insufficient or poorly distributed data.
- Decision Tree Regression: Used when the data has non-linear shape. It breaks down the data into smaller subsets capturing the non-linearity while incrementally developing an associated decision tree.
- Random Forest Regression: Used for large datasets with higher dimensionality. It is effective in dealing with missing data and maintains accuracy when a large proportion of the data is missing.

Classification

Classification algorithms are used to identify the category or class of an observation. It is a form of supervised learning where the categories of the objects are already known for training and testing purposes. The algorithms learn from the input data and use this learning to make a prediction in a discrete form. This is useful in classifying pavements based on their characteristics and surface condition. The major classification algorithms include:

- Logistic Regression: Used when the dependent variable is binary or dichotomous.
- K-Nearest Neighbors: It assumes that similar things exist in close proximity. Usually used for a smaller dataset.
- Support Vector Classifier (SVC): Finds optimal hyperplanes in an N-dimensional space (with N being the number of features) that distinctively classifies the data points.
- Kernel-SVM: The use of kernels enables these models to effectively operate in high-dimensional feature space. Common kernel functions include linear, polynomial, radial basis, and sigmoid functions.
- Naïve Bayes: It assumes that the presence of a particular feature in a class is not related to the presence of any other function. It is very effective for larger datasets.
- Decision Tree Classification: The data is continuously split according to a certain parameter until a classification decision has been made. These are often fast and accurate.
- Random Forest Classification: Consists of a large number of decision trees where the final decision is made based on the votes from each individual tree.

Clustering

Clustering algorithms are used to group sets of objects based on their similarities in a multidimensional space. It is a form of unsupervised learning where unlabeled data without any preprocessing is used to train the models. The output of these models is usually not exactly known before the training. Therefore, clustering algorithms are mainly used to discover some structures in the dataset which is useful in investigating the spectral and textual properties of pavements. The major clustering algorithms include:

- **K-Means Clustering:** It is an iterative process that tries to minimize the distance between the points within a cluster. It is very good at capturing the data structure if the clusters have a spherical-like shape.
- **Gaussian Mixture Models (GMM):** It considers both the distance and variance of the data points. It can also handle very oblong clusters in contrast to the K-Means clustering.

Deep Learning

Deep learning algorithms are used to extract information from very large datasets of images, videos, texts, and audios. These models consist of multiple layers of non-linear information processing. These algorithms are inspired by the structure and function of the brain called artificial neural networks. These models require access to a large amount of data to be effective. These can be, therefore, leveraged in analyzing big image data sources such as optical satellite imagery and SAR data. State-of-the-art deep learning algorithms include:

- Convolutional Neural Network (CNN): Mainly used for image processing, recognition and classification, video recognition, pattern recognition, natural language processing tasks, and text analytics.
- Recurrent Neural Networks (RNN): It is a powerful tool to process sequential and temporal data such as: sound, time series data, and written natural language.
- Long Short-Term Memory Networks (LSTM): It is a sophisticated RNN algorithm which can store the patterns in the memory for an extended period of time, allowing the model to selectively recall or delete data.
- Deep Belief Networks (DBN): These are useful for image and video-sequence recognition, motion-sensitive data, and classifying high-resolution satellite image data.

1.3 Points of Departure

1.3.1 Potential of Optical Imagery in Pavement Monitoring

A great majority of the optical imagery based research efforts have used airborne systems such as: aircrafts and unmanned aerial vehicles (Herold et al. 2004; Mettas et al. 2015) to evaluate road surface condition. Spaceborne optical imagery applications in pavement monitoring are found in (Mei et al. 2014; Mohammadi 2012; Noronha et al. 2002; Nussbaum et al. 2006; Pan et al. 2016; Shahi et al. 2017). Most of these approaches are fundamentally based on characterization of road surfaces based on spectral information extracted from the multispectral or hyperspectral imagery. While they have shown promising performance in classification tasks, none of the studies have attempted to estimate pavement condition. The availability of very high-resolution multispectral

and hyperspectral optical imagery offers the opportunity to augment spectral-based approaches with textural information derived from these imagery (Kupidura 2019).

To this end, we will explore the capabilities of optical imagery in pavement monitoring. This study will advance the current understanding on the use of uncertain sensor data in pavement maintenance decision making. The findings of this study will also establish the value that uncertain satellite-inspections may add to an existing pavement maintenance decision making system.

1.3.2 Capabilities of Machine Learning in Improving Estimation Accuracy

Most of the research efforts in using ML to improve data driven processes in pavement management are focused on modeling pavement performance. Widely applied algorithms include ANN (Abdelaziz et al. 2020; Chandra et al. 2013; Sollazzo et al. 2017; Yamany et al. 2020), Decision Tree (Kang et al. 2010; Zeiada et al. 2020), Random Forest (RF) (Gong et al. 2018; Marcelino et al. 2019, 2020), Support Vector Machine (SVM) (Georgiou et al. 2018; Kargah-Ostadi and Stoffels 2015; Zeiada et al. 2020; Ziari et al. 2016a), and Recurrent Neural Network (Choi and Do 2020). The application of these algorithms in evaluating pavement condition can be found in (Nitsche et al. 2012; Tsai et al. 2021; Wang et al. 2017; Zhang et al. 2018).

While all these studies were able to accomplish their respective objectives, a high disparity in performance of ML models were observed in terms of the reported accuracy and the size of the dataset used. The lack of a standard guideline to apply ML for pavement applications may be hindering the adoption of these techniques in practice. Furthermore, since these research efforts are highly scattered, a complete picture and understanding of the ML literature in pavement performance monitoring is still missing. Therefore, it is difficult to quantitatively evaluate the true performance of these algorithms.

To address this gap, we will identify all the studies attempted to estimate pavement condition using ML techniques and quantitatively establish an overall performance for each of these algorithms. The findings and recommendations derived from this study will serve as a reference for the practitioners aiming to use ML in solving pavement management problems.

1.3.3 Leveraging Machine Learning and SAR Imagery to Assess Pavement Condition

SAR has been researched widely for monitoring pavements and fields closely related to pavements. These applications include: detection of pavement surface deformation (Goel and Adam 2014; Hoppe et al. 2014), sinkhole formation (Bruckno et al. 2013), rock slope monitoring (Bruckno et al. 2013; Hoppe et al. 2014), and pavement surface distress detection (Hoppe et al. 2014). These studies, however, are applicable at a project level. Network scale applications of SAR data in evaluating road surface condition are found in (Meyer et al. 2020; Suanpaga and Yoshikazu 2010). Meyer et al. (2020) developed a regression model to estimate pavement roughness using high-resolution X-band SAR data. Suanpaga and Yoshikazu (2010) developed a multinomial logit model using medium resolution L-band SAR data. Although both of these studies reported SAR to be a strong candidate for an alternative method of pavement condition assessment, their findings about the most suitable polarization (VV vs. HH) for studying pavements are highly contrasting. The models presented in these studies are also not the most efficient in leveraging big data. Furthermore, the capabilities of publicly available moderate to high resolution (5 meter) C-band SAR data have not been researched in estimating pavement condition. C-band data is a very good compromise between cost and resolution as compared to high resolution X-band SAR data which is only available commercially.

Therefore, we will exploit the opportunities that SAR has to offer and leverage machine learning capabilities to derive a tool to estimate pavement roughness at a network level. The study will specifically focus on establishing a SAR processing framework that is more suitable to address pavement related issues such as traffic, snow, and tall objects near roads.

1.4 Research Objectives

The primary objective of this research is to assess the capabilities of satellite data and machine learning in evaluating road surface condition. To specifically address the research gaps identified in the previous section, the primary objective is further broken down into the following research objectives (ROs):

- RO1.** Quantify the value of including satellite imagery in optimal inspection and maintenance strategies over the pavement life cycle.
- RO2.** Evaluate the performance of machine learning algorithms in predicting pavement condition as compared to traditional techniques.
- RO3.** Develop a machine-learning based approach to assess pavement condition using publicly available Synthetic Aperture Radar (SAR) data.

1.5 Research Methods

Based on the review, we identify two opportunities: application of satellite data, and machine learning to enhance pavement condition assessment practices. In the next two chapters (i.e., Chapter 2 and 3), we independently explore the capabilities of optical imagery and machine learning in the context of improving pavement asset management. In Chapter 4, we combine the findings from both these topics to develop a ML-based model to estimate pavement roughness

using SAR data. While the use of optical imagery and SAR data in Chapter 2 and 4 are interchangeable, we decided to explore this particular combination (i.e., SAR and ML) because ML models require a great volume of data and SAR data is accessible at no direct cost to us; as compared to the high-resolution multispectral imagery, which is expensive to acquire. The following section provides a very brief overview of the methods proposed for each of these chapters.

1.5.1 Incorporation of Uncertain Satellite Data in Pavement Monitoring Decision-Making

Data obtained from optical satellite imagery is characterized as imprecise, and uncertain. we analyze whether the inclusion of such information obtained from satellites reduces the pavement life-cycle cost by optimizing the pavement inspection and maintenance decision making. We model the problem as a Partially Observable Markov Decision Process (POMDP). POMDP offers a sound mathematical framework for sequential decision making under uncertain dynamic environments, which can be conveniently adapted for the pavement maintenance decision making. The solutions obtained from this process will result in optimum policies on when and how to collect and use pavement condition data. Part of this study has been published at *Reliability Engineering and System Safety* (Seites-Rundlett et al. 2021).

1.5.2 Meta-Analysis of the Pavement Performance Prediction Models

To evaluate the true effectiveness of machine learning models as compared to the traditional regression-based techniques, a meta-analysis of relevant studies is performed. Meta-analysis is a statistical approach to combine quantitative research findings from multiple empirical studies. In this approach, the effect sizes from different studies are combined to increase power and capture

the true effect, thus allowing to summarize and compare empirical research studies. The process of conducting a meta-analysis includes a comprehensive review of literature, extracting the data of interest, standardizing effect sizes, and estimating the overall effect size (Alruqi and Hallowell 2019). A detailed analysis of the model architecture is carried out to identify optimal values of hyperparameters in modeling pavement roughness. The effect of sample size on the predictive performance is also investigated to derive recommendations on a minimum sample size. This study has already been published at the *Transportation Research Record: Journal of the Transportation Research Board* (Bashar and Torres-Machi 2021).

1.5.3 Apply Deep Learning to Estimate Pavement Condition using SAR data

Historical in-situ pavement condition data, pavement features, and radar signals extracted from a processed SAR imagery are used to train a pavement roughness estimation model using state-of-the-art deep learning techniques. The study will focus more on the development of a framework that will effectively address the challenges associated pavement applications such as: filtering out traffic and other noises from the images without compromising the texture and edge of the linear road features. The developed model is envisioned to work as follows: for a given SAR image and road features, the model will estimate pavement roughness, associated prediction intervals, and road quality class. This paper is under review at the *Automation in Construction*.

1.6 Dissertation Organization

This dissertation is organized in a three-paper format with the addition of an introduction and a conclusion chapter at the beginning and end of the dissertation. Chapter 1 introduces the problem, describes background, identifies research needs, and defines the specific objectives to address the

knowledge gaps. Chapter 2, 3, and 4 follow the format of a journal research paper, where each of these papers addresses one of the three objectives described in section 1.4. Chapter 5 summarizes the findings and contributions resulting in from three of the preceding chapters.

CHAPTER 2. QUANTIFYING THE VALUE OF SATELLITE-BASED PAVEMENT MONITORING IN PARTIALLY OBSERVABLE STOCHASTIC ENVIRONMENTS

2.1 Abstract

Accurate and timely assessment of pavement condition is critical to determine optimal maintenance plans. Due to the high costs of ground-based inspections, agencies often limit their monitoring to major roads, as required by federal regulations. As a result, the condition of some elements of the road network, such as ancillary roads, often remains unknown. Satellites, capable of rapidly collecting information over wide areas, can be a cost-effective alternative to monitor pavement condition. This wide coverage, however, comes at the expense of lower levels of accuracy. The objective of this study is to quantify the value of satellite-based information in optimal inspection and maintenance strategies. To account for the uncertainties associated with satellite observations, the system is modeled as a partially observable Markov decision process (POMDP) to determine optimal life-cycle inspection and maintenance policies. To estimate the value of information obtained from satellite inspections, two cases representing current pavement condition practices were simulated: (1) as an alternative to inspect highways, roads which are traditionally monitored with annual automated distress surveys, and (2) as an option to inspect local or ancillary roads, which are not typically monitored. Results indicate that satellite observations result in up to 6.5% reductions in cost if it is used to make maintenance and inspection decisions over the pavement life cycle. Savings are higher for non-monitored roads, as compared to major roads that are annually inspected with automated distress surveys. Satellite information was found to become valuable at 70% level of accuracy when used in combination with more accurate systems.

2.2 Introduction

The Federal Highway Administration (FHWA) mandates the State Departments of Transportation (DOT) to annually monitor and report pavement condition in terms of the International Roughness Index (IRI) for their entire national highway system, freeways, expressways, and principal arterials (FHWA 2016). The DOTs are also required to collect rutting data for asphalt pavements, faulting data for jointed concrete pavements, and the percentage of surface cracks for all pavement types. These inspections result in a high volume of pavement condition data that the DOTs use to prepare short- and long-term plans for maintaining and improving the road network. The collection of such detailed and accurate condition data, however, comes at a great cost. Michigan DOT spends \$2.5M on average annually on pavement data collection and analysis and Illinois, Kansas, and Wisconsin DOTs reported their annual budget for data collection to be around \$750k to \$1M (Hicks et al. 2011). According to a 2019 synthesis study on automated pavement condition surveys, the average cost of data collection ranges from \$34 to \$199/mi (Pierce and Weitzel 2019).

Given the budgetary constraints the transportation agencies face, spending a significant amount of public funds on evaluating the entire road network annually with high certainty is far from optimal. A newly constructed section, for example, may not need a detailed evaluation in the years immediately after construction. Moreover, the high costs of inspection often limit the capabilities of agencies to monitor the condition of the ancillary components of the highway system (e.g., ramps, auxiliary lanes, and frontage road pavements). As a result, the condition of ancillary roads often remains unknown to decision-makers. Optimizing the inspection and monitoring activities for the pavement network can therefore significantly reduce the cost of managing pavement assets.

Satellite data can provide valuable information for evaluating pavement condition in a cost-effective manner (Fagrhi and Ozden 2015; Hoppe et al. 2014; Li et al. 2017). Based on the average miles of roadway per square mile of land in the United States (FHWA 2015), a processed high resolution (i.e., 30 cm) satellite image costs about \$30.70/mile of pavement. While this is already cheaper than collecting pavement condition data using the ground-based methods described above, this cost goes down significantly for urbanized areas with more roads per square mile (e.g., \$11.10/mile for New Jersey).

A significant number of studies over the last decade have established a preliminary foundation for the use of satellite imagery in mapping road surface condition (Emery 2014; Pan et al. 2016; Shahi et al. 2017). These studies have found evidence of the potential of satellite imagery to inspect road surface condition, with the main limitation of this information being the limited accuracy of satellite-based information. In this paper, we explore the capabilities of satellite-based monitoring as an alternative inspection option that could supplement ground-based monitoring and reduce the overall costs of pavement inspections. The objective of this study is to quantify the value of satellite-based information in optimal inspection and maintenance strategies over the life cycle of pavements. To account for the uncertainties inherent to satellite-based inspections, we use Partially Observable Markov Decision Process (POMDP) to estimate optimal inspection and maintenance strategies over the pavement life cycle. Compared to traditional Markov Decision Process that assumes pavement condition to be known with certainty, POMDP considers the decision-maker can only observe this partially, through noisy observations. POMDPs offer an ideal mathematical framework because it allows to incorporate incomplete information obtained from sources such as

satellite data and the stochastic deterioration process in the design of optimal maintenance and inspection strategies (Papakonstantinou et al. 2018).

2.2.1 Objectives

The primary objective of this study is to quantify the value of information obtained from satellite imagery in optimal maintenance and inspection strategies over the pavement life cycle. To quantify the value of incorporating highly uncertain satellite data, the system is modeled as a POMDP problem and two cases are evaluated to reflect the current practice for (1) highways (i.e., roads which are typically monitored annually using automated distress surveys), and (2) local or ancillary roads (i.e., roads which are not typically monitored). The study derives recommendations on using the estimated value of information to make investment decisions while adopting a new technology with a certain degree of sensor accuracy.

2.3 Background

This section summarizes the literature and highlight the gaps on the use of optical satellite imagery to evaluate pavement condition. It also presents the theoretical premises of the POMDP framework used to estimate the value of satellite-based information in optimal maintenance and inspection strategies.

2.3.1 Optical Satellite Imagery in Pavement Monitoring

Satellite imagery captured within visible light and near-infrared wavelengths can be broadly classified as panchromatic, multispectral, and hyperspectral. A panchromatic image is essentially black and white and consists of just one band offering high spatial resolution. Multispectral images supply rich spectral information, as it consists of 3-10 broad discrete bands including red, green,

blue, yellow, infrared, and short-wave infrared. Panchromatic and multispectral images can be combined to produce a pansharpened image that improves the image quality by sharpening the coarse resolution of multispectral images. Hyperspectral images, on the other hand, capture hundreds of very narrow bands of visible lights and provide more detailed information than the other types of satellite images.

Previous research has explored the use of these three image types and varying degrees of spatial resolution to assess pavement condition. Shahi et al. (2017) investigated the capabilities of pansharpened imagery in classifying the level of pavement damage using an object-based image processing technique. The proposed model estimated the condition of asphalt pavements with an 83% accuracy, although only two levels of damage were considered. Also, the model exhibited limitations to differentiate between the road surface and colored vehicles due to their spectral similarity. In a similar study, Pan et al. (2016) used multispectral images and, in particular, the reflectance and slope of the spectral signature, to classify the aging of asphalt pavements using three categories (i.e., light, medium, and heavily aging). The results indicated that aging pavements are characterized by different spectral features at wavelengths ranging from 0.35 to 2.5 μm . This study, however, did not find a clear trend to differentiate among different distress types. Seites-Rundlett et al. (2021) extended the use of multispectral imagery in estimating pavement roughness by developing an Evidence Theory-based model which effectively address the conflicts that may rise while combining data from multiple sensors with varying degrees of uncertainty.

Other applications of hyperspectral imagery in pavement research can be found in Herold et al. (2008); Herold and Roberts (2005); Mohammadi (2012); and Noronha et al. (2002). Mohammadi (2012) found that the hydrocarbon absorption bands (i.e., the wavelengths ranging from 1.7082 to

1.7323 μm) are the most suitable for identifying different states of aging asphalt. Noronha et al. (2002), however, concluded that using hyperspectral imagery for road surface condition assessment is 'not perfect' due to the uncertainties associated with within-class spectral variability between different distress types. The spectral similarities between pavement surfaces and roofing materials also make this approach limited.

Regardless of the type of image and methodology used, all these studies have found evidence of the potential of satellite imagery to inspect road surface condition, with the main limitation of this information being the limited precision of satellite-based information. None of these studies, however, explore this issue in detail, nor provide guidance on the use satellite-based information on pavement management decision-making. This study addresses this gap by objectively quantifying the value of the information provided by satellite imagery in optimal pavement management scenarios in which the decision-maker has the option to choose inspection and maintenance strategies that minimize costs over the pavement life cycle.

2.3.2 Maintenance Scheduling in Uncertain Stochastic Environments

In the management of pavement infrastructure, decision-makers seek to make inspection and maintenance decisions that minimize the overall costs over the pavement life cycle. In this process, decision-makers use deterioration models to predict the condition of pavements over time and use this information to inform maintenance strategies. These deterioration models, however, have inherent uncertainties that may result in erroneous forecasts of pavement condition. To reduce the uncertainty in these condition forecasts, agencies may decide to inspect the condition of pavements. The optimal management of infrastructure should therefore consider optimal maintenance and inspection strategies that minimize the overall costs over the pavement life cycle.

To account for uncertainties in forecasting infrastructure deterioration, Markov Decision Process (MDP) has widely been used for sequential decision-making in pavement management (Butt et al. 1987; Carnahan 1988; Golabi et al. 1982; Medury et al. 2014). The MDP framework assumes that inspections always reveal the accurate condition of the system with certainty. This may be true for certain applications, but it is unrealistic when it comes to monitoring pavement assets (Amin and Amador-Jiménez 2016). Since the true deterioration of pavements is never observed completely, even an accurate stochastic deterioration model will lead to non-optimum maintenance actions (Papakonstantinou et al. 2018). Another major limitation of MDP is that it requires an inspection to be performed before taking an action at every time step. In this framework, the decision-maker observes the state (s) of a system and takes an action (a) from a finite set of available actions. As a result, the system transitions to a new state (s') and the agent receives a reward (r), as shown in Figure 2.1. This transition, however, only depends on the current state and the action taken and does not take into account the history of preceding states and actions.

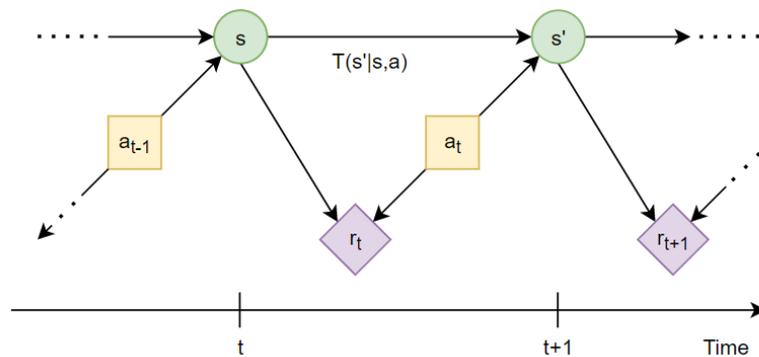


Figure 2.1 Sequential decision-making within an MDP framework

To address these limitations, the Partially Observable Markov Decision Problem (POMDP) was introduced as a generalization of MDP (Smallwood and Sondik 1973; Sondik 1978). POMDPs offer a flexible and mathematically sound decision-making process that does not impose

constraints such as mandatory periodic inspections and perfect observations (Papakonstantinou et al. 2018). In this uncertain environment, infrastructure condition is described using a probability distribution, referred to as belief, over the possible states of the system. This belief is a sufficient statistic of the history and observations –the belief provides the decision-maker with the same amount of information as the full history of maintenance and inspection actions (Papakonstantinou and Shinozuka 2014a). Early applications of POMDPs in infrastructure maintenance scheduling are found in Corotis et al. (2005); Ellis et al. (1995), and Madanat and Ben-Akiva (1994).

Despite its powerful capabilities, this framework was not widely used until recently, since it was very difficult to solve a complex POMDP representing real-life scenarios (Papakonstantinou and Shinozuka 2014b). Recent advances in the field of robotics have led to the development of modern algorithms to approximate POMDP solutions more effectively. Recent research have presented large-scale applications of POMDP and demonstrated the robustness of this framework in deriving optimal inspection and maintenance strategies for infrastructure systems (Andriotis et al. 2020; Kim et al. 2018; Memarzadeh et al. 2015; Memarzadeh and Pozzi 2016; Papakonstantinou and Shinozuka 2014a; Schöbi and Chatzi 2016). All these studies, however, have used synthetic transition probabilities and cost data to demonstrate different applications of the model.

In this paper, we model the POMDP problem with transition probabilities derived from actual field performance of pavements and costs estimated from maintenance history databases. Before discussing the specifics of the model, we introduce the notations and explain the POMDP model in detail in the following section.

2.3.3 POMDP Framework

A POMDP is typically defined as follows (Kamalzadeh and Hahsler 2019):

$$\mathcal{P} = (S, A, T, R, \Omega, O, \gamma, b_0) \quad (2-1)$$

where, $S = \{s_1, s_2, \dots, s_n\}$ is a set of states of a system; $A = \{a_1, a_2, \dots, a_n\}$ is a set of possible actions that an agent can take; T is a set of conditional transition probabilities $T(s'|s, a)$ for the state transition $s \rightarrow s'$; $R: S \times A \rightarrow \mathbb{R}$ is the reward function; $\Omega = \{o_1, o_2, \dots, o_n\}$ is a set of observations the agent has access to; O is the emission probability (i.e., a set of conditional observation probabilities). $O(o|s', a)$ gives, for each action and resulting state, a probability distribution over the observations; $\gamma \in [0,1]$ is a discount factor that discounts the future rewards and relates them to present value; and $b_0 = P(s_0)$ is the initial belief state.

Figure 2.2 illustrates the POMDP sequential decision-making process. At a time t , the environment is in some state $s \in S$. If the agent decides to take an action $a \in A$, the environment transitions to state $s' \in S$ with probability $T(s'|s, a)$. Based on the action and resulting state, the agent receives an observation $o \in \Omega$ with probability $O(o|s', a)$. Finally, a cost or reward $R(s, a)$ is awarded to the agent. When this process is repeated, the agent tries to choose the best action each time, so the expected future discounted reward is maximum, or the cost is minimum.

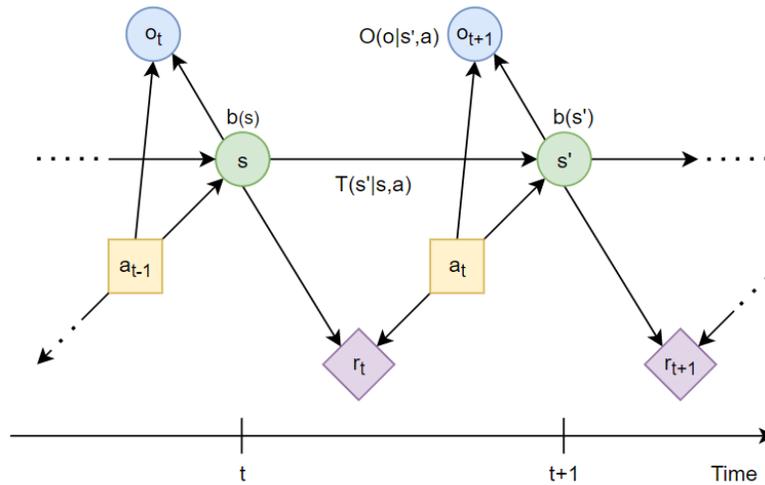


Figure 2.2 Sequential decision process in a POMDP model

In POMDP, the identification of the current state involves uncertainty and therefore, the agent's knowledge about the state is described by a probability distribution called 'belief' $b(s)$ over the state space S (Smallwood and Sondik 1973; Sondik 1978). As the system evolves due to the agent's action, the belief about the resulting state is updated based on the previous belief state, executed action, and received observation. At time $t + 1$, the belief $b(s) \in B$ can be updated to $b(s')$ using Bayes' rule:

$$b(s') = \frac{O(s', a_t, o_{t+1}) \sum_{s \in S} T(s, a_t, s') b(s)}{\sum_{s'' \in S} O(s'', a_t, o_{t+1}) \sum_{s \in S} T(s, a_t, s'') b(s)} \quad (2-2)$$

The behavior of an agent is driven by a 'policy', which is a function $\pi: B \rightarrow A$ that relates beliefs to actions (Memarzadeh and Pozzi 2016). The value of policy π at belief b i.e., V^π is estimated as the expected sum of discounted rewards while being in belief b and taking actions according to that policy. The optimal policy π^* can be computed using Bellman's equation (Bellman 1957):

$$V^*(b) = \max_{a \in A} \left[\sum_{s \in S} b(s) R(s, a) + \gamma \sum_{o \in \Omega} P(o|b, a) V^*(b') \right] \quad (2-3)$$

Where, b is the belief $b(s)$ and b' is the updated belief $b(s')$ according to Equation 2. The conditional probability $P(o_{t+1}|b, a_t)$ can be computed as follows:

$$P(o|b, a) = \sum_{s \in S} O(s', a, o) \sum_{s \in S} T(s, a, s') b(s) \quad (2-4)$$

Solving the POMDP involves identifying the optimal policy π^* for all possible beliefs, or at least for the current belief, to identify optimal action $a_t = \pi^* b(s)$.

Algorithm Selection to Solve POMDP

Algorithms to derive exact solutions for POMDPs are computationally very expensive, and thus, only effective in computing optimal policies for very small problems (Papakonstantinou and Shinozuka 2014a). A breakthrough in approximating POMDP solution was achieved after Pineau et al. (2003) introduced a point-based value iteration algorithm. This approach solves the fundamental constrain in solving POMDP (i.e., the assumption of having an infinite number of states due to the continuity of the belief space). State of the art point-based value iteration algorithms include PERSEUS (Spaan and Vlassis 2005), SARSOP (Kurniawati et al. 2008), HSVI (Smith and Simmons 2012), and FRTDP (Smith and Simmons 2006). Papakonstantinou et al. (2018) evaluated the performance of these algorithms in POMDP planning and observed that SARSOP is the best algorithm in terms of efficiency and accuracy. Since this paper is concerned with approximation of the value function instead of a thorough pursuit of the best possible policy, we use SARSOP to solve the POMDP model presented in the Methodology section.

2.4 Methodology

To optimize the maintenance and inspection strategies considering the uncertainties associated with satellite data, the life cycle of a pavement section is modeled as a POMDP. The optimal solution of the POMDP model is then used to estimate the value of a satellite-based monitoring system. The value of a satellite-based monitoring system, however, will depend on the cost and accuracy of the system relative to the existing system. To account for this, this project evaluates different combinations of monitoring systems with varying degrees of sensor accuracy.

2.4.1 Data Collection

A 10-square-mile area near downtown Denver, Colorado, United States, was selected as the study area for this project. This area was selected because: (i) pavement condition data was available for the roads in this area; (ii) it had a representation of road sections in good, fair, and poor condition; and (iii) there were satellite images collected at a similar time than the field condition data. Condition data for asphalt pavements located in this area was collected from the Colorado Department of Transportation (CDOT) for the years 2013 to 2018. Condition of the roads were classified based on their International Roughness Index (IRI) and rutting values using FHWA criteria (Arhin et al. 2015) as shown in Table 2-1. The condition of the pavements was verified by reviewing the CDOT Online Transportation Information System video logs.

Table 2-1 Pavement classification criteria based on roughness condition

Criteria	Good	Fair	Poor
IRI (in/mile)	< 95	95 – 170	> 170
Rutting (in)	< 0.20	0.20 – 0.40	> 0.40

High-resolution imagery from the WorldView3 satellite was obtained from DigitalGlobe. The satellite data consisted of a 30 cm resolution 1 band 11-bit panchromatic image which was captured on April 20, 2015. The cost of acquiring this image was \$58 per square mile. Figure 2.3 illustrates how pavements in different conditions appear under black and white panchromatic satellite sensor (top row) as compared to their ground-based digital images (bottom row) obtained from the CDOT Online Transportation Information System video logs. Panchromatic satellite images (shown in

the top row images in Figure 2.3) show differences in pavement roughness as rough surfaces appear brighter on panchromatic satellite images as compared to smooth surfaces (Bashar and Torres-Machi 2022). These results suggest that the pixel brightness of panchromatic satellite images can be used to infer information of pavement roughness condition.

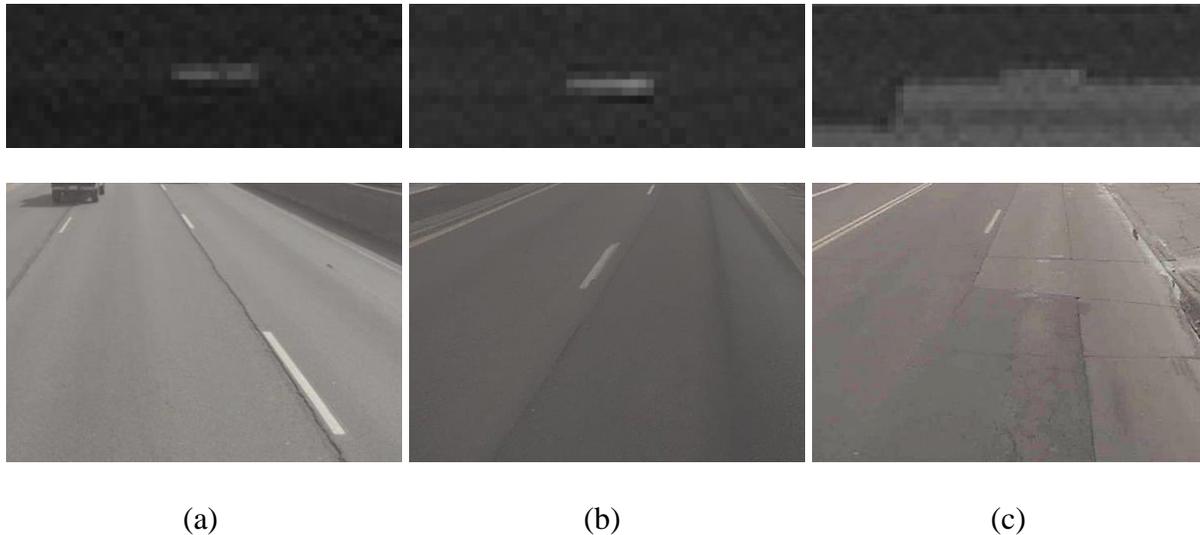


Figure 2.3 Panchromatic satellite images (top row) showing differences in pixel brightness for pavements in (a) good, (b) fair, and (c) poor condition. Ground-based digital images for the same sections are shown in the bottom row for comparison

2.4.2 Optical Satellite-Image Analysis

Pavement research studies have traditionally used spectral indices and texture metrics derived from multispectral imagery as the key parameters to characterize pavement condition (Mohammadi 2012; Pan et al. 2016). The following sub-sections explain the detailed process of extracting these parameters from multispectral imagery.

Spectral Indices

Spectral indices refer to the combinations of spectral reflectance from two or more wavelengths. Spectral indices in the form of simple reflectance, reflectance difference, ratio, and normalized

difference are widely used for detection and classification problems using multispectral and hyperspectral imagery. The sensor bands found in the WorldView3 imagery are coastal blue (CB), blue (B), green (G), yellow (Y), red (R), red edge (RE), near infrared 1 (NIR1), and near infrared 2 (NIR2). The wavelength of these bands' ranges from 400 to 1,040 nm. Since the oxidation process of the asphalt pavements and exposure of rocky components are typically displayed by the appearance of iron-oxide absorption features at wavelengths 520 and 870 nm (Shahi et al. 2015), pavement studies have usually focused on bands within these wavelengths (i.e., *B* and *NIR*).

Textural Metrics

Texture is defined as the spatial variation of greyscale levels in a neighborhood as a function of the resolution of the image. This spatial feature is commonly used in image analysis because it does not require prior image segmentation, as compared to other spatial features such as shape or size (Kupidura 2019). Including textural metrics, along with spectral information, has been found to significantly improve the accuracy of pixel-based classification problems (Bekkari et al. 2012). The process of texture analysis involves a kernel, which is a rectangular moving window with odd number of pixels in both directions of an image. The texture metric is computed based on the estimated probability of each of the pixel brightness values within the kernel. The calculated texture value is then assigned to the center pixel and the kernel moves one pixel over. The process is repeated until every pixel in the image has been assigned a texture value.

Two types of kernel-based textures, namely: first-order (occurrence) and second-order (co-occurrence) metrics, are typically used in analyzing satellite imagery (Warner 2011). First-order textures are estimated based on the counts of different digital number (DN) values irrespective of the location of pixels within the kernel. Commonly used occurrence metrics are mean, data range,

variance, entropy, and skewness. The mean is essentially a smoothing operation that accounts for the arithmetic average of the brightness values of a kernel. Data range and variance account for variability in kernel values, whereas entropy is calculated based on the distribution of the pixel values in the kernel. Skewness measures the symmetry around mean. These metrics are estimated based on Anys et al. (1994):

$$\text{Data range} = i_{max} - i_{min} \quad (2-5)$$

$$\text{Variance} = \sum_{i=0}^{N_g-1} (i - M)^2 P(i) \quad (2-6)$$

$$\text{Entropy} = \sum_{i=0}^{N_g-1} P(i) \times \ln P(i) \quad (2-7)$$

Where, $P(i)$ is the probability of occurrence of each pixel value i , M is the mean pixel value, N_g is the number of distinct grey levels in an image.

Second-order textures depend on the angular relationship and distance between two neighboring pixels. The co-occurrence matrix is estimated based on the counts of these pixel relationships. The second-order measures are estimated based on Haralick et al. (1973):

$$\text{Variance} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 \times P(i, j) \quad (2-8)$$

$$\text{Homogeneity} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} \times P(i, j) \quad (2-9)$$

$$\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j)(i - j)^2 \quad (2-10)$$

$$Dissimilarity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j) \times |i - j| \quad (2-11)$$

$$Entropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j) \log(P(i, j)) \quad (2-12)$$

$$Second\ Moment = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \{P(i, j)\}^2 \quad (2-13)$$

Where, $P(i, j)$ is the probability of occurrence of pixel value pairs i and j ; μ is the mean of P .

These texture metrics were extracted from pavements in the panchromatic image. A manual marking process was followed to exclude road markings, vehicles, trees, and other noises on pavement surfaces. Based on the definition of these texture metrics, we expect pavements in good condition to have higher homogeneity in pixel values and lower values of data range, variance, and entropy than pavements in poor condition.

2.4.3 Modeling POMDP

We consider a single-component system consisting of a one-mile pavement section. This section defines the 8-tuple $(S, A, T, \Omega, O, R, \gamma, b_0)$ needed to model a POMDP (Equation 2-1).

States of the System (S) and Initial Belief (b₀)

Pavement condition is described using three discrete states based on their International Roughness Index (IRI): (1) *Good*, (2) *Fair*, and (3) *Poor* as described in Table 2-1. The probability of the system initially being in any of these three states is assumed to be uniform.

Maintenance Actions (A) and Transition Probability Matrices (T)

The states were assumed to have stationary transition dynamics (i.e., the transition from one state to another is independent of time and rate of pavement deterioration). The available maintenance actions include: (1) *do nothing*, (2) *minor repair*, (3) *major repair*, and (4) *reconstruction*. Transition probability matrices (T) for each of these maintenance options are derived from the CDOT pavement condition dataset. These actions result in the following transition probability matrices:

$$T(1) = \begin{bmatrix} 0.74 & 0.25 & 0.01 \\ 0 & 0.82 & 0.18 \\ 0 & 0 & 1.00 \end{bmatrix}$$

$$T(2) = \begin{bmatrix} 0.86 & 0.14 & 0 \\ 0.62 & 0.35 & 0.03 \\ 0 & 0 & 1.00 \end{bmatrix}$$

$$T(3) = \begin{bmatrix} 0.94 & 0.06 & 0 \\ 0.86 & 0.13 & 0.01 \\ 0 & 0.50 & 0.50 \end{bmatrix}$$

$$T(4) = \begin{bmatrix} 1.00 & 0 & 0 \\ 1.00 & 0 & 0 \\ 1.00 & 0 & 0 \end{bmatrix}$$

Observations (Ω) and Emission Probability Matrices (\mathbf{O})

Three inspection choices are available: (1) *no inspection*, (2) *satellite-based inspection*, and (3) *annual distress surveys*. If *no inspection* strategy is chosen, then the same observation “unknown” will be obtained, regardless of the true state of the system. The observation matrix for this case will be:

$$O(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The observation matrix for *satellite-based inspections* is established based on histograms of pixel brightness for pavements in *good*, *fair*, and *poor* conditions, as observed in the panchromatic image. The *satellite-based inspection* is assumed to yield five levels of possible outcomes, corresponding to five levels of pixel brightness measured in terms of Digital Number (DN) values: (1) *very low*, (2) *low*, (3) *medium*, (4) *high*, and (5) *very high*. These levels are determined by dividing the range of DN values into five classes, as shown in Figure 2.4. Pixel brightness values lower than 1% of the pixels from the pavements in fair condition are considered *very low*, whereas the pixel brightness values greater than 99% of the pixels from the same curve are considered to be *very high*. The space between these two extremes is divided into three equal classes which are denoted as *low*, *medium*, and *high*.

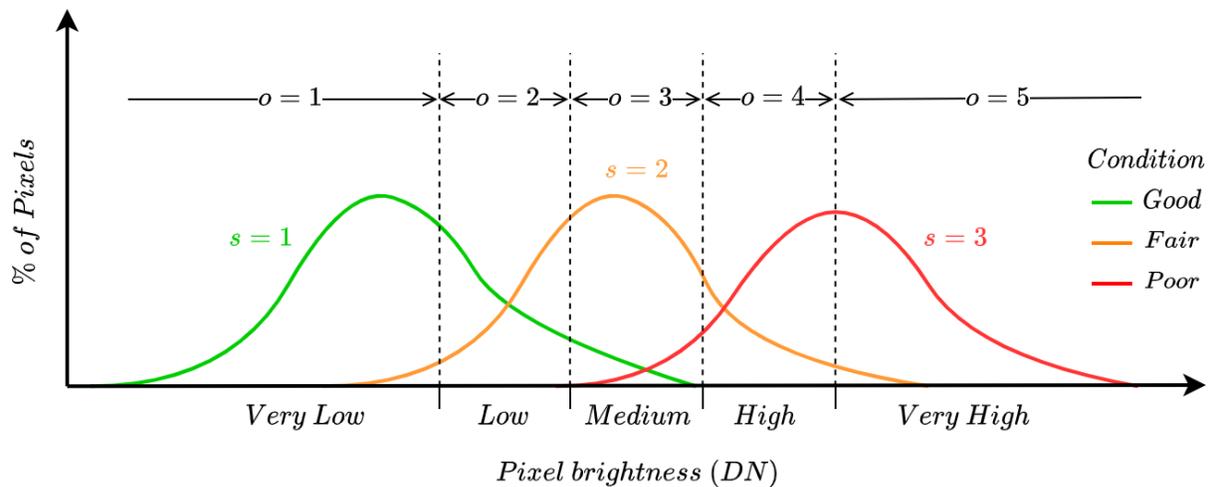


Figure 2.4 Histograms resulting from satellite observations are used to describe pavement states based on different levels of pixel brightness values

The observation probabilities O^{so} are estimated by calculating the area under the curve for each state (s) that belongs to a particular observation level (o). Therefore, the observation matrix for *satellite-based inspection* will have the structure depicted below. O_{12}^{s0} , for example, corresponds

to the area below the histogram of pavements in good condition ($s = 1$) within the range of low pixel brightness values ($o = 2$). A similar reasoning is used to estimate the remaining elements of the observation probability matrix for satellite-based inspections.

$$O(2) = \begin{bmatrix} O_{11}^{so} & O_{12}^{so} & O_{13}^{so} & O_{14}^{so} & O_{15}^{so} \\ O_{21}^{so} & O_{22}^{so} & O_{23}^{so} & O_{24}^{so} & O_{25}^{so} \\ O_{31}^{so} & O_{32}^{so} & O_{33}^{so} & O_{34}^{so} & O_{35}^{so} \end{bmatrix}$$

The *annual distress surveys* can be used to estimate *IRI* and therefore, would yield the same number of outcomes as that of the system states, resulting in this case in a 3 x 3 matrix. Although annual distress surveys are more accurate than satellite-based inspections, some level of inaccuracy can still be expected. A 90% accuracy is assigned to these ground-based observations (i.e., the probability of observing the correct state is 0.9) and the remaining probability is distributed over the other states.

$$O(3) = \begin{bmatrix} 0.90 & 0.08 & 0.02 \\ 0.05 & 0.90 & 0.05 \\ 0.02 & 0.08 & 0.90 \end{bmatrix}$$

Rewards (R) and discount factor (γ)

An analysis period of 25 years with a discount factor (γ) of 0.95 is considered to account for the time value of money. The proposed discount factor is aligned with typical values (i.e., 0.95-0.98) used for infrastructure projects (Andriotis et al. 2020). The costs (negative rewards) associated with different actions were derived from CDOT historical project costs (Table 2-2). While the costs of the inspection actions and reconstruction are independent of the condition of the pavement, the costs for minor and major repairs are assumed to increase by 1.5 and 2 times with each deteriorating level of condition. The costs of inspection actions are established based on the

average costs discussed in the Introduction section. The user/penalty reward was estimated based on the increase in vehicle operating cost that drivers pay annually due to poor conditions of a road with an annual average daily traffic (AADT) of 5,000 (Barnes and Langworthy 2003). No user/penalty cost was assumed when the pavement was in a state of good condition.

Table 2-2 Costs associated with different actions in 1000\$ per lane mile

Condition States	1	2	3	
Maintenance Rewards (r_m)	1: Do Nothing	0	0	0
	2: Minor Repair	-25	-38	-50
	3: Major Repair	-175	-263	-350
	4: Reconstruction	-1050	-1050	-1050
Inspection Rewards (r_i)	1: No Inspection	0	0	0
	2: Satellite-based Inspection	-0.03	-0.03	-0.03
	3: Annual Distress Survey	-0.1	-0.1	-0.1
User/Penalty Rewards (r_D)	0	-34	-48	

2.4.4 Accuracy of Satellite Observations

The emission probability matrix (O) defined for the POMDP model can further be used to characterize observational uncertainties associated with satellite images in classifying pavement

conditions. As defined by Andriotis et al. (2021), the following distribution of observation accuracy (p) can be used to study the effect of observational uncertainty, where $0 \leq p \leq 1$.

$$O'(2) = \begin{bmatrix} p & (1-p)/2 & (1-p)/2 & 0 & 0 \\ 0 & (1-p)/2 & p & (1-p)/2 & 0 \\ 0 & 0 & (1-p)/2 & (1-p)/2 & p \end{bmatrix}$$

The current level of accuracy of satellite observations is estimated by comparing $O(2)$ and $O'(2)$. This distribution is also used to perform a sensitivity analysis to evaluate how the value of satellite information changes with improvements in accuracy (i.e., reductions in uncertainty).

2.4.5 Value of Satellite-based Monitoring

Incorporating satellite data in the pavement maintenance decision-making would result in different outcomes based on how these data would complement the existing pavement monitoring system. The information obtained from satellite data, for example, will have a different value for roads which are not typically monitored, as compared to major highways which are inspected annually using costly and highly accurate distress surveys. Therefore, to quantify the Value of Satellite-based Monitoring (*VoSBM*) two extreme cases are considered. The first represents the current practice for major highways in the US, in which pavements are monitored annually using automated distress surveys. This case represents a scenario of high monitoring costs and satellite inspections are explored as a supplement of annual distress surveys. The second scenario represents current practice of local and ancillary roads, whose pavements are not monitored. This case represents a scenario of low (i.e., zero) monitoring costs and satellite inspections are explored as the only method of monitoring.

Case 1: Satellite Observations Supplementing Annual Distress Surveys

To estimate the value of satellite information when used in addition to annual distress surveys, two POMDP control settings are modeled and compared: $Setting_{opt}$ and $Setting_{perm}$. The former accounts for scenarios where the decision maker will have the option to select inspection actions from a set of available actions, whereas the latter reflects the current practice of performing mandatory annual inspections using automated distress surveys.

For $Setting_{opt}$, three inspection choices are available: (1) *no inspection*, (2) *satellite-based inspection*, and (3) *annual distress surveys*. These inspection options are modeled together with the four maintenance actions (i.e., *do nothing*, *minor repair*, *major repair*, and *reconstruction*) at belief points suggested by the POMDP solution. Based on the combined options of inspections and maintenance alternatives (i.e., three inspection and four maintenance alternatives), there are 12 possible actions an agent can take at each decision period in $Setting_{opt}$. However, *reconstruction and annual distress survey*, and *reconstruction and satellite-based inspection* actions are excluded from the analysis to reflect a practical approach, as the *reconstruction* action results in a good pavement condition with certainty and, therefore, inspections are redundant and not needed following a reconstruction.

For $Setting_{perm}$, *annual distress surveys* are conducted by default at every decision step. The same four maintenance actions are available for this setting (i.e., *do nothing*, *minor repair*, *major repair*, and *reconstruction*). Combining these maintenance alternatives with the mandatory inspection results in a total of 4 system actions available for the decision-maker in $Setting_{perm}$.

Quantifying the $VoSBM$ will allow transportation agencies to make informed decisions about investing in satellite-based monitoring systems as a low-cost inspection method. $VoSBM$ can be used as an objective metric reflecting the benefits of considering satellite data in the decision-support system (Andriotis et al. 2020; Papakonstantinou et al. 2019). As defined by Andriotis et al. (2021), the $VoSBM$ of this scenario can be estimated as the difference between the value functions of a system:

$$VoSBM_1 = V_{opt}^* - V_{perm}^* \quad (2-14)$$

Where, V_{perm}^* is the value function of a system with mandatory *annual distress surveys* and V_{opt}^* is the value function of the same system with the flexibility of choosing *satellite-based inspections*. This equation can be used to estimate $VoSBM$ at every possible belief point that the system may pass through during the planning horizon. A $VoSBM$ value lower than the cost of adopting a satellite-based monitoring system would indicate that there is no benefit in investing in a satellite-based monitoring system and in such case, maintenance scheduling based on annual distress survey measurements would be preferred (Andriotis et al. 2020).

Case 2: Satellite Observations being the Only Method of Inspection

To quantify the $VoSBM$ for roads which are not typically monitored, two more POMPD control settings are modeled and compared: $Setting_{blind}$ and $Setting_{sat}$. The *blind* setting accounts for a scenario in which roads are not inspected using any monitoring techniques. For this case, *no inspection* is the only inspection option available. This inspection is uninformative and is described by the emission probability matrix $O(1)$. The $Setting_{sat}$ accounts for scenarios when the decision-maker has access to satellite observations to make pavement maintenance decisions. In this setting, two inspection actions are available: (1) *no inspection*, and (2) *satellite-based inspection*. The

available maintenance actions for both settings are the same as in case 1 (i.e., *do nothing*, *minor repair*, *major repair*, and *reconstruction*). The $VoSBM$ for this case is estimated by comparing the value of optimal policies for $Setting_{sat}$ and $Setting_{blind}$.

$$VoSBM_2 = V_{sat}^* - V_{blind}^* \quad (2-15)$$

Where, V_{blind}^* is the value function of a system where the roads are not inspected and V_{sat}^* is the value function of the same system with the flexibility of choosing *satellite-based inspections*. If the $VoSBM$ is greater than the cost of adopting a satellite-based monitoring system, adopting a satellite-based monitoring system for the roads which are not typically monitored would result in lower maintenance and inspections costs over the life cycle of the roads.

2.4.6 Expected Life-Cycle Cost

The value of optimal policies, as described in the previous section, is determined by solving the POMDPs for an infinite horizon. In the context of pavement management, however, life-cycle costs are more relevant, as they provide meaningful insights for decision-making. Therefore, the total expected life-cycle inspection and maintenance costs for the pavement section described in the Methodology section was estimated by simulating the converged policies for both cases. 10,000 trajectories for a horizon of 25 years were simulated with a uniform probability of the pavement initially being on any of the three condition states. To estimate the feasibility of investments to adopt a satellite-based monitoring system with a certain degree of accuracy, we compared the mean expected life-cycle costs and their 95% confidence intervals for both the cases.

2.5 Results

2.5.1 Sensor Sensitivity to Pavement Conditions

The analysis of pixel brightness values, measured in terms of the DN, showed that the pavement sections with higher IRI values exhibited greater brightness values across all the bands, as compared to the pavements in fair and good condition (Figure 2.5a). Opposite trends were observed for sections in poor rutting conditions (Figure 2.5b). Rutting tends to make the road surfaces appear darker and therefore, pavements with high rutting showed lower brightness values. These trends show that pixel brightness can be useful in investigating individual pavement distresses. However, the contrasting behavior of pixel brightness values for different distresses makes the spectral indices approach limited in evaluating the overall condition of the section.

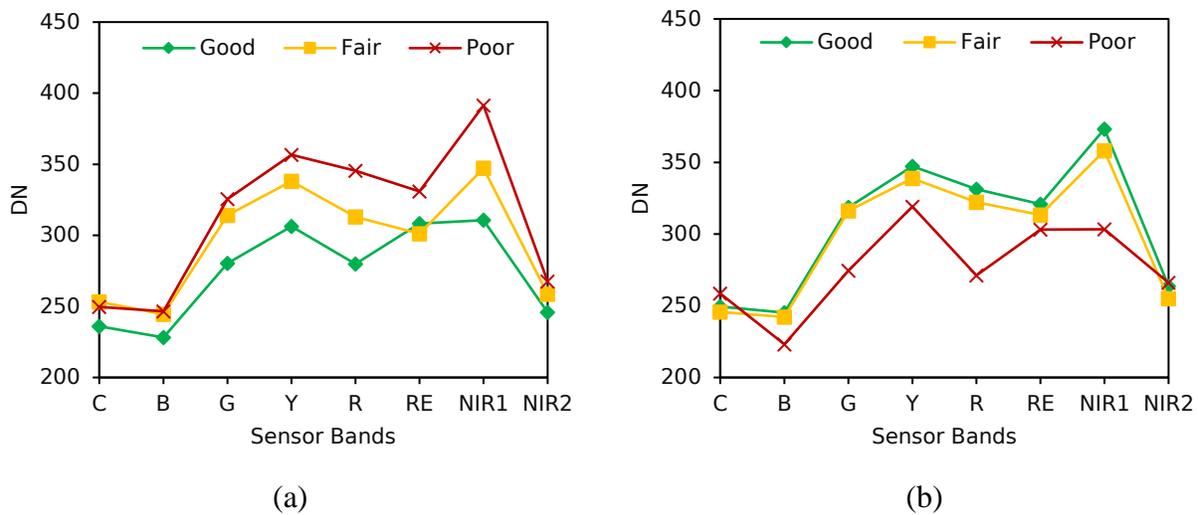
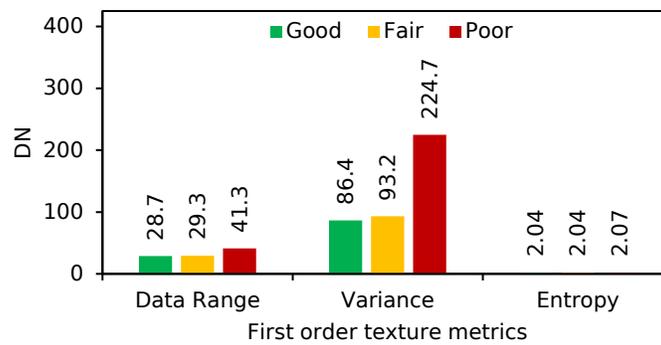


Figure 2.5 Average DN values across different bands for pavements, when they are classified based on (a) IRI, and (b) rutting

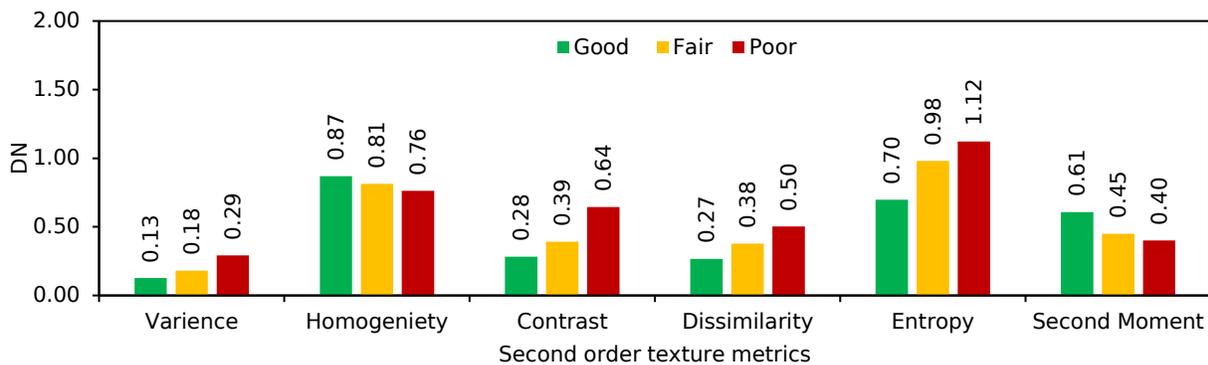
2.5.2 Texture Analysis

The results obtained from texture analysis indicated a strong correlation between the texture metrics and the deterioration of pavements. Both the first and second-order measures showed that

pixel brightness values increase with higher levels of pavement roughness. As the pavement degraded from a good to poor condition, the pixels became more heterogenous – the pixel values were distributed over a larger range (Figure 2.6a). The pixel brightness levels (i.e., DN) for the locations with high surface roughness showed a greater variance as compared to the pavements in fair to good condition. The co-occurrence-based measures confirmed these findings as a similar trend is observed from Figure 2.6(b). Variance, contrast, and entropy of pavements in poor condition were found to be higher, whereas the texture of pavements in good condition was more homogenous.



(a)



(b)

Figure 2.6 (a) First and (b) Second-order texture metrics for pavements in good, fair, and poor roughness condition.

2.5.3 Accuracy of Satellite Observations

A total of 42 sections with an equal number of good, fair, and poor pavements were analyzed to develop histograms of pixel brightness for pavements in different condition states (Figure 2.7). Ideally, as conceptualized in Figure 2.4, the means of these histograms are supposed to be apart from one another, so that the pixel brightness values can be used to characterize pavement condition. While this is true in Figure 2.7 for the pavements in ‘good’ and ‘fair’ condition – suggesting that pixel brightness can be used to infer information of pavement condition; the overlap of the histograms (specially the one of ‘poor’ pavements) demonstrates the high uncertainty of satellite-based observations.

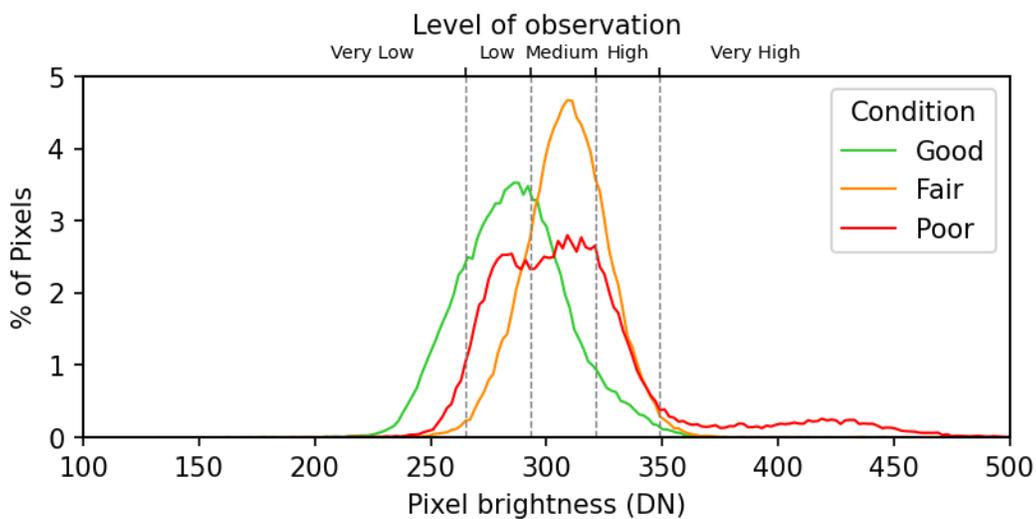


Figure 2.7 DN histograms for pavements in good, fair, and poor conditions and their corresponding levels of satellite observations (very low, low, medium, high, and very high)

The observation matrix for *satellite-based inspection* action was derived by calculating the areas under each curve for different outcome levels:

$$O(2) = \begin{bmatrix} 0.19 & 0.44 & 0.30 & 0.06 & 0.01 \\ 0.01 & 0.19 & 0.57 & 0.22 & 0.01 \\ 0.04 & 0.31 & 0.36 & 0.18 & 0.11 \end{bmatrix}$$

Comparing this observation matrix $O(2)$ with the synthetic observation matrix $O'(2)$, we observe that actual satellite observations are 19% accurate in classifying good pavements, 57% accurate for fair pavements, and 11% accurate for the pavements in poor condition. These numbers, however, do not truly represent the accuracy of satellite data itself, rather it reflects the approach that has been used to estimate this matrix. Changing the number of outcome levels would significantly increase the observation accuracy, as they would be aggregated over broader classes.

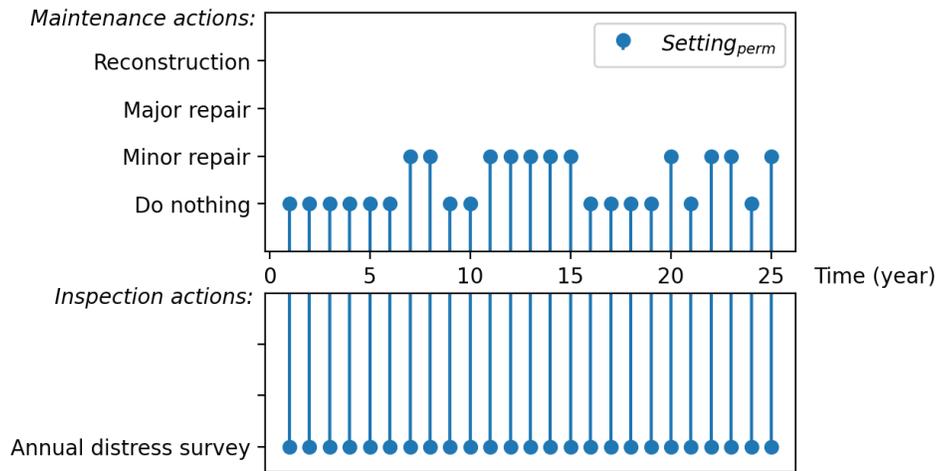
2.5.4 Evaluation of Optimal Policies

Case 1: Satellite Observations Supplementing Annual Distress Surveys

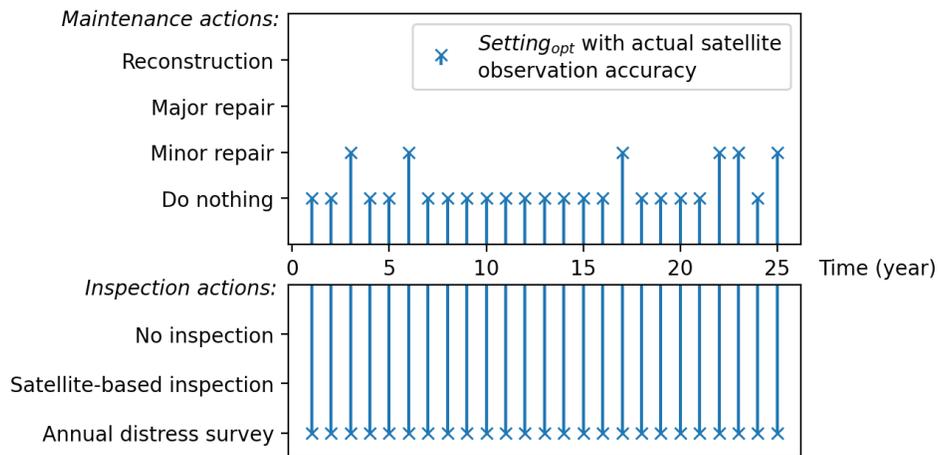
Realizations of converged optimal inspection and maintenance policies for both settings are shown in Figure 2.8. For both cases, the decision-makers adopt policies that minimize the total expected life-cycle inspection and maintenance cost of the respective systems. For *Setting_{perm}*, annual distress surveys are conducted every year (i.e., this result is obvious because this is the only inspection option available to the decision-maker). In terms of maintenance, optimal policies recommend no maintenance actions for the first 6 years (Figure 2.8a). After that, the decision-maker should perform minor repairs frequently to keep the pavement in good condition.

For the actual accuracy of the satellite inspections, derived from the histograms presented in Figure 2.7, the decision-maker should opt to keep using annual distress surveys at every time step, despite having access to other inspection actions for *Setting_{opt}* (Figure 2.8b). This is primarily because of the low accuracy of satellite inspections in differentiating between good, fair, and poor pavements (depicted in the overlapping histograms). The high uncertainty in satellite observations results in a policy similar to a situation where an agent takes random actions, which, in turn, results

in high life-cycle costs. As a result, in an optimal scenario, the decision-maker should decide to perform expensive but accurate distress surveys, to minimize the life cycle cost.



(a)



(b)

Figure 2.8 Policy realization for (a) $Setting_{perm}$, and (b) $Setting_{opt}$ with current satellite accuracy

To estimate the impact of satellite accuracy in optimal inspection strategies, additional cases were explored. The goal of these simulations was to define the required level of accuracy in satellite inspections for this technology to be competitive in optimal management strategies. This analysis

showed that optimal policies change drastically if the satellite inspections have higher accuracy. To demonstrate this, a policy realization with 80% accurate satellite inspections is shown in Figure 2.9. With less uncertainty, the decision-maker decides to perform satellite inspections at most of the time steps. The expensive annual distress survey is only chosen at year 11, when the decision-maker need more accurate information to return the pavement to a good state after doing minor repairs in two consecutive years. This also shows that minor repairs every 4-6 years based on satellite inspections are sufficient to keep the pavement in a good condition. For neither of the settings, however, the decision-maker chose a *major maintenance* or *reconstruction* action. This is reasonable considering the very high cost of these actions as compared to the *minor repairs* when the decision-maker is concerned about optimizing the cost over a life of only 25 years.

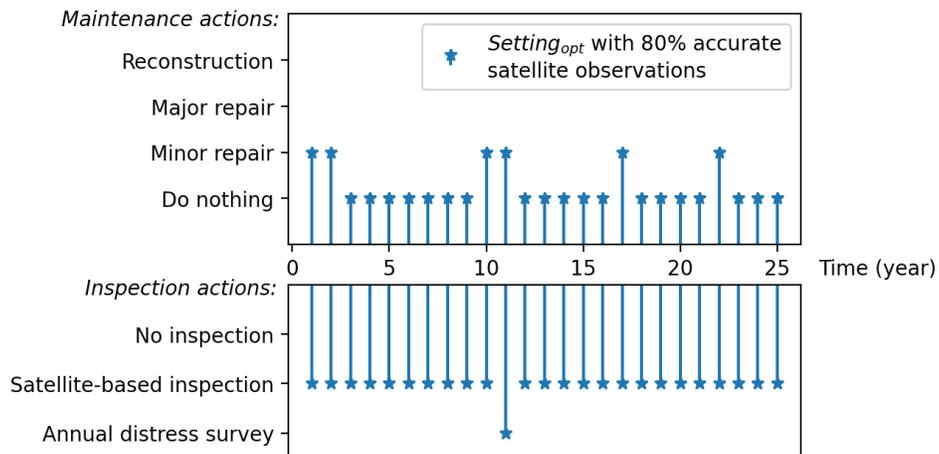


Figure 2.9 Policy realization for $Setting_{opt}$ with 80% accurate satellite inspections

The optimal maintenance and inspection action at a time step is governed by the belief about the system states at that time. The optimal policies as a function of the belief space are illustrated in Figure 2.10. The belief space is represented by an equilateral triangle and the belief at a certain time step is read using the smaller triangles by following the gridlines up to the corresponding

sides. The color of the circles inside the smaller triangles indicates the optimal action for corresponding beliefs about the state of a system. For the actual accuracy of satellite inspections, the optimal policy is dominated by the belief of a pavement being in fair condition (Figure 2.10a). When the observations result in about a 30-40% chance of the pavement being in fair condition, the decision-maker decides to do a minor repair to improve the condition of the pavement to a good state. A more practical policy is observed when the satellite inspections result in more accurate observations (Figure 2.10b). When the observations indicate the probability of the pavement being in a poor condition is less than about 20%, then the decision-maker decides to perform satellite inspections to save costs. With greater belief about the pavement being in a poor condition, indicated by the red and blue region, the decision-maker opts for more costly annual distress surveys so that they have more accurate information to make maintenance decisions.

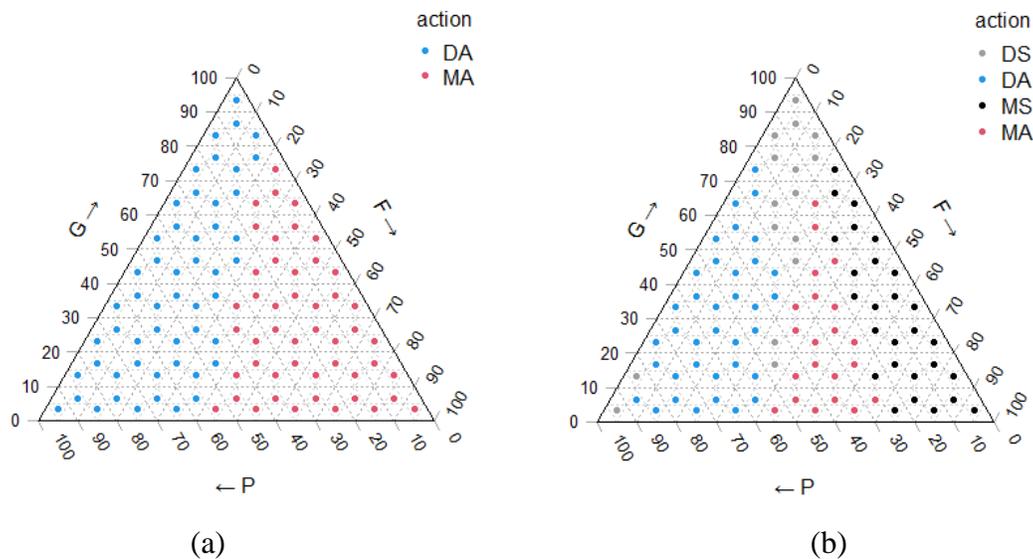
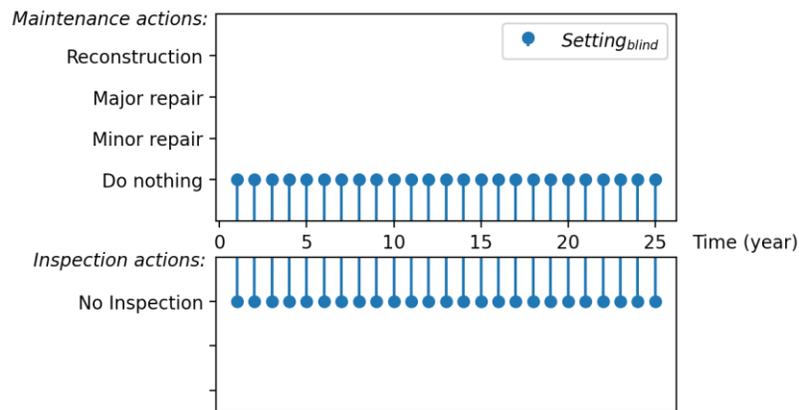


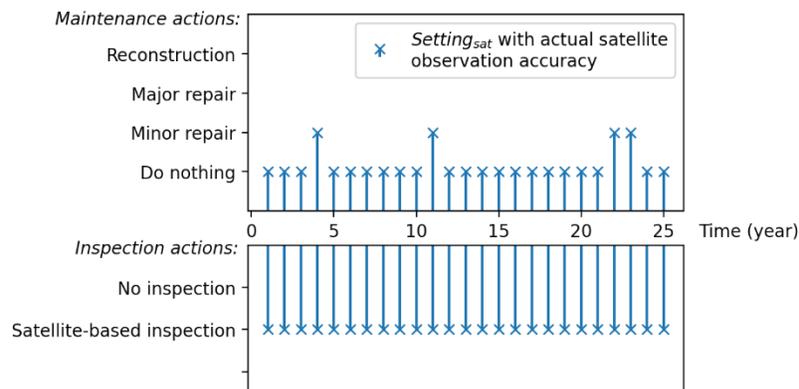
Figure 2.10 Policy as a function of belief about pavement condition for the optional inspection action setting with (a) actual, and (b) 80% satellite inspection accuracy, where D = Do nothing, M = Minor repair, A = Annual distress survey, and S = Satellite-based inspection

Case 2: Satellite Observations being the Only Method of Inspection

Optimal policies for case 2 settings are relatively less complex as compared to case 1. For $Setting_{blind}$, the decision-maker has no information about the condition of the roads and the optimal decision is to perform no maintenance actions over the service life (Figure 2.11a). When the satellite inspection system is available, the decision-maker decides to perform satellite inspections every year and use the data to optimally take *Minor Repair* actions at a regular interval to minimize life cycle costs (Figure 2.11b).



(a)



(b)

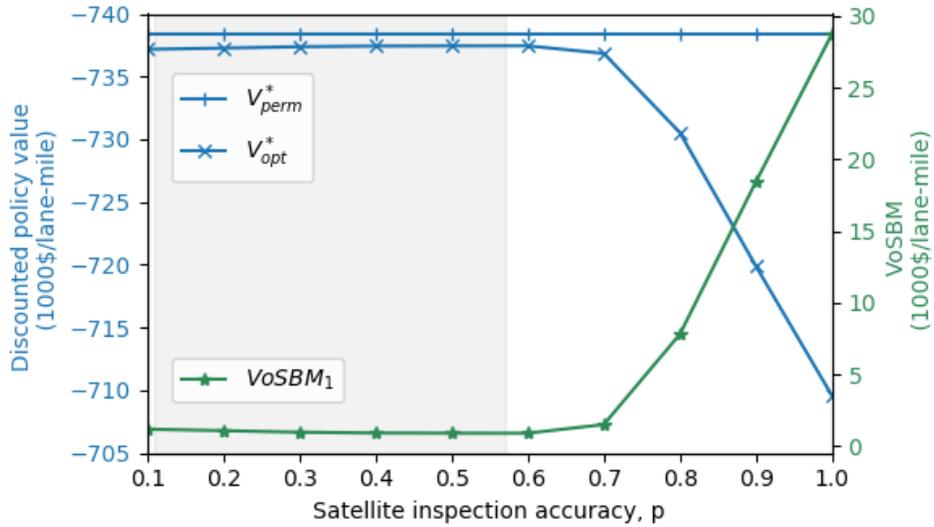
Figure 2.11 Policy realization for (a) $Setting_{blind}$, and (b) $Setting_{sat}$ with actual satellite observation accuracy

2.5.5 Value of Satellite-based Monitoring

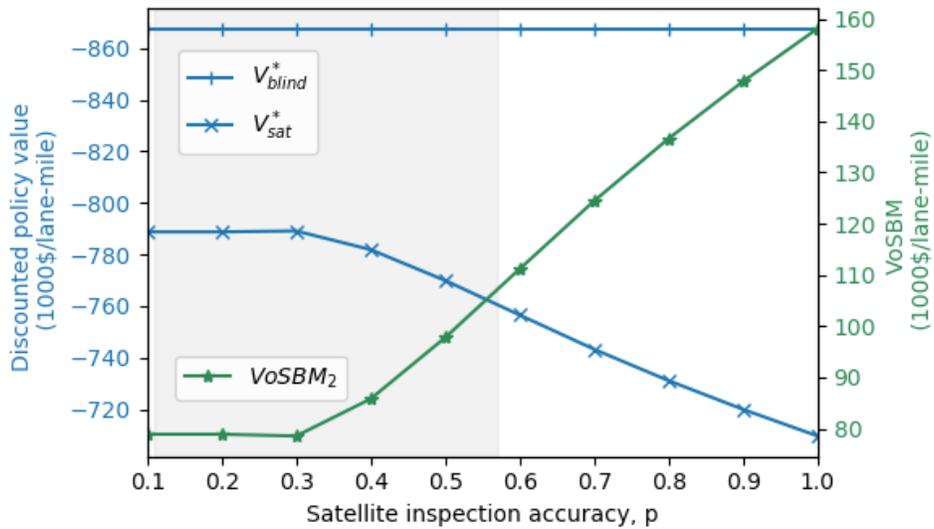
As discussed in the previous section, the accuracy of the satellite inspections significantly influences the choice of maintenance actions and thus the value of the policy. The $O'(2)$ observation matrix was used to evaluate the effect of satellite inspection accuracy (p) on the value of optimal policies for an infinite horizon. For case 1, the trend in total discounted value of the optimal policies for $Setting_{opt}$ indicate that information obtained from satellite inspections only become valuable when $p > 0.7$. Beyond this value, the increase in satellite inspection accuracy results in a significant reduction of the total life-cycle cost (V_{opt}^* in Figure 2.12a). The value function for $Setting_{perm}$ remain the same across all the levels of p , as it does not involve satellite inspections. Based on the assumption of 90% accurate annual distress surveys, as described in the Methodology section, the $VoSBM$ of the system is estimated to range from 0.2% to 4% of the value of $Setting_{opt}$ for $p = 0.7$ to $p = 1.0$ (i.e., perfect satellite inspections), respectively. In monetary terms, $VoSBM$ ranges from \$1,500 to \$28,774 per lane mile over an infinite planning horizon for case 1.

For case 2, satellite information adds value with as little as 30% accuracy, which is within the current range of accuracy (i.e., illustrated by the shaded region in Figure 2.12b). Satellite data for this case is significantly more valuable with an estimated $VoSBM$ ranging from \$79,000 to \$158,000 per lane mile, which is about 10 to 22% of the $Satellite_{sat}$ policy values. Since this is the value satellite inspections bring in the decision-making process, the lifetime cost of adopting a satellite-based monitoring system needs to be lower than these amounts for the system to be cost-effective. These amounts, extrapolated at a network level, will help transportation agencies decide

if establishing a satellite-based monitoring system with a certain level of associated uncertainty will result in better inspection and maintenance decision-making for their road network.



(a)



(b)

Figure 2.12 Optimal value functions for (a) case 1, and (b) case 2, and their corresponding values of satellite-based monitoring at different levels of satellite inspection accuracy, where the shaded region indicates the current range of accuracy

2.5.6 Expected Life-Cycle Cost

The simulation results, shown in Table 2-3, indicate that the total expected life-cycle cost are lower when satellite data is included in the decision-making process (i.e., *Setting_{opt}* and *Setting_{sat}*) as compared to the scenarios representing current practices (i.e., *Setting_{perm}* and *Setting_{blind}*). While the use of satellite data for case 1 results in about 0.75% reductions in life-cycle costs, these savings are significantly higher (i.e., 6.5%) for non-monitored roads (i.e., case 2). Simulation results for both cases also confirm that life-cycle cost can be further reduced with improvements in satellite inspection accuracy, as indicated by the costs estimated for a scenario with an 80% accuracy.

Table 2-3 Total expected life-cycle cost estimates with 95% confidence intervals based on 10,000 simulations for a 25-year service life

Case	Setting	Accuracy of Satellite Inspections	Expected Life-Cycle Cost (\$1,000/lane-mile)
	<i>Setting_{perm}</i>	-	-519.9 ± 4.1
1	<i>Setting_{opt}</i>	Actual	-516.0 ± 4.1
		80%	-514.5 ± 3.9
	<i>Setting_{blind}</i>	-	-601.6 ± 2.8
2	<i>Setting_{sat}</i>	Actual	-562.3 ± 3.4
		80%	-516.3 ± 3.9

The expected value of satellite-based monitoring systems based on these simulations indicate that adopting such systems for non-monitored roads (i.e., case 2) are economically more feasible as compared to roads currently monitored using annual distress surveys (i.e., case 1). The expected *VoSBM* is the maximum cost a transportation agency should plan to invest to cover the costs of acquiring, installing, operating, and maintaining a satellite-based monitoring system. Therefore, as illustrated in Figure 2.13, the transportation agencies can invest about \$35,000 to \$43,700 per lane-mile of road in acquiring satellite data at current level of accuracy for situations similar to case 2. For case 1, however, the expected *VoSBM* ranges from $-\$1,800$ to $\$9,000$ per lane-mile. A negative *VoSBM* indicates that setting up an additional satellite-based monitoring system may end up costing the transportation agencies more than the value it is supposed to bring in. Figure 2.13 also indicates that improving the system accuracy increases the value of satellite information significantly, allowing transportation agencies to invest more towards the adoption and development of the system.

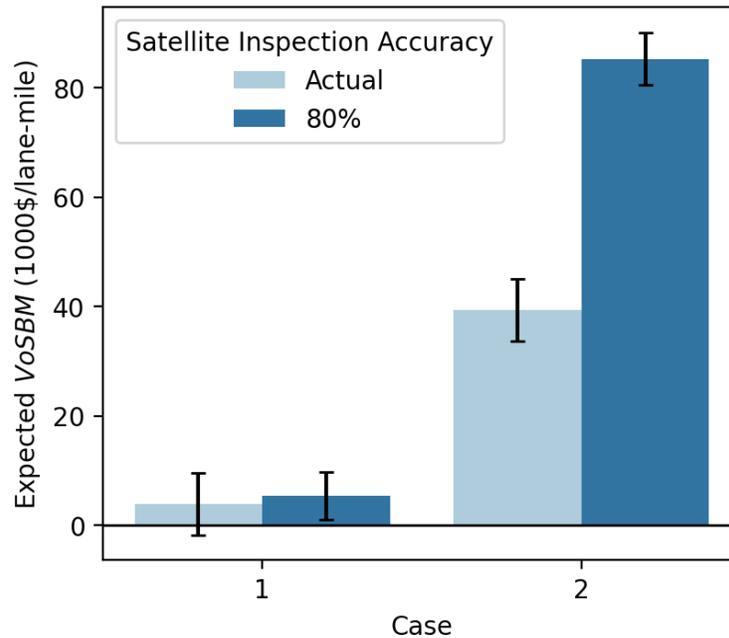


Figure 2.13 Expected value of satellite-based monitoring with 95% confidence intervals based on 10,000 simulations for a 25-year service life

2.6 Summary and Conclusions

This study evaluates the value of using optical satellite imagery in pavement inspection and maintenance decision-making within the context of POMDPs. Optimum life cycle policies were modeled considering stochastic controls, uncertain satellite data, actual field performance of pavements, and their maintenance costs. The effect of observational uncertainty on policy outcomes was also evaluated to derive recommendations on feasible investments to set up a satellite-based monitoring system. The major outcomes of this work are briefly outlined below:

- Rough surfaces show higher pixel brightness values as compared to the pavements in good and fair roughness condition. Pavements with poor rutting condition exhibit lower pixel brightness values across all the bands as rutting introduces darker shade to the pavement surface.

- Reflectance-based evaluation of road surface condition becomes complex under the presence of multiple distresses at the same time due to the contrasting behavior of pixel brightness values for different distresses. As a result, spectral indices are more useful in studying individual distresses as compared to an overall condition of a road surface.
- Pavements become lighter in panchromatic image as they degrade, with mean values of pixel brightness increasing with pavement roughness. They also become less uniform as they degrade, showing increased values of data range, variance, and entropy. All the texture metrics followed a logical relation with pavement condition, indicating strong evidence that satellite data has the potential to be used as a technique to investigate the deterioration of pavements.
- Satellite information is significantly more valuable for roads which are not typically monitored as compared to roads monitored using highly accurate distress surveys. For non-monitored roads, satellite data results in about 6.5% reductions in expected lifecycle cost of the pavements, as compared to 0.75% reductions for roads monitored annually with distress surveys.
- The lifecycle maintenance and inspection cost of the pavements are significantly influenced by observational uncertainty. Satellite inspections become valuable at 70% level of accuracy when used in combination with observations from 90% accurate distress surveys. Including less accurate satellite data does not influence the decision-makers' behavior, as it decides to perform more expensive but accurate distress surveys to reduce the life-cycle costs.
- The value of a satellite-based monitoring system was estimated to range from 0.2 to 4.0% of the total cost of a monitoring system when used for monitored highways, and 10 to 22%

when used for non-monitored (i.e., local and ancillary) roads. This outcome is critical for transportation agencies, as the investment needed to adopt a satellite-based inspection system should be lower than these amounts for the whole system to be cost-effective.

- Optimal policies simulated over a service life of 25 years indicate that satellite-based inspections at current accuracy may or may not (e.g., negative $VoSBM$) result in a cost-effective system, depending on the existing monitoring system of roads it is adopted for. It is, however, important to point out that this accuracy is not necessarily a limitation of the satellite data itself, but rather how it has been used to characterize the pavement condition states (e.g., the pixel brightness-based histogram approach). Therefore, improvements in satellite data processing and frameworks to estimate pavement condition using satellite data will add more value to these observations without increasing the cost and make the system more affordable and appealing to transportation authorities.

2.6.1 Limitations and Future Research

The research presented in this paper will advance the knowledge on the feasibility of using high-resolution optical satellite imagery in the management of pavement assets. This research, however, is limited by several factors. We assumed stationarity in deterioration transitions, while in reality, the transitions are non-stationary as the pavements deteriorate at a faster rate with age (Bashar et al. 2019). Also, characterizing pavement condition with 3 discrete condition states do not address the Markovian property of independence from history (Papakonstantinou and Shinozuka 2014b). Therefore, potential extensions of this work include modeling time dependent POMDPs where the number of states would be augmented by the combinations of pavement conditions and different deterioration rates. Improving the accuracy of the histogram-based approach by considering

different texture metrics instead of the individual pixel brightness values would also significantly improve the observation accuracy obtained from satellite inspections. Combining panchromatic and multispectral images to produce a pansharpened image will also allow us to leverage both high-resolution texture and spectral information in establishing a more accurate observation matrix.

CHAPTER 3. PERFORMANCE OF MACHINE LEARNING ALGORITHMS IN PREDICTING THE PAVEMENT INTERNATIONAL ROUGHNESS INDEX

3.1 Abstract

Significant research efforts have documented the capabilities of machine learning (ML) algorithms to model pavement performance. Several challenges, however, limit the implementation of ML by practitioners and transportation agencies. One of these challenges is related to the high variability in the performance of ML models as reported by different studies and the lack of quantitative evidence supporting the true effectiveness of these techniques. The objective of this paper is twofold: to assess the overall performance of traditional and ML techniques used to predict pavement condition, and to provide guidance on the optimal architecture and minimum sample size required to develop these models. This paper analyzes three ML algorithms commonly used to predict IRI: Artificial Neural Network (ANN), Random Forest (RF), Support Vector Machine (SVM), and compares their performance to traditional techniques. An inverse variance heterogeneity based meta-analysis is performed on twenty studies conducted between 2001 and 2020. The results indicate that ML algorithms capture on average 15.6% more variability than traditional techniques. RF is the most accurate technique with an overall performance value of 0.995. ANN is also identified as a highly effective technique that has widely been used and provides accurate predictions with both small and larger sample sizes. For ANN algorithms, a single hidden layer with nodes equal to 0.3 to 2 times the number of input features is found to be sufficient in predicting pavement deterioration. A minimum sample size equal to 50 times the number of input variables is recommend to model pavement deterioration using ML.

3.2 Introduction

Predicting pavement deterioration is an essential component of pavement management systems since it determines future pavement condition and maintenance needs. Pavement deterioration is typically measured using distress indices for cracking, rutting, patching, and roughness. Historically, Pavement Serviceability Index (PSI), Pavement Condition Index (PCI), Ride Quality Index (RQI) have also been used to model pavement performance. However, International Roughness Index (IRI) is the most well-recognized pavement performance indicator since it is used by transportation agencies throughout the world as a standard to measure road surface roughness (Du et al. 2014; Michigan Department of Transportation 2017). IRI measures the pavement surface deviations along the road that impacts vehicle suspension movement.

Pavement deterioration has traditionally been modeled using linear, non-linear, and multiple linear regression analysis, Markov chains, mechanistic-empirical, survivor curves, semi-Markov and Bayesian models (Abaza et al. 2001; Li et al. 1997; Li 2018; Osorio-Lird et al. 2018). These traditional approaches face certain challenges to account for all the factors affecting pavement deterioration and to handle pavement inventory information, characterized by imprecise, uncertain, ambiguous, subjective, and incomplete data (Flintsch and Chen 2004). On the other hand, transportation agencies are increasing their capacity to collect more data, though the capacity of traditional techniques to process and utilize this information is still limited. In the light of these challenges, Machine Learning (ML) is an appealing alternative to predict pavement deterioration as it offers significant productivity improvements over traditional techniques. ML is the science of making computers learn and act intelligently and improving their learning by feeding them data in the form of observations and real-world interactions (Faggella 2020). These algorithms can process

large volume of data with high degree of accuracy, handle noisy and complex data, tackle non-linear problems and once trained, can make predictions and generalizations in real time (Darko et al. 2020). ML excels in automation and pattern recognition and thus holds significant potential towards building a modern and robust pavement management system.

Although several research efforts in the last decade have established ML as a superior alternative to current techniques (Koch et al. 2015; Salehi and Burgueño 2018; Spencer et al. 2019), transportation agencies show reluctance in adopting these techniques due to their perception of ML being a black box, the challenges of integrating ML with the existing systems, and the lack of quantitative evidence supporting the true effectiveness of using these techniques (Abduljabbar et al. 2019; Flintsch and Chen 2004; Pell et al. 2015). Furthermore, a complete picture and understanding of the ML literature in pavement performance modeling is still missing. There exists a high variability in the performance of ML models as reported by previous studies. Therefore, it is difficult to quantitatively establish the true performance of these algorithms.

3.2.1 Objectives and Scope of the Study

The objective of this paper is twofold: to assess the overall performance of traditional and ML techniques used to predict pavement condition, and to provide guidance on the development of ML algorithms by identifying the optimal architecture and minimum sample size required for these models. In practical terms, this study serves as a guidance and state-of-the-art reference for practitioners, researchers, and highway agencies to assist them in using ML techniques as a tool to predict pavement performance.

The scope of this paper includes studies predicting IRI using ANN, RF, SVM and traditional techniques. Other pavement performance indicators and ML algorithms were not included because

of the limited literature available to perform a comparative analysis. The traditional models considered in this study include regression analysis (linear, non-linear, multiple linear, ridge random parameter, quadratic, partial least square), mechanistic-empirical, and sigmoid models. These models were selected from the studies implementing ML algorithms, as they were originally used to compare the capabilities of both techniques. This approach allows for a relatively fair comparison among the techniques' outcomes.

3.3 Overview of the ML Techniques used in Performance Modeling

In the recent past, significant research efforts have been shifted towards the use of ML algorithms for pavement deterioration modeling. In this section, we briefly introduce the most prominent ML algorithms used in pavement performance modeling and the metrics used to measure their performance.

3.3.1 Machine Learning Algorithms for Pavement Performance Modeling

Notable number of researchers have used Artificial Neural Network (ANN) to predict pavement performance (Abdelaziz et al. 2020; Bayrak et al. 2004; Chandra et al. 2013; Sollazzo et al. 2017; Yamany et al. 2020). Other ML algorithms used in modeling pavement performance include Decision Trees (Kang et al. 2010; Zeiada et al. 2020), Ensemble Trees (Inkoom et al. 2019; Rodriguez-Lozano et al. 2020), Random Forest (RF) (Gong et al. 2018; Marcelino et al. 2019, 2020), Support Vector Machine (SVM) (Georgiou et al. 2018; Kargah-Ostadi and Stoffels 2015; Zeiada et al. 2020; Ziari et al. 2016a), and Recurrent Neural Network (Choi and Do 2020). The description is focused on ANN, RF, and SVM since these algorithms are the ones most frequently

used in modeling IRI. The main characteristics of these algorithms and their architecture (defined by the model hyperparameters) are described in detail in the following subsections.

Artificial Neural Network (ANN)

Artificial Neural Networks are based on the idea of emulating the human brain and nervous system (Kobbacy 2012). The architecture of neural nets involve a network of interconnected neurons, arranged in layers, where the neurons receive and pass information to others via activation functions (Ceylan et al. 2014). Each of synaptic connections has its own numeric weight that can be adjusted based on experience, making the neural nets adaptive to inputs and capable of learning. The common hyperparameters for ANN include the number of hidden layers, number of neurons in each hidden layer, learning rate, number of epochs, momentum, activation function, number of epochs and batch size.

Random Forest (RF)

Random Forest algorithms aggregate the results from multiple decision trees, where the trees in the forest run in parallel without any interactions among them. This process ensures that the model is not heavily dependent on a particular feature. It also provides a better framework to prevent overfitting as compared to ANNs, since each tree uses a random sampling method to draw data from the original dataset while generating its splits (Chakure 2019). The major hyperparameters for a random forest algorithm include number of trees in the forest (`n_estimators`), maximum number of features considered for splitting a node (`max_features`), maximum number of levels in each decision tree (`max_depth`), minimum number of data points placed in a node before the node

is split (`min_samples_split`), and minimum number of data points allowed in a leaf node (`min_samples_leaf`).

Support Vector Machine (SVM)

Support Vector Machine algorithms identify optimum hyperplanes in a high-dimensional space which classify the data points. Regression is performed considering data points that are within the decision boundary lines around the hyperplane. These algorithms use different mathematical functions (i.e., kernels) to transform the input data to a required form. Typical kernel functions are linear, quadratic, cubic, sigmoid, and gaussian. These kernels often lead to a time-consuming optimization process when the nature of the data is unknown. To overcome this disadvantage, Üstün et al. (2006) introduced a universal Pearson VII function which resulted in a better generalization performance of the SVMs.

3.3.2 Algorithms' Performance Measures

The available success metrics for regression models include Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC), Mallor's C_p , adjusted R^2 , mean absolute error (MAE) and root mean square error (RMSE). Since R^2 and RMSE are the most commonly reported performance measures and are used in this study as measures of effectiveness. R^2 measures the variance in dependent variable explained by the independent variables, while RMSE measures the error in model's prediction.

3.4 Methodology

A comprehensive literature search was undertaken to identify studies predicting IRI with ML techniques. Studies using ANN, RF, and SVM algorithms and other study inclusion criteria,

detailed subsequently, were selected for a meta-analysis. A detailed analysis of the model architecture was carried out to identify optimal values of hyperparameters. The effect of sample size on the predictive performance was also investigated to derive recommendations on a minimum sample size.

3.4.1 Meta-Analysis

Meta-analysis consists of a statistical approach to combine quantitative research findings from multiple empirical studies. In meta-analysis, the effect sizes from different studies are combined to increase power and capture the true effect, allowing thus to summarize and compare empirical research studies. Meta-analysis has been applied extensively in medical research and several studies in the field of transportation and construction research have also implemented this methodology (Alruqi and Hallowell 2019; Elvik 2005; Papadimitriou and Theofilatos 2017). This study performs a meta-analysis to capture the true effectiveness of ML algorithms in modeling pavement deterioration. The meta-meta-analysis used in this research is based on Elvik (Elvik 2005) and Doi et al. (Doi et al. 2015). The process of conducting a meta-analysis includes a comprehensive review of literature, extracting the data of interest, standardizing effect sizes, and estimating the overall effect size (Alruqi and Hallowell 2019).

Systematic Literature Search

As part of the literature search, the authors referred to several databases to identify candidate studies: Transportation Research Record, American Society of Civil Engineers, Web of Science and ScienceDirect. The keywords used in the search process include *modeling/predicting pavement performance, machine learning, soft computing, roughness, IRI, artificial neural*

network, ANN, random forest, RF, support vector machine, and SVM. The following criteria were used to include a study in the analysis process:

1. Studies predicted IRI using ANN, RF, or SVM algorithms.
2. Studies included quantitative findings and statistical models reporting standard errors.
3. R^2 and RMSE were preferred over other statistical indicators.
4. Journal papers were preferred over conference papers. However, highly informative conference papers and reports were included in case of a shortage in specific model types.

Data extraction

Each of the papers retrieved in the literature search was studied in detail to identify the following information: algorithms used, algorithm architecture, sample size, unit used, testing, training, validation, and overall performance of the algorithms, and standard error of each estimate of performance. For the studies that reported the results of multiple trials with slightly modified architecture of the same algorithm, only the model with best performance was considered in this analysis.

Standardize Effect Size

The next step in the process consisted of extracting the effect size from individual studies (e.g., correlation values). Since the studies report a correlation among continuous variables, the correlation coefficient (r) was used as the effect size index (Borenstein et al. 2009). The R^2 values reported by the studies were used to estimate r for the corresponding models (with $r = \sqrt{R^2}$). For correlation-based effect sizes, meta-analyses are not directly performed on r itself, rather it is

converted to Fisher's z scale and all the analyses are performed using the transformed value. The correlation coefficient of j^{th} can be transformed to Fisher's z using the following equation:

$$z_j = 0.5 \times \ln\left(\frac{1 + r_j}{1 - r_j}\right) \quad (3-1)$$

The variance of z_j can be estimated as:

$$v_j = \frac{1}{n_j - 3} \quad (3-2)$$

Where, n is the sample size used in j^{th} study.

The standard error of Fisher's z is calculated as:

$$SE_j = \sqrt{v_j} \quad (3-3)$$

Estimate Overall Effect Size

In meta-analysis, the overall performance is typically estimated using either a fixed-effect (FE) or a random effects (RE) model. The FE model assumes that all the studies have one true effect size (i.e., all factors that could influence the correlation coefficient are the same in all the studies and therefore the true effect size is the same in all the studies). In this approach, each study is weighted (w_j) by the inverse of the square of standard error of the effect size to ensure that more accurate studies have greater impact on the overall effect size. The following formula is used to calculate the weights:

$$w_j = \frac{1}{SE_j^2} \quad (3-4)$$

The overall effect size is estimated as:

$$Z = \frac{\sum(w_j z_j)}{\sum w_j} \quad (3-5)$$

In many systemic reviews, including this study, the FE model assumption is implausible. In such cases a RE model is typically applied, as it accounts for the heterogeneity present among the studies. To determine heterogeneity among studies the following statistical test is performed:

$$Q = \sum w_j z_j^2 - \frac{(\sum w_j z_j)^2}{\sum w_j} \quad (3-6)$$

Where, Q is the chi-square heterogeneity statistic with $(J - 1)$ degrees of freedom.

If Q is found to be significant, the variance between studies is considerable. The percentage of variability (I^2) in effect estimates due to heterogeneity can be estimated using the following equation:

$$I^2 = \left[\frac{Q - (J - 1)}{Q} \right] \times 100\% \quad (3-7)$$

Where, J is the number of studies included in the analysis.

A random variance factor is included in Equation (3-4) to adjust the weights for RE model:

$$w_j = \frac{1}{SE_j^2 + \tau^2} \quad (3-8)$$

Where, τ^2 is the random variance of the heterogeneity and it is estimated as follows:

$$\tau^2 = \frac{Q - (J - 1)}{\sum w_j - \frac{\sum w_j^2}{\sum w_j}} \quad (3-9)$$

Doi et al. proposed an improved inverse variance heterogeneous (IVhet) meta-analysis approach, since the RE model is known to suffer from underestimating statistical error and spuriously

producing overconfident estimate intervals (Doi et al. 2015). The IVhet method uses a quasi-likelihood method to estimate the variance of the overall effect. The variance of overall effect (V) is estimated using Equation (3-10).

$$V = \sum_{j=1}^J \left[\left(\frac{\frac{1}{v_j}}{\sum \frac{1}{v_j}} \right)^2 (v_j + \tau^2) \right] \quad (3-10)$$

The 95% confidence intervals (CI) of the overall effect is estimated by:

$$Z \pm (1.96 \times \sqrt{V}) \quad (3-11)$$

Fisher's z is less frequently used in research and may increase difficulty in interpreting results (Alruqi and Hallowell 2019). Therefore, the Z and the associated confidence intervals are transformed back to r using Equation (3-12).

$$R = \frac{e^{2Z} - 1}{e^{2Z} + 1} \quad (3-12)$$

Where, R is the mean correlation coefficient across studies.

3.4.2 Analysis of Hyperparameters and Sample Size

Most of the ML learning algorithms have a set of hyperparameters that control the architecture and behavior of these models. The values of hyperparameters are set before the learning process begins, as these values cannot be directly trained from the data. This is a critical step in the modeling process since the optimum hyperparameters can significantly improve the quality of the model and help preventing overfitting. For example, if the learning rate of a neural network is too large, the performance of the model will oscillate over training epochs, and if it is set too small the model may never converge to its optimal solution (Brownlee 2019a).

Grid search and random search are commonly used approaches in optimizing hyperparameters, where all or a set of possible combination of hyperparameters are iteratively considered to find either the minimum (e.g. loss) or the maximum (e.g. accuracy) of a function. Since there exists no analytical formula to get an idea about these values, the hyperparameters of the selected models were studied in relation to the algorithms performance metrics to develop domain knowledge (i.e., an average setting which has resulted in a better performance).

Effect of Sample Size

ML models benefit from larger dataset since the models learn patterns from the data (Halevy et al. 2009). However, acquiring a great amount of data can be both expensive and time consuming. It is thus crucial to understand the effect of sample size in the algorithm performance. Although the sample size depends greatly on the complexity of the learning algorithm and the intricacy of the problem, using statistical heuristic methods such as a factor of the number of input features can be used to approximate the sample size. Using domain expertise or averaging the sample size over multiple similar studies is also a good approach to get an idea about the sample size (Brownlee 2019b). To further analyze this issue, this study analyzed the sample sizes used by the researchers to identify its effect on the models' accuracy and derive recommendations on the minimum sample size required in ML models.

3.5 Results

Twenty studies conducted between 2001 and 2020 are selected a detailed analysis (Table 3-1). Each study is coded by concatenating the last name of the first author and the year of publication. Multiple algorithms from a single study are coded with a number in parenthesis after their

corresponding study identification to avoid confusion in the analysis process. Most of these studies used Long-Term Pavement Performance (LTPP) to build their models, although field data from India, Greece, Taiwan, and Portugal are also observed. This variation in the data sources introduces a level of heterogeneity among the studies. Therefore, an inverse variance heterogeneity based meta-analysis is performed.

Table 3-1 Summary of the selected studies for meta-analysis.

Study Identification	Models Used	Pavement Type	Data Source	Reference
Abdelaziz 2020	ANN, LR	Flexible	LTPP	(Abdelaziz et al. 2020)
Bayrak 2004	ANN	Rigid	LTPP	(Bayrak et al. 2004)
Chandra 2013	ANN, NLR	Flexible	Field Data (India)	(Chandra et al. 2013)
Choi 2004	ANN, MLR	Flexible	LTPP	(Choi et al. 2004)
El-Hakim 2013	ANN, MEPDG	Rigid	LTPP	(Abd El-Hakim and El-Badawy 2013)
Georgiou 2018	ANN, SVM	-	Field Data (Greece)	(Georgiou et al. 2018)
Gong 2018	RF, RR	Flexible	LTPP	(Gong et al. 2018)
Hossain 2020	ANN	Rigid	LTPP	(Hossain et al. 2020)
Kargah-Ostadi 2015	ANN, SVM	Flexible	LTPP	(Kargah-Ostadi and Stoffels 2015)
Kaya 2020	ANN, SM	Flexible and Rigid	Iowa DOT PMIS	(Kaya et al. 2020)
Lin 2003	ANN	Flexible	Field Data (Taiwan)	(Lin et al. 2003)
Marcelino 2019	RF	-	LTPP	(Marcelino et al. 2019)

LTPP and Portuguese				
Marcelino 2019b	RF	Flexible	Road Administration Database	(Marcelino et al. 2020)
Mazari 2016	ANN	Flexible	LTPP	(Mazari and Rodriguez 2016)
Ozbay 2001	ANN	Flexible	LTPP	(Ozbay and Laub 2001)
Sollazzo 2017	ANN, LR	Flexible	LTPP	(Sollazzo et al. 2017)
Yamany 2020	ANN, LR, RPR	Flexible	LTPP	(Yamany et al. 2020)
Zeiyada 2020	ANN, SVM, LR, QLR, PLSR	Flexible	LTPP	(Zeiyada et al. 2020)
Ziari 2016	SVM	Flexible	LTPP	(Ziari et al. 2016a)
Ziari 2016b	ANN	Flexible	LTPP	(Ziari et al. 2016b)

Note: LR = Linear regression, NLR = Non-linear regression, MLR = Multiple linear regression, RR = Ridge regression, RPR = Random parameter regression, QLR = Quadratic linear regression, PLSR = Partial least square regression, MEPDG = Mechanistic-Empirical Pavement Design Guide, SM = Sigmoid Model

3.5.1 Overall Performance

The analysis resulted in an overall performance score in terms of correlation coefficient with 95% CI for each of the techniques. The results are presented in Figures 2.1 to 2.4 using forest plots, where the correlation coefficient is shown on x-axis and the studies are listed on y-axis. The size of the square boxes represents weights assigned to each of the studies. The diamond represents point estimate of the overall performance of each algorithm and its length shows associated 95% confidence intervals.

The overall performance of the traditional techniques in predicting the IRI is estimated to be 0.791 (95% CI = 0.595 – 0.898), as shown in Figure 3.1. Gong et al. (Gong et al. 2018) used 11,715

data points in their modeling process, which is significantly high as compared the other studies and thus carried a weight of 62.8% in determining the overall performance. A wider diamond shape ranging from 0.595 to 0.898 indicates that the traditional techniques across different studies resulted in highly varying accuracies.

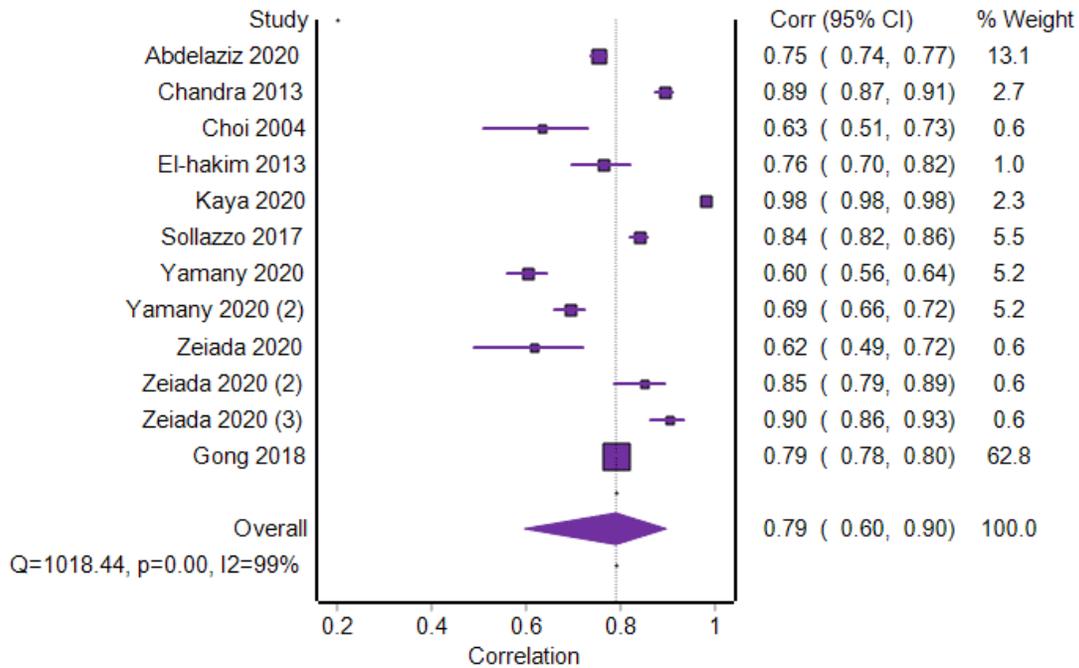


Figure 3.1 Forest plot illustrating the performance of traditional approaches.

The overall performance of the ANN models is 0.930 (95% *CI* = 0.863 – 0.946). Among the ANN studies it is worth mentioning that Choi et al. (Choi et al. 2004) and Lin et al. (Lin et al. 2003) used a small sample size (i.e., 117 and 125, respectively) and thus showed a high variance in the analysis (Figure 3.2). RF models resulted in a very high overall performance of 0.995 (95% *CI* = 0.981 – 0.999), with all the individual studies achieving a correlation coefficient greater than 0.96 (Figure 3.3). The accuracy of SVMs is estimated at 0.916 (95% *CI* = 0.672 – 0.981), which is slightly less than the performance of ANNs. However, SVMs show a higher

variance as compared to ANNs and RFs, since the studies exhibited a varying performance of SVMs ranging from 0.500 to 0.964 (Figure 3.4).

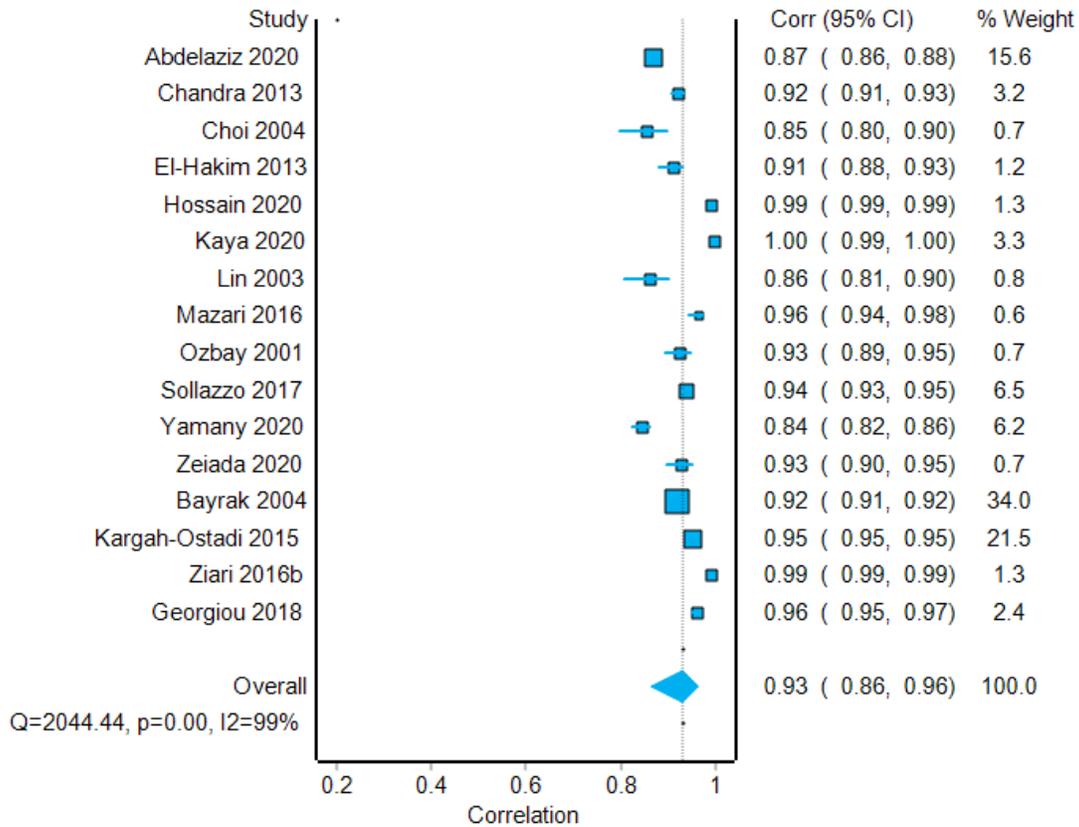


Figure 3.2 Forest plot illustrating the performance of ANN algorithms.

These results indicate that ML algorithms performed significantly better since they were able to explain, on average, 15.6% more variance in IRI than traditional techniques. RF is found to be the most accurate approach in predicting IRI, although the number of studies using this algorithm are considerably less compared to the other ML techniques. ANN is also highly effective and shows a consistent performance over a large number of studies.

The Q test was found to be significant ($p - value < 0.05$) for all the cases, indicating that considerable heterogeneity exists among the true performance estimations across the studies, thus

justifying the selection of inverse variance heterogeneity approach of meta-analysis. This heterogeneity stems primarily from the use of different data sources and the variations in the hyperparameters used to tune the algorithms.

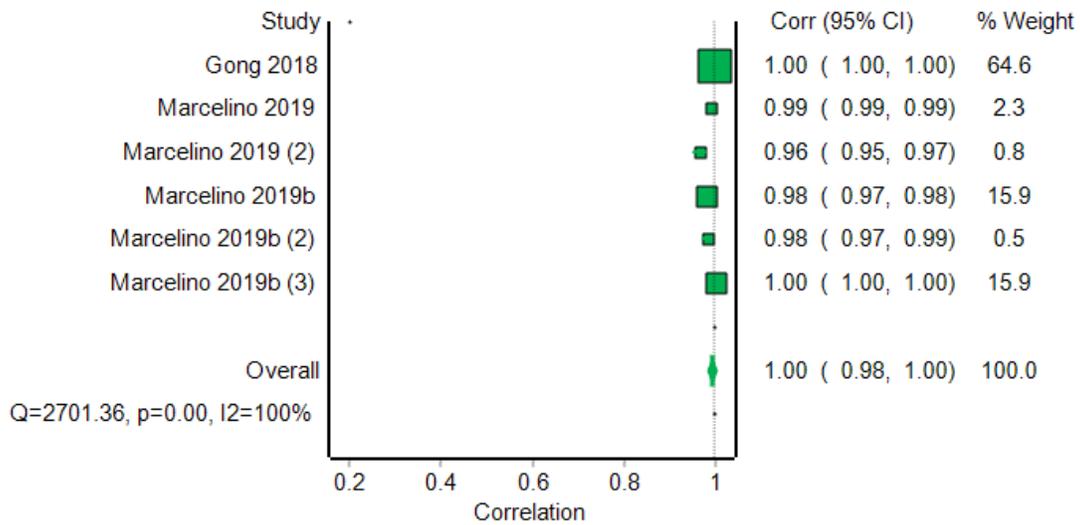


Figure 3.3 Forest plot illustrating the performance of RF algorithms.

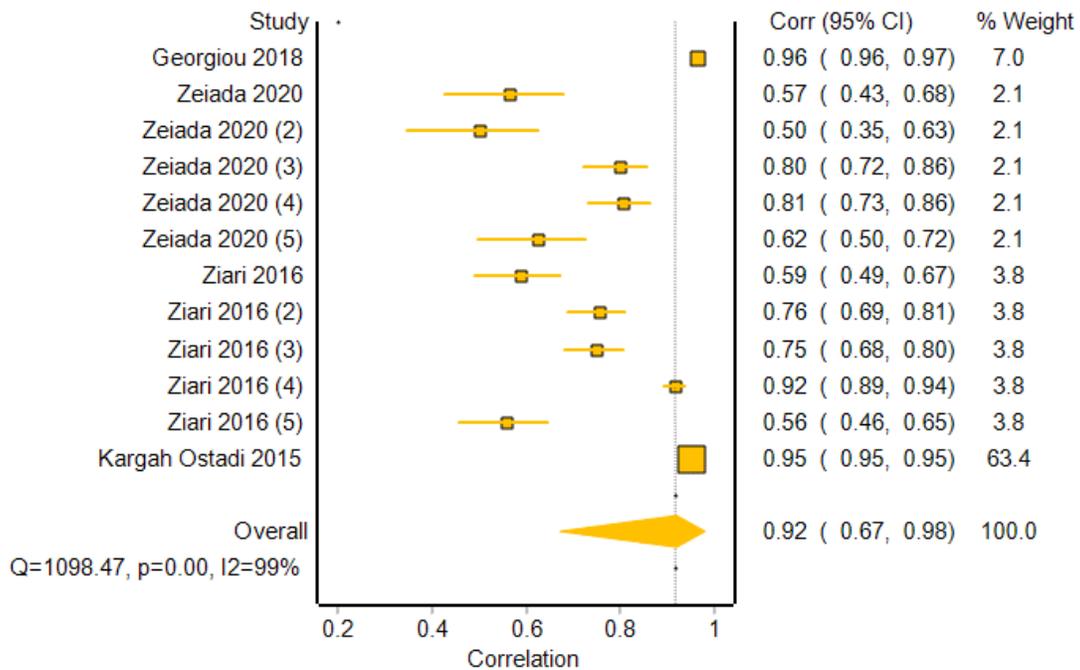


Figure 3.4 Forest plot illustrating the performance of SVM algorithms.

3.5.2 *Optimal Hyperparameters*

The following section provides guidance on the optimal architecture and minimal sample size required to develop accurate ML algorithms.

Artificial Neural Network

The reviewed studies showed consistency in some of the ANN hyperparameters. For example, most of the studies reported either a Tan-Sigmoid or a Log-Sigmoid activation function for their final model. The number of epochs varied from about 1,000 to 5,000. However, when it came to choosing the number of hidden layers and the number of neurons in these hidden layers, the studies used a wide variety of approaches. Therefore, these two items have been investigated in a greater detail.

Number of hidden layers

Existing literature indicates that most of the time a single hidden layer is sufficient enough to approximate any continuous function (Cybenko 1989; Hornik et al. 1989). Although some researchers (Flood and Kartam 1994; Ripley 2007) recommend more than one layer to model complex functions, studies (Chester 1990; Lapedes and Farber 1988) have concluded that two hidden layers are enough. A very similar pattern was observed in the reviewed models, since more than 72% of models reached a level of satisfactory accuracy with just a single layer of hidden layer as shown in Figure 3.5(a). About 17% of the models needed to use three layers to achieve a reasonable accuracy. However, using more than one hidden layers significantly increases both the training time and the chances of getting trapped in a local minima (Masters 1993). Based on the

reviewed models, it is recommended to use a single hidden layer to model pavement performance with ANN.

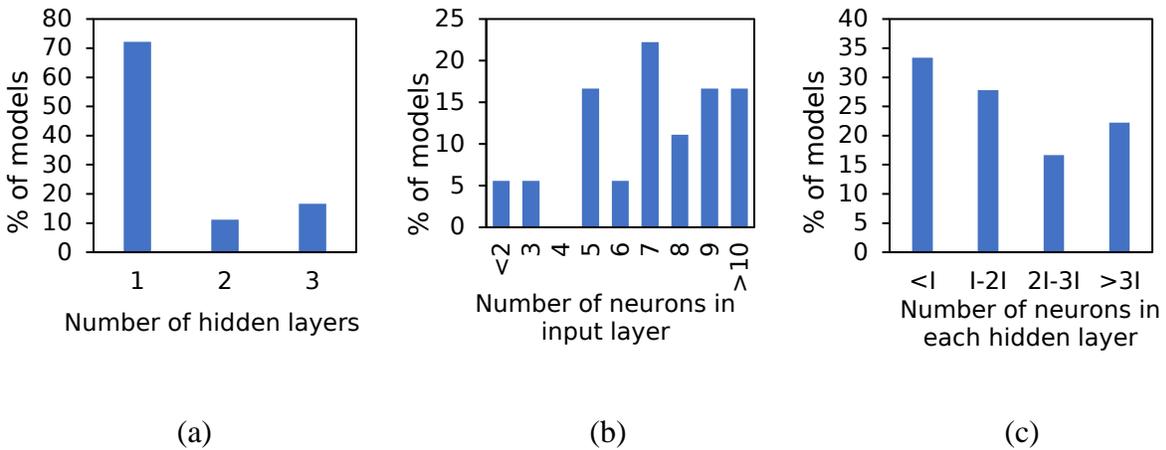


Figure 3.5 Percentage of ANN models with (a) number of hidden layers, (b) number of neurons in input layer, and (c) number of neurons in hidden layer as a function of input features.

Number of nodes in hidden layers

While the number of nodes in the input and output layers are determined by the number of input features and output parameters, there is no explicit approach to determine the number of nodes in a specific hidden layer. More than two-third of the reviewed ANN models used 7 or more features, (Figure 3.5(b)). This is significantly higher as compared to the traditional models where only one or two features were typically used (Kaya et al. 2020; Yamany et al. 2020).

For the number of nodes in hidden layers, an initial value of 75% of input features can be used (Salchenberger et al. 1992). An average value of the number of nodes in input and output layers can also be used to approximate the number of nodes in hidden layers (Berke and Hajela 1993). Studies suggest limiting the maximum number of neurons in hidden layer to $(2I + 1)$, where I is the number of nodes in input layer (Hecht-Nielsen 1987). Most of the reviewed models followed either of these three rules of thumb while iterating with different number of nodes. Figure 3.5(c)

shows the percentage of models reaching their best performance with the number of neurons in a hidden layer as a function of the number of input features (I). About 33% the models reached optimum accuracy with a smaller number of nodes than their input layer. For the 22% of the models with more than $3I$ neurons in the hidden layers, it was observed the iterations with $I/3$ neurons in the hidden layers produced accuracies very close to the optimum model (Ziari et al. 2016b). Based on the reviewed studies, starting with about 30% and then gradually increasing up to 2-times the number of neurons in hidden layer should significantly improve the performance of the model.

Random Forest

Validation curves are typically developed to identify initial values of these hyperparameters, although several rules of thumb are found in the literature. For example, Breiman (Breiman 2001) recommends to use the square root of the total number of predictors to estimate `max_features` for classification problems. While two thirds of the reviewed models used 100% of the predictors as `max_features`, Gong et al. (Gong et al. 2018) used a grid search with a 4-fold cross validation approach to identify `max_features` and `n_estimators` based on the resulting MSE values. Marcelino et al. (Marcelino et al. 2020) used the default `n_estimators` value of 10 and achieved very high R^2 values ranging from 0.954 to 0.995 for different models. The average of these optimum hyperparameters in the reviewed models are summarized in

Table 3-2, which can be used as an educated starting point in the tuning process of the RFs to reduce the optimization time.

Table 3-2 Summary of the optimum hyperparameters for the RF models.

Hyperparameter	Average	Range	
		Min	Max
n_estimators	355	10	800
max_features (as a % of the total features)	66%	7%	100%
max_depth	None*	None	15
min_samples_split	2*	2	10
min_samples_leaf	1*	1	2

Note: *Indicates Mode

Support Vector Machine

Observed hyperparameters for SVM regression models include kernel function K , parameters C and γ . The C parameter is responsible for adding penalty for each misclassified data point and γ parameter controls the distance of influence of training data points. Typical ranges of these hyperparameters are $0.1 < C < 100$ and $0.0001 < \gamma < 10$ (Yıldırım 2020). A k-fold cross-validation and trial-and-error approach is used to identify the specific optimal values depending on the application. Georgiou et al. (Georgiou et al. 2018) obtained the maximum predictive performance for their SVM models at a $C = 5$ and $\gamma = 0.01$. The performance analysis based on

kernel type indicated that Pearson VII Universal and Gaussian Kernels performed better as compared to the other kernel types in terms of both R^2 and RMSE values as shown in Figure 3.6.

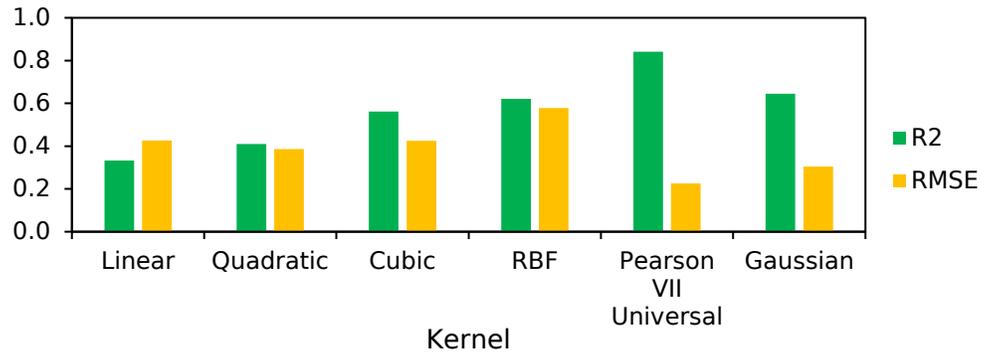


Figure 3.6 Predictive performance of SVM algorithms with different kernels.

3.5.3 Effect of Sample Size

Mazari et. al (Mazari and Rodriguez 2016) used only 95 samples to predict IRI using three input features: structural number, Equivalent Single Axle Load (ESAL), and age. They were able to achieve an overall R^2 of 0.927 and RMSE of 0.057 m /km. Choi et al. (Choi et al. 2004) and Ozbay et al. (Ozbay and Laub 2001) also reported similar performance with sample sizes lower than 120. However, studies with similar sample sizes trained with higher number of input variables have generally resulted in a poor performance (Zeiada et al. 2020; Ziari et al. 2016b). Figure 3.7(a) shows the sample size relations with the reported R^2 values. Although no strong correlation is observed, the range of R^2 values narrow down towards a higher accuracy as sample size increases. A similar pattern is also observed in RMSE values for sample sizes larger than 200 (Figure 3.7(b)). A detailed investigation of the models with higher accuracy ($R^2 > 0.9$) showed that they used approximately 25 to 75 times more training samples as compared to the number of input variables. Therefore, on average a minimum sample size equal to 50 times the number of input variables

should be sufficient. Most of the ANN models used a 70-15-15 ratio to split the dataset into training, testing, and validation sets. All the studies used an 80-20 train-test ratio for the RF algorithm, while for the SVM models the studies adopted either a 5 or 10-fold cross validation approach.

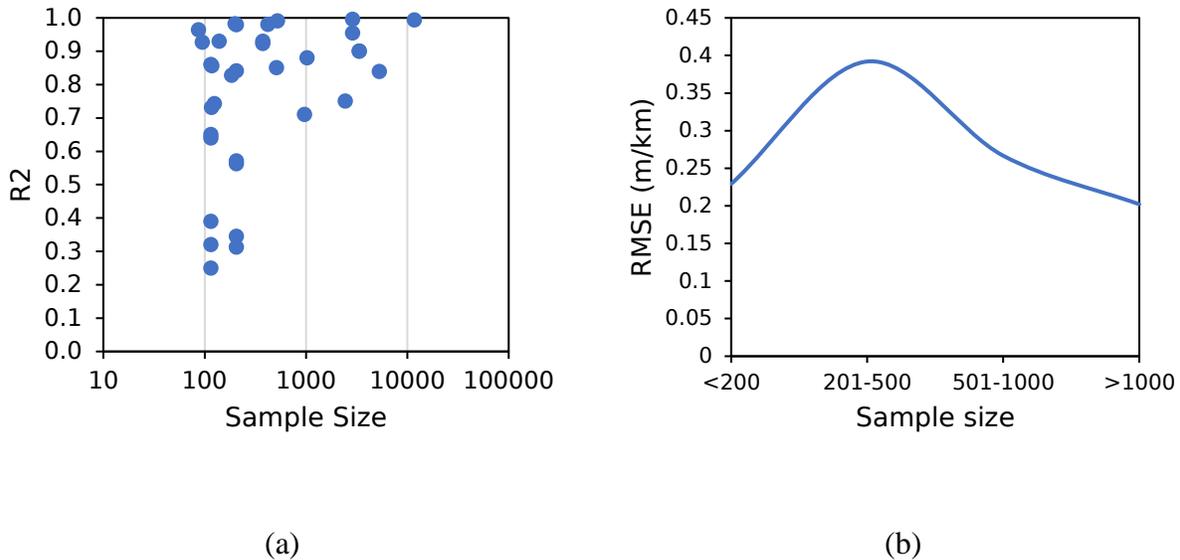


Figure 3.7 Sample size correlations ML algorithms' predictive performance.

3.6 Conclusions and Recommendations

This paper is aimed at evaluating the performance of ML algorithms in predicting pavement performance and providing guidance on the optimal model hyperparameters and minimum sample size required for these models. Twenty studies predicting IRI using ANN, RF, SVM and traditional techniques that satisfied study acceptance criteria were selected for a detailed analysis. An inverse variance heterogeneity based meta-analysis was performed to identify the overall performance of these algorithms as compared to traditional techniques. The notable findings of this research are summarized below:

3.6.1 Conclusions

- ML algorithms outperformed traditional regression-based approaches in predicting IRI. ML algorithms could capture on average 15.6% more variability in IRI than traditional techniques.
- RF was found to be the most accurate ML technique with an overall performance of 0.995. ANN and SVM were also found to be highly accurate in predicting IRI with performance values of 0.930 and 0.916 respectively, although the SVM models showed a very high degree of variability.
- Overall, we recommend ANN to model IRI, since its performance has been proven to be very accurate over a significant number of studies with both small and larger sample sizes. RF is more suitable when the user is suspecting the data to overfit or when a very degree of accuracy is desired.

3.6.2 Recommendations

- For ANN algorithms, using a single hidden layer with nodes equal to 0.3 to 2 times the number of input features should be sufficient in predicting pavement performance.
- The optimum hyperparameters for RF models were found to be $n_estimators = 355$, $max_features = 66\%$ of the total features, $min_samples_split = 2$ and $min_samples_leaf = 1$.
- For SVM models, use a Pearson VII universal of Gaussian kernel for higher accuracy.
- A minimum sample size equal to 50 times the number of input variables should be used for modeling pavement performance using a ML approach.

3.6.3 Limitations and Future Research

This study does not consider the effect of data pre-processing on the performance of the predictive models. Data pre-processing typically involves data cleaning, handling of missing values, feature sampling, dimensionality reduction, and data standardization which are not typically reported on the papers. These steps can significantly influence the predictive performance of the ML algorithms and are highly dependent upon the skills of a modeler. Since most of the studies only reported the best iteration of their model, the results may not reflect the true effect of sample size and hyperparameters on R^2 and RMSE values. Further research providing guidance on best practices on data pre-processing would be valuable for new practitioners and researchers.

The number of RF models included in the analysis were considerably less as compared to the other ML techniques that have received more attention from researchers (i.e., artificial neural network and support vector machine). Despite the limited number of publications reporting the use of RF, these models were found to be highly accurate in this meta-analysis. With RF models being a promising technique in predicting IRI, we may expect to see an increased application of these algorithms. It would be interesting to update this meta-analysis and analyze the accuracy of RF as it becomes established in the community and more relevant studies become available.

This study only focused on IRI prediction models. Future research is encouraged to analyze prediction models based on other indicators (e.g., overall condition indices such as Pavement Condition Index) as well as other aspects of the pavement management process such as the evaluation of maintenance alternatives and the optimization of maintenance programs. These processes may also benefit from the approach followed in this study, since similar algorithms have been used in previous research in these areas. An overall summary of these studies conducted

worldwide will help establishing ML as a superior alternative to the traditional techniques in the pavement management process. For future research, it is also suggested to analyze how this improvement in prediction accuracy may ultimately impact maintenance programs.

CHAPTER 4. A DEEP LEARNING FRAMEWORK TO ESTIMATE PAVEMENT ROUGHNESS USING SYNTHETIC APERTURE RADAR DATA

4.1 Abstract

Because of the high costs of ground-based pavement condition methods used to monitor pavement condition, transportation agencies often limit distress surveys to their major roads. As a result, the condition of local and ancillary roads remains unknown to decision-makers. This study addresses this gap by exploring the capabilities of publicly available Synthetic Aperture Radar (SAR) data to estimate pavement roughness. This paper introduces a novel framework to address the challenges of using SAR images in evaluating pavement condition. The trunk highway network in Minnesota is analyzed to develop deep learning models that predict International Roughness Index (IRI) and associated prediction intervals. This analysis found that SAR images have a strong potential in quantifying pavement condition. The deep learning models were able to predict IRI with a mean absolute error of 14.6 inches/miles and provide intervals of pavement condition that capture actual IRI values with an accuracy of 81%.

4.2 Introduction

Accurate and timely assessment of pavement condition is critical in the management of transportation infrastructure, as it determines maintenance needs and funding requirements. The transportation network in the United States comprises 3.9 million miles of built street, roads, and highways: 43% of which are in a poor or mediocre condition (ASCE 2021; TRIP 2018). While users demand more in terms of quality, safety, and accountability, the state Departments of Transportation (DOTs) are faced with challenges of aging pavements, deteriorating networks, and insufficient budgets to inspect and maintain such a large and complex network. Due to the high

costs of collecting pavement condition data using ground-based approaches, DOTs often limit their monitoring to the major roads of a network, as required by federal regulations (FHWA 2016). As a result, the condition of the ancillary components of a highway system such as ramps, auxiliary lanes, and frontage road pavements remain unknown to decision-makers. This raises the need for alternative solutions to monitor the condition of ancillary roads in a cost-effective manner.

Satellite remote sensing has the potential to provide pavement condition information that could complement the ground-based measurements and reduce monitoring costs. Past attempts in extracting road condition from remote sensors have mainly focused on optical satellite imagery (Mettas et al. 2015; Shahi et al. 2015). These approaches, however, are limited by the high cost of very high-resolution images, and the complications associated with processing optical images such as cloud covers, lighting, and weather conditions. Spaceborne Synthetic Aperture Radar (SAR) data effectively addresses these issues. Radar signals can penetrate clouds and image the whole earth during both day and night regardless of the weather condition. Moreover, C-band SAR data from Sentinel-1 satellite are available for public use at zero cost to the user. Previous studies have established SAR imagery to be successful in detecting changes in road surface with millimeter accuracy (Li et al. 2017). However, no studies so far have explored the potential of this publicly available bigdata in pavement monitoring. Indeed, the traditional computation techniques currently used in modeling pavement condition are ineffective in leveraging big datasets (Bashar and Torres-Machi 2021; Koch et al. 2015). With the flourishing of big-data applications, deep learning has emerged as a valuable tool for data-driven decision making in the management of infrastructure assets (Li et al. 2022; Tong et al. 2018). Deep learning algorithms constantly learn patterns from data and are highly effective in progressively extracting higher level features from complex

datasets using multiple layers of neurons. In this research, we aim to leverage the capabilities of deep learning algorithms to estimate pavement condition at a network level using state-of-the-art SAR technology.

4.2.1 Objectives

The primary objective of this study is to establish a framework to estimate pavement roughness using satellite-based SAR data and deep learning algorithms. To accomplish this, we first explored radar signal processing techniques to derive an optimal approach in processing SAR imagery for pavement condition evaluation purposes. Signals extracted from SAR imagery are then combined with relevant pavement features and modeled using deep learning algorithms to estimate pavement condition. The proposed framework was packaged as a software with a graphical user interface to facilitate its implementation by transportation agencies.

4.3 Challenges in Using SAR to Monitor Pavements

Radar technology, especially Ground Penetrating Radar (GPR), has been widely used for wide variety of pavement applications including modeling pavement deterioration (Batrakova et al. 2018), detecting subsurface cracks (Batrakov et al. 2021; Tong et al. 2017), moisture damage (Ma et al. 2021), measuring layer thicknesses (Al-Qadi and Lahouar 2005), and material density (Plati and Loizos 2013). Despite having a similar working principle, the use of SAR technology in pavement applications, however, is not well established. SAR sensors transmit microwave signals at a slanted angle and measure the backscattered signal to characterize features on earth surface (Munawar et al. 2021). Each pixel of the radar image is composed of phase and amplitude information. Phase indicates the distance between the sensor and the reflecting surface and is

number of practical challenges that are described below and addressed in the proposed framework. The first challenge is related to traffic noise, as pavement backscatters are greatly affected when vehicles and other objects are present on the road. In the presence of traffic (Figure 4.2), the SAR signal will suffer a double bounce effect and result in higher backscatter coefficients represented with brighter pixels. A smooth pavement may therefore appear brighter due to the presence of traffic, objects, trees, and tall buildings near the roads. Therefore, it is essential to filter out the reflected signals from traffic and other similar obstructions on or near the roads to accurately model road surface condition from SAR backscatters.

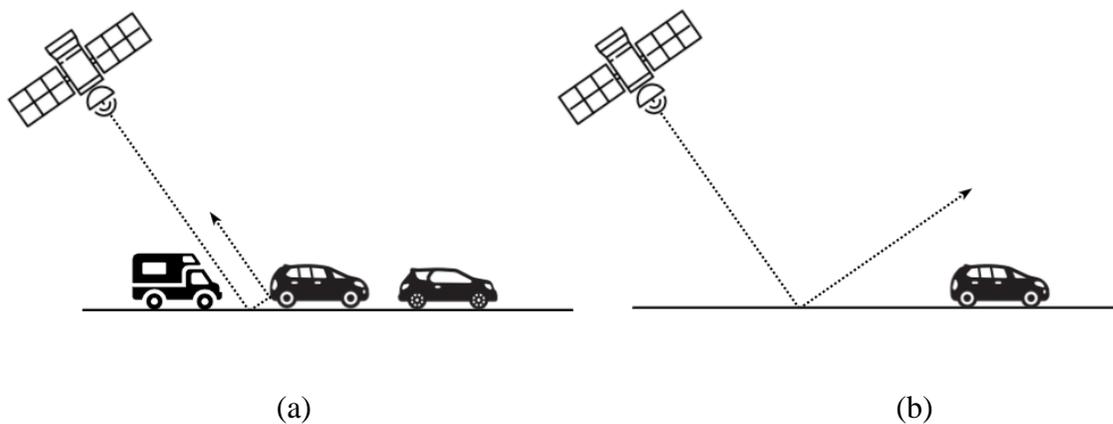


Figure 4.2 SAR backscattering in the (a) presence, and (b) absence of traffic. Image adapted from (Karimzadeh and Matsuoka 2020)

Similar to traffic noise, SAR images suffer from speckle noise when backscatters from different individual ground scatterers interfere with each other, resulting in either strong or weak return signals. This gives the SAR images a grainy appearance. To ensure accurate relationships between pavement condition and SAR responses, it is necessary to remove these speckles from SAR images. Lee filter is commonly used as an effective solution to suppress speckles in SAR images (Jaybhay and Shastri 2015). Lee filter, however, fails to preserve the edges and texture of the linear

features well, which are critical in roadway applications. While pavement related studies (Karimzadeh and Matsuoka 2020; Meyer et al. 2020) have applied several different filters to deal with speckles, the performance of these filters have not been evaluated quantitatively.

Also, there is no agreement on the most effective polarization of radar signals to capture pavement roughness. The polarization (i.e., orientation of the plane of oscillation) of a propagating signal affects how a signal interacts with an object on the ground. Since SAR has its own source of illumination, it can control the polarization of both the transmitted and backscattered signal. A vertical-vertical (VV) polarization indicates that the radar signals are transmitted and received vertically. Similarly, a vertical-horizontal (VH) polarization means the radar signals are transmitted vertically and received horizontally. Meyer et al. (Meyer et al. 2020) found VV polarization to be highly sensitive to rough surface scattering and recommended it for investigating roads and paved surfaces. Suanpaga and Yoshikazu (Suanpaga and Yoshikazu 2010), however, found HH polarization to be the most useful for modeling the International Roughness Index (IRI) of pavements.

Furthermore, the terrain contained in the pre-processed SAR images introduce geometric distortions due to the side-looking imaging technique of SAR systems. This results in over and under exposed pixels creating a barrier in correlating backscatter strengths to condition of the pavements located in different terrains. To address these challenges, we propose a structured approach that effectively improves post-processing of SAR images for pavement applications.

4.4 Proposed Framework

This paper introduces a novel framework to leverage SAR imagery and deep learning in estimating pavement roughness. The proposed framework (summarized in Figure 4.3) provides a process that

improves the standard SAR data processing method (Meyer 2019) to better address the issues associated with using SAR to monitor pavements. Our framework provides guidance on the polarization channel that should be used to capture pavement roughness, which filters should be applied to remove speckles without compromising the linear road features, and how to remove traffic noise and the effect of terrain to accurately model pavement condition from SAR backscatters. Once these processes are completed, data is modeled using deep learning algorithms and results in a predictive tool that is developed, tested, and ultimately deployed. The critical components of the proposed framework are discussed in detail in the subsequent sections.

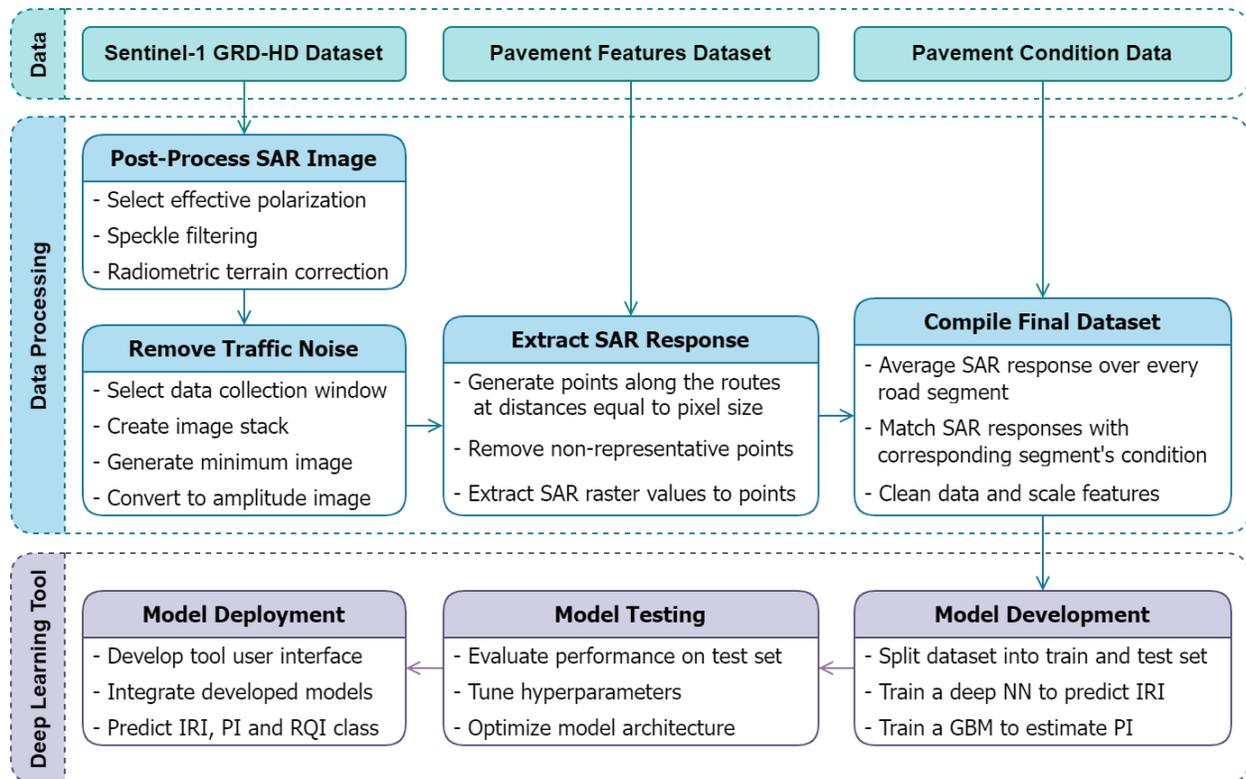


Figure 4.3 Proposed framework to estimate pavement condition using SAR imagery

4.4.1 Data Processing

Post-Process SAR Image

The proposed framework leverages SAR imagery captured by the Sentinel-1 satellite and, more specifically, the pre-processed Level-1 ground range detected high resolution dataset acquired from the Alaska Satellite Facility (“Copernicus Sentinel data. Retrieved from ASF DAAC. Processed by ESA” 2019). The acquired imagery typically have geometric and radiometric distortions due to the oblique observation geometry. These data, therefore, requires post-processing before they can be analyzed in a geographic information system (GIS) environment. Standard routine in post-processing these data include applying precise orbit file, radiometric calibration, speckle filter, radiometric terrain flattening, and geometric terrain correction. In this paper, we recommend a standard post-processing routine for pavement applications. Readers interested in a more detailed review of these processes can refer to (Meyer 2019).

Select Effective Polarization

Radar sensors typically collect data in multiple polarizations. The backscatters received for the same object on the ground varies based on the polarization channel of a sensor. Therefore, using the image captured in a polarization that is more sensitive to pavement roughness is of utmost importance in modeling IRI using SAR backscatters. Given the lack of agreement on what polarization channel is more effective for pavement applications, the first step of the proposed framework is to explore the suitability of Sentinel-1 polarization channels (i.e., VV and VH). SAR responses along the roads from both the VV and VH images were compared against their corresponding levels of roughness to quantify the ability of these channels at capturing differences in pavement condition.

Speckle Filtering

To remove speckles, especially from the pavement pixels, six different adaptive filters were considered in this study: Lee, Frost, Gamma-map, Intensity Driven Adaptive Neighborhood (IDAN), Refined Lee, and Lee Sigma. The goal of this analysis is to identify the filter that is most effective in suppressing speckles from pavement pixels while preserving the sharpness of edges and linear road features. The effectiveness of these filters was assessed using the following metrics:

- Speckle Noise Index (SNI): This index measures the intensity of speckle noise in an image. Lower SNI values indicate better speckle noise suppression. SNI is defined as follows (Crimmins 1985):

$$SNI = \frac{\sigma}{\mu} \quad (4-1)$$

Where, μ and σ are the mean and standard deviation of the filtered image.

- Equivalent Number of Looks (ENL): To smooth out noises, ground range detected (i.e., phase information removed) SAR images are subject to multi-looking (i.e., averaging the intensity of neighboring pixels) during the pre-processing. This concept of multi-looking was used to coin the term Equivalent Number of Looks (ENL), which is a measure of the degree of speckle suppression in post-processing. While ENL is similar to SNI, the second power in the formulation is useful in differentiating among similarly performing filters. Higher ENL indicates greater speckle suppression at the expense of edges and texture information. The choice of an ideal filter is, therefore, a compromise between noise removal and details preservation. ENL is estimated as (Parrilli et al. 2012):

$$ENL = \left(\frac{\mu}{\sigma}\right)^2 \quad (4-2)$$

- Normalized Mean (NM): This metric is used to evaluate if a filter results in an unbiased estimate. It is estimated as follows (Oliver and Quegan 1997), with NM values close to 1 indicating that the original information was perfectly preserved (Guo et al. 2018).

$$NM = \frac{\mu_{filtered}}{\mu_{original}} \quad (4-3)$$

Where, $\mu_{filtered}$ and $\mu_{original}$ is the mean of the pixel values before and after filtering the image.

Radiometric Terrain Correction

Each pixel of a Level-1 pre-processed SAR image essentially indicates the value of a backscatter coefficient (σ_0) resulting from the measured return signals. As a result, this image is often referred to as a Sigma Naught image. This image, however, suffers from the effect of topography, resulting in misleading σ_0 values for locations where the signals are affected by an uneven terrain. Rather than capturing straight-down, the SAR sensors use a side-looking imaging technique which causes geometric distortions leading to geolocation errors. This worsens in the presence of slopes, resulting in deceptive σ_0 . Since the proposed framework is based on measures of SAR amplitude (i.e., strength of the backscatter), it is critical to apply radiometric terrain correction to ensure accurate measurement of backscatters. Radiometric terrain correction refers to the process of removing the influence of topography from SAR images. This process moves the SAR pixels into correct spatial relationship to each other and the corrected backscatter coefficients are denoted by γ_0 . Therefore, the resulting image is referred to as a Gamma Naught image, where each pixel of the image indicates the value of corrected backscatter coefficient γ_0 .

Remove Traffic Noise

To remove traffic or any other temporary noise from the pavement pixels, the framework recommends an image stacking solution. With this approach, multiple images collected within a time window are bundled together. The stack is then used to generate a minimum intensity projection image where each pixel intensity is the minimum of all the pixels at that location across all the images in the stack. Traffic or other temporary objects on road create stronger backscatter (i.e., brighter pixels). Since the minimum intensity projection filters out the brighter spots which are not present in all the images, the temporary noises are removed while the brighter signals from permanent objects are preserved as they are similarly bright in all the images of the stack. Including a large number of images in the stack would increase the probability of filtering out heavy traffic noise. Given the proposed stacking solution requires a time window for image acquisition, a seasonal variability analysis of SAR responses was performed to derive recommendations on how to select this time window for a specific region. An example of this method applied to the pavements in Minnesota is described in the ‘Case Study’ section.

Extract SAR Responses

To extract backscatters from SAR images along the roads, a road network shapefile is first created based on the location information stored in the pavement features dataset. Then, reference points are generated along the road lines at a distance equal to the size of a pixel (i.e., spatial resolution) as illustrated in Figure 4.4(a) with a satellite image in the background. These reference points are carefully reviewed to remove any points where the backscatters are not representative of the pavement condition. For example, traffic signals, signposts, overpasses, or any other visible objects on or near the road are not included in the extraction, as they cause double bounce scatters

and result in stronger backscatters. An example of this is shown by overlaying the reference points on top of a SAR image in Figure 4.4(b), where an overpass causes significantly higher backscatters that result in a high pixel value (i.e., bright pixels). The final reference points are then used to extract γ_0 values along the roads. Pavement conditions are typically reported every 0.1 mile, extracted γ_0 values are, therefore, averaged over every 0.1 mile.

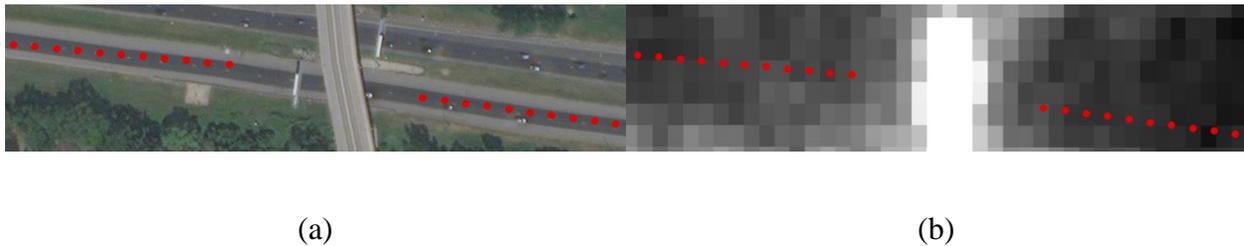


Figure 4.4: Road reference points overlaid on top of (a) satellite, and (b) processed SAR image

Compile Final Dataset

The average γ_0 values are then labeled with the IRI for corresponding sections. Additional features of these sections such as surface type (i.e., concrete or asphalt), age (i.e., measured as number of years since last major maintenance or construction), thickness of the surface layer, thickness of the base layer, and average annual daily traffic (AADT) are included as pavement features in the final dataset.

4.4.2 Deep Learning Tool

Model Development

To leverage the improvements resulted from the proposed framework a Deep Neural Network model is developed to estimate pavement IRI from the processed SAR imagery. To account for the uncertainties associated with the point predictions of IRI, a Gradient Boosting Machine model

is also developed. The Gradient Boosting Machine model is used to estimate prediction intervals for corresponding estimations of IRI from the Deep Neural Network model. For both models, the dataset is split into 80% for training and 20% for testing.

IRI Prediction

The Keras API with Tensorflow backend is used to define a sequential Deep Neural Network model which uses the feedforward backpropagation algorithm to learn from the training samples. The input layer consisted of 6 neurons with 1 neuron in the output layer. A normalization layer is added before the input layer to scale the features for efficient computation. Several different combinations of number of hidden layers, number of neurons in each hidden layers, and activation functions are tested to identify the optimum model architecture. Adam optimizer with a decaying learning rate starting from 0.001 is used to train the model to facilitate both better optimization and generalization. To prevent the model from overfitting, a smaller batch size of 100 samples is used. The training is stopped early for the same purpose by monitoring the performance of the model on a validation set with 20% of training samples. The optimum architecture of the final Deep Neural Network model consisted of 2 hidden layers with 24 neurons in the first and 18 neurons in the second hidden layer. For both the hidden layers, Rectified Linear Unit (ReLU) activation resulted in the best performance.

Prediction Intervals

A Gradient Boosting Machine (GBM) model is trained to estimate the errors produced by the Deep Neural Network model. GBM algorithm makes predictions by averaging results obtained from an ensemble of decision trees. These trees are completely different from one another based on the features they use to make decisions at each node. Each of these trees are trained sequentially in a

way that they try to minimize the errors made by the previous trees, which results in a successive decrease of error in subsequent tree ensemble. This leads to a greater prediction accuracy (James et al. 2013) and both faster and efficient computation as compared to neural networks (Barua et al. 2021). GBM is also commonly used to estimate prediction intervals to quantify the uncertainties associated with point estimates (Chakraborty et al. 2020). Therefore, to estimate the prediction intervals for the point IRI estimates, the errors are calculated first by squaring the difference between the predicted and actual IRI. Then the Gradient Boosting Regressor algorithm from the scikit-learn library is used to fit the GBM model for errors. A grid-search approach covering a range of learning rates, number of boosting states, minimum number of samples required to split an internal node, minimum number of samples required to be at a leaf node, and maximum depth of individual regression estimators is used to optimize the model. The standard deviation for each IRI prediction is computed by taking the root of the error predicted by the Gradient Boosting Machine model. The standard deviation is finally adjusted to construct the prediction interval around a predicted IRI.

Model Testing

The most commonly reported metrics to evaluate the goodness-of-fit of regression models in pavement research are the coefficient of determination (R^2), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) (Kargah-Ostadi 2014; Sollazzo et al. 2017; Yamany et al. 2020; Zeiada et al. 2020; Ziari et al. 2016b). R^2 measures the variance in target variable explained by the independent variables. Although it is often very misleading as inclusion of more variables always result in higher R^2 values, it was reported in this paper considering similar studies. MAE describes the average error and RMSE is more useful in limiting larger errors as they assign relatively higher

weight to larger errors (i.e., the errors are squared before averaging). The performance of the models during the training and testing phases were evaluated in terms of the following metrics:

$$R^2 = 1 - \frac{\sum_{i=1}^n (IRI_i - \widehat{IRI}_i)^2}{\sum_{i=1}^n (IRI_i - \overline{IRI})^2} \quad (4-4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |IRI_i - \widehat{IRI}_i| \quad (4-5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (IRI_i - \widehat{IRI}_i)^2} \quad (4-6)$$

4.5 Case Study

To evaluate the capabilities of the proposed framework, a case study analyzing the Minnesota's trunk highway network was undertaken. Minnesota Department of Transportation's (MnDOT) trunk highway system is composed of approximately 14,300 roadway miles of pavement. The entire trunk highway system is surveyed annually to record pavement roughness and surface distresses since the late 1960s (MnDOT 2019). For this project, pavements within the Metro District, covering an area of 3,237 square miles were analyzed.

4.5.1 Pavement Condition and Feature Data

The condition of pavements in the area of study was surveyed using a digital inspection vehicle driven on the outer lane of all trunk highways (MnDOT 2019). Three laser sensors mounted on the front bumper of the vehicle recorded roughness and faulting on both the wheel paths and center of the lane. IRI is estimated as the ratio of a standard vehicle's accumulated suspension motion (inches) and the distance traveled by the vehicle during the measurement period (miles) (Suanpaga and Yoshikazu 2010). This process follows the ASTM E 1926 specifications, where a quarter-car

is driven along the longitudinal profile at a speed of 50 miles/hour and the suspension deflection is estimated using measured profile displacement and standard car structure values (Sayers 1995). Smooth roads result in smaller accumulation of suspension deflection resulting low IRI and rough roads result in high IRI values as illustrated in Figure 4.5. Two lasers mounted on the back of the vehicle were used to capture 3D images of the pavement surface for rut measurements. A camera mounted on the back of the vehicle was used to capture pavement distresses such as cracking and patching. The distresses were recorded at every 1/8 inches as the van travelled at a driving speed, although the measurements were processed at every 0.1 mile. For this study, the pavement condition dataset included IRI data for the entire trunk highway network at every 0.1-mile. In addition to this, pavement features such as age, surface type, layer thicknesses, base type, traffic, and maintenance history (i.e., time and type of last maintenance activity), reference points and their coordinates for the corresponding 0.1-mile segments were compiled to produce a pavement features dataset.



Figure 4.5 US-169 pavement surface showing locations with (a) low, and (b) high IRI values.

In terms of pavement condition indices, this study analyzed pavement roughness (i.e., measured in terms of IRI) and Ride Quality Index (RQI). We decided to use IRI because it is a well-recognized pavement performance indicator and transportation agencies around the world use IRI to measure road surface roughness (Bashar and Torres-Machi 2021; Michigan Department of Transportation 2017). RQI, in turn, is estimated to reflect the users' perceived roughness while driving on a road. To develop a correlation between IRI and RQI, MnDOT asked 32 citizens to rate 120 test sections with different levels of roughness. After driving on each of the 0.25-mile test sections, the panelists rated the quality of their rides on a scale of 0 to 5 based on how they felt about the roughness of these roads. Based on these ratings, the following equations were developed to estimate RQI for asphalt and concrete pavements (MnDOT 2011):

$$RQI_{asphalt} = 5.697 - 0.264 \times \sqrt{IRI} \quad (4-7)$$

$$RQI_{concrete} = 6.634 - 0.353 \times \sqrt{IRI} \quad (4-8)$$

Where, IRI is the International Roughness Index of the pavements in inches/mile.

RQI is an unitless quantity estimated on a numeric scale of 0 to 5, where 5 represents the smoothest ride possible. Newly constructed roads have RQI values greater than 4, whereas pavements are typically rehabilitated for a terminal RQI value of 2.5. MnDOT road categories based RQI are given in Table 4-1. RQI was deemed a valuable indicator of condition, in addition to IRI, because it allows to categorize roughness into a few ordinal categories. Also, RQI is one of the indices currently used by MnDOT for decision-making purposes.

Table 4-1 RQI performance categories

RQI Range	Performance Measure Category
4.1 – 5.0	Very Good
3.1 – 4.0	Good
2.1 – 3.0	Fair
1.1 – 2.0	Poor
0 – 1.0	Very Poor

4.5.2 SAR Imagery

For this project, 91 SAR images captured by Sentinel-1 satellite were obtained from the Alaska Satellite Facility (ASF) (ASF DAAC 2021). The Sentinel-1 constellation is comprised of two polar orbiting satellites (1A and 1B) which images the earth using a C-band SAR sensor. To keep traffic interferences to a minimum, images from 1A satellite were analyzed in this project as it passes over the study area during midnight. The details of the collected data are summarized in Table 4-2.

The acquired SAR imagery, in conjunction with the pavement features, and condition dataset were then processed using the framework proposed in Section 4.4. The Data Processing module of the framework resulted in a dataset consisting of 5,774 samples of road segments. For each segment, the dataset included surface type (asphalt/concrete), surface age in years, pavement layer thickness in inches, base thickness in inches, annual average daily traffic (AADT), γ_0 , and IRI. The thickness of the pavements ranged from 2 to 16 inches with base layers ranging from 0 (i.e., no base layer) to 17 inches. The age of the pavements ranged from 0 (i.e., newly constructed) to 66 years. However, only a smaller number of sections were found to have higher levels of roughness, as

MnDOT maintains the trunk highway network at a very high standard. This resulted in a right-skewed distribution of the IRI values as shown in Figure 4.6(a). The extracted γ_0 values were also overserved to have a similar distribution with a slightly longer upper tail (Figure 4.6b).

Table 4-2 Description of the acquired SAR data

Item	Description
Sensor	Sentinel-1A
Band	C
Wavelength	5.6cm
Spatial Resolution	10m × 10m
Revisit Frequency	12 days
Path	165
Frame	144
Acquisition Mode	Interferometric Wide (IW) swath
Flight Direction	Ascending
Polarization	VV + VH
Level of Preprocessing	L1 Ground Range Detected High Resolution
Number of Images Collected	91
Period Covered	Jan 2017 – Dec 2019
Time of acquisition	00:05

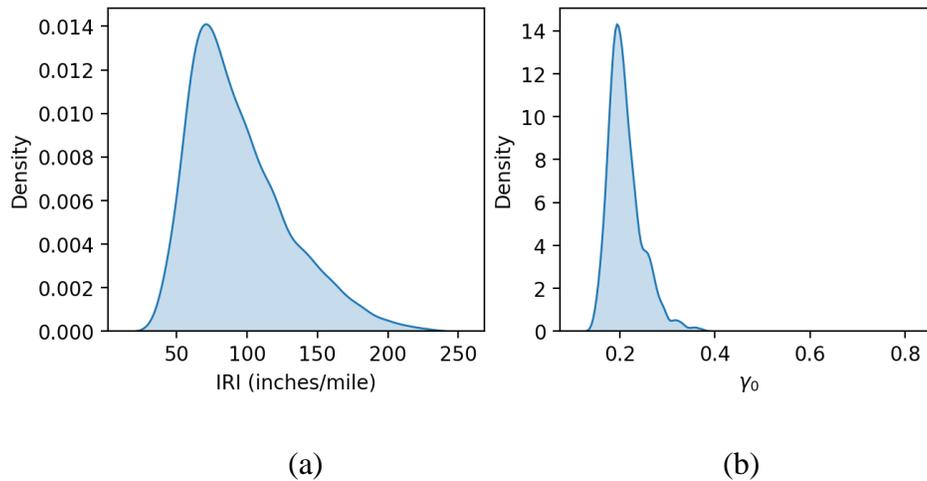


Figure 4.6 Distribution of (a) IRI, and (b) γ_0 values in the final dataset.

4.6 Results

4.6.1 Data Processing

This section describes the improvements in processed SAR data, specifically for the purpose of evaluating pavement condition, resulting from the proposed methodology.

Selection of Appropriate Polarization

The extracted γ_0 values were observed to have a clear pattern when grouped together based on their RQI class (Figure 4.7). Roads in poor condition exhibited stronger backscatters as compared to the roads in better condition, which is consistent with the concepts illustrated in Figure 4.1 (i.e., rough surfaces scatter higher energy as compared to smooth surfaces). This trend is a strong indication of the potential of SAR data in evaluating pavement condition. Figure 4.7 shows that the differences in backscatters for pavements in different condition is more evident in VV polarization compared to the VH polarization. Therefore, using the VV image would be more suitable in

modeling pavement condition. This observation is aligned with the recommendations found in the literature (Karimzadeh and Matsuoka 2020; Meyer et al. 2020).

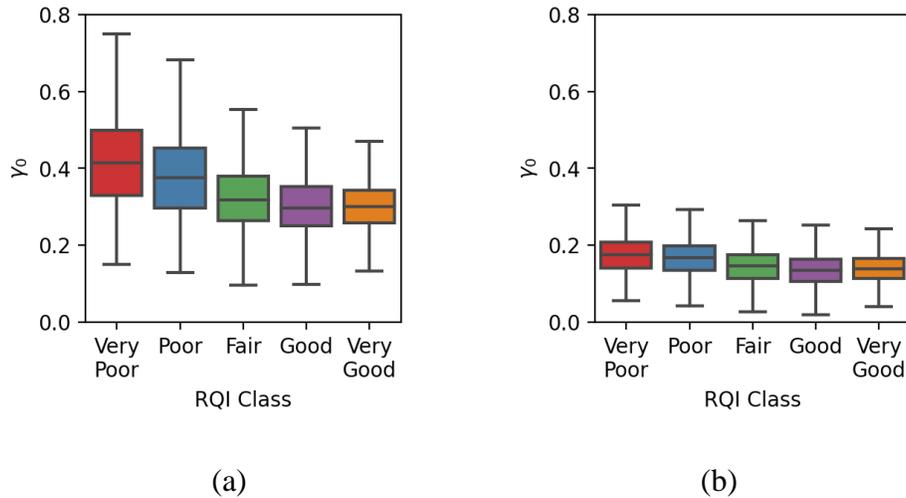


Figure 4.7 Backscatters in (a) VV, and (b) VH polarization for pavements in different condition

Speckle Suppression Performance

The performance of six speckle filters (i.e., Lee, Refined Lee, Lee Sigma, Gamma-map, Frost, and Intensity-Driven Adaptive Neighborhood (IDAN)) were tested to identify the most effective filter in suppressing speckles along the roads. While Lee filter is commonly used for filtering narrow road segments (Karimzadeh and Matsuoka 2020), comparative analysis of the filtered pavement pixels showed that IDAN and Refined Lee perform better than Lee in suppressing speckles across all the performance metrics (Figure 4.8). IDAN resulted in significantly less speckles ($SNI = 0.77$) and offered higher equivalent number of looks ($ENL = 1.68$) as compared to Refined Lee ($SNI = 1.03, ENL = 0.93$). Both IDAN ($NM = 1.08$) and Refined Lee ($NM = 1.07$) performed similarly in preserving original information along the roads. However, when it came to preserving the linear features and texture information, Refined Lee performed significantly better than IDAN

and Lee (Figure 4.9). Since preserving this information is critical for roads, especially for narrower roads, Refined Lee filter is recommended to effectively suppress speckles along the road pixels.

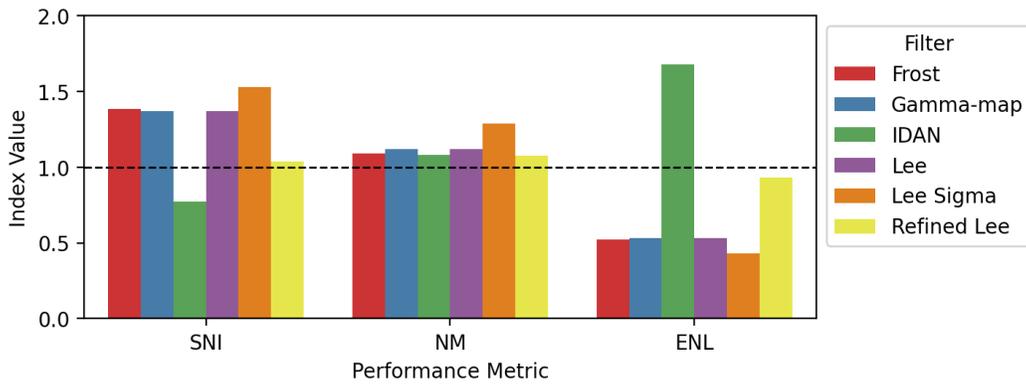


Figure 4.8 Performance of filters in suppressing speckles in pavement pixels

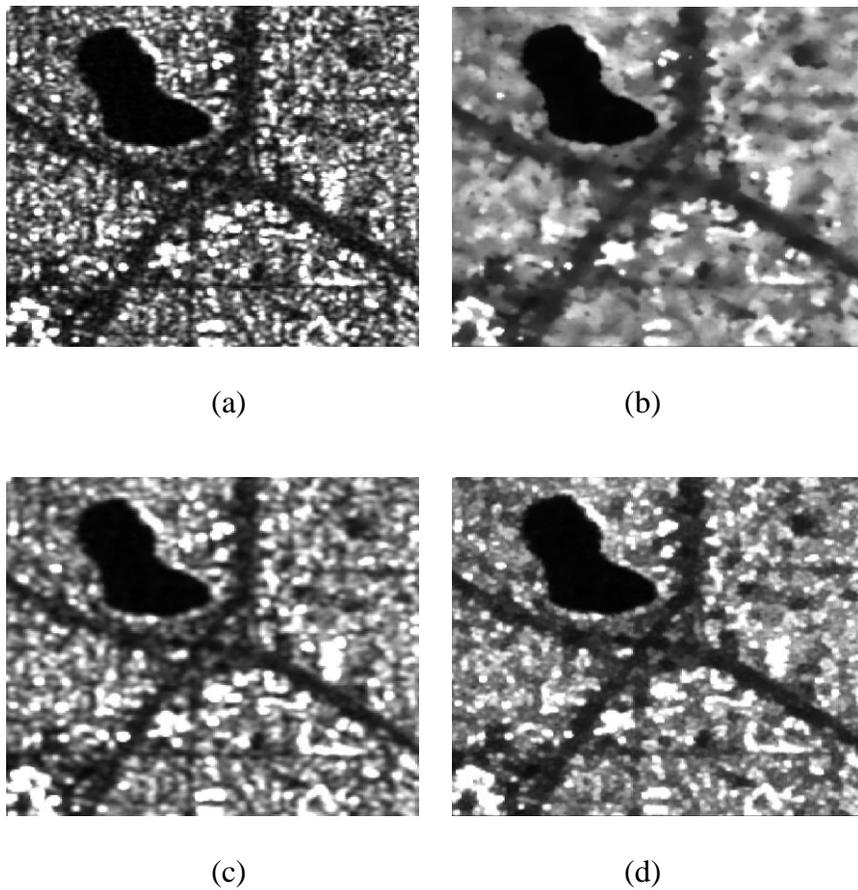


Figure 4.9 (a) Original image as compared to (b) IDAN, (c) Lee, (d) Refined Lee filtered image

Effect of Radiometric Terrain Correction

Radiometric terrain correction was found to be effective in removing the slope impacts on the SAR backscatters. While the backscatters from the highway network considered in this case study were not affected due to its flat terrain, Figure 4.10a shows that the roads located near the Mississippi riverbank were severely affected by the over exposed pixels. A radiometric terrain correction removes the influence of terrain on measured radar brightness (Figure 4.10b). Removing such local biases is essential in establishing meaningful insights from pavement backscatters over a large network. Therefore, it is recommended to apply a radiometric terrain correction as part of the SAR image post-processing in pavement applications.

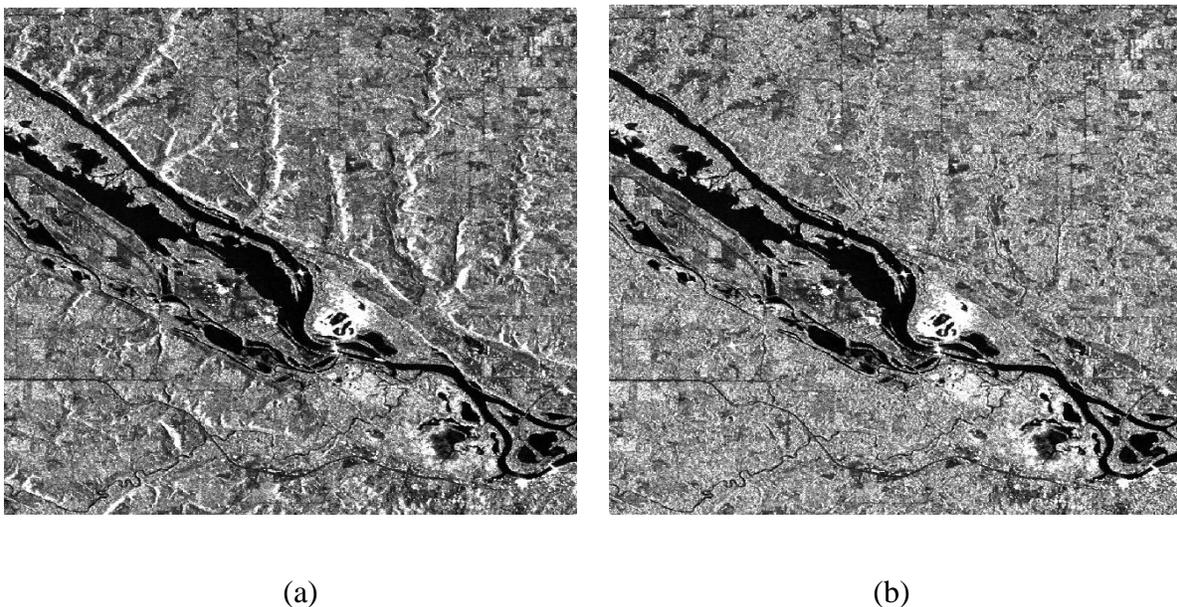


Figure 4.10 Processed SAR image (a) without, and (b) with radiometric terrain correction

Seasonal Variability of SAR Response

Weather conditions such as snowfall and stagnant water in pavements can significantly influence the backscatter signals in SAR data. To better understand the impacts of weather conditions, we

investigated the seasonal variations in SAR backscatter. The objective of this analysis is to identify the appropriate window for SAR data acquisition to avoid the effects of weather on SAR backscatter. One SAR image for each season for the years 2017 to 2019 were used to extract γ_0 values at road reference points after making necessary radiometric and geometric adjustments. Backscatters in winter were constantly lower across all the years as compared to the other seasons, possibly because of the snow reflecting most of the incident signal away. The same is true for spring 2018, when the Twin Cities area received about 26.1 inches of snowfall at the time the image was captured. This snowfall was significantly higher than the ones recorded in 2017 and 2019, which were less than 8 inches over the month of April. These results confirm that snowfall significantly impacts the SAR backscatters.

The backscatter pattern in Summer and Fall were found to be the most consistent over the years. Historical weather data for this area, however, indicates trace amount of snowfalls during the months of September and October (MnDNR 2022). Therefore, the SAR images captured during the summer (i.e., June-August) would be more appropriate to avoid the effects of snowfall. It is also recommended to carefully review the weather conditions for the dates of image acquisition at a specific location to exclude the images including snow from analysis. The remaining analyses of this project has been conducted based on the images acquired during a summer season only.

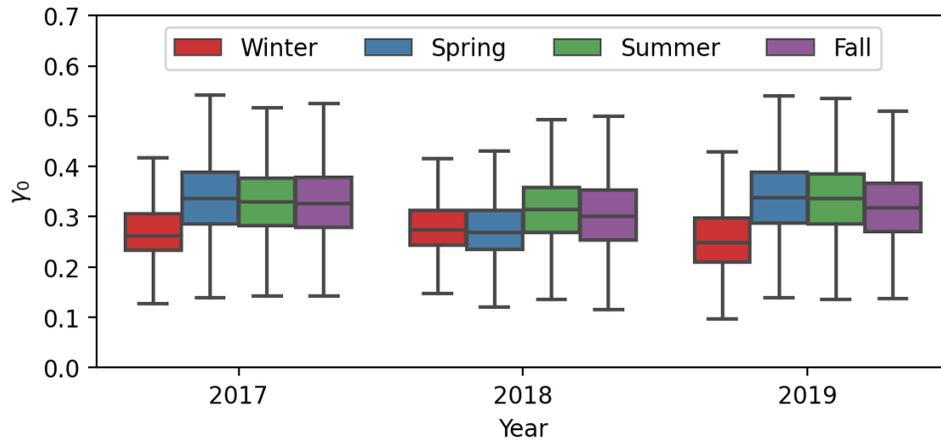


Figure 4.11 Seasonal variations in VV backscatter values from pavements over the study period
Removing Traffic Noise

Images from Sentinel-1A collected during the months of June, July, and August were used to create stacks for different years. These stacks were then used to generate minimum intensity projection images for corresponding years. A visual comparison of the optical satellite images, individual SAR images, and the corresponding minimum intensity projection image indicated that the proposed methodology is highly effective in removing traffic and other temporary noises from the pavement pixels. For example, for the section shown Figure 4.12(a), a SAR image captured on June 4, 2018, had a noise on the road surface (Figure 4.12(b)). While it cannot be confirmed as a noise coming from traffic, it was not present in any of the other images on the 2018 stack. The minimum intensity projection image, shown in Figure 4.12(c), was able successfully remove this temporary noise while preserving the backscatters coming from the permanent object such as the signposts. A careful inspection of all the minimum intensity projection images revealed a similar performance. Therefore, the proposed solution is recommended to effectively minimize traffic and other temporary noises from the road surfaces.

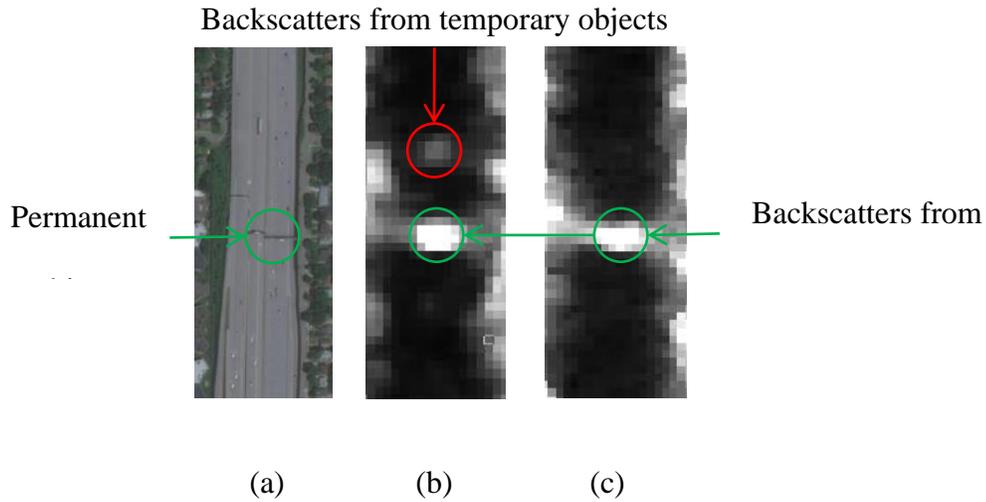


Figure 4.12 (a) Satellite image, (b) an individual SAR image, and (c) the minimum intensity projection image generated from a stack.

4.6.2 Deep Learning Tool

IRI Prediction

The optimal architecture of the Deep Neural Network model was found to be 6-24-18-1 with ReLu as the activation function for both the hidden layers. The model was able to achieve an *RMSE* of 19.41 inches/mile, an *MAE* of 13.96 inches/mile with and an R^2 of 0.68. As illustrated in Figure 4.13, a similar performance was obtained for the test set, indicating that the model does not suffer from overfitting. The predictive performance of the model was further investigated by analyzing the residuals. The residuals were observed to be randomly distributed along the range of predicted values, as shown in Figure 4.14a, indicating that the model does not suffer from heteroscedasticity. The Q-Q plot (Figure 4.14b) also confirms that the residuals are normally distributed. The right tail deviating upwards, however, is indicative of an inferior performance of the model for high IRI values (i.e., residuals are high for higher IRI values).

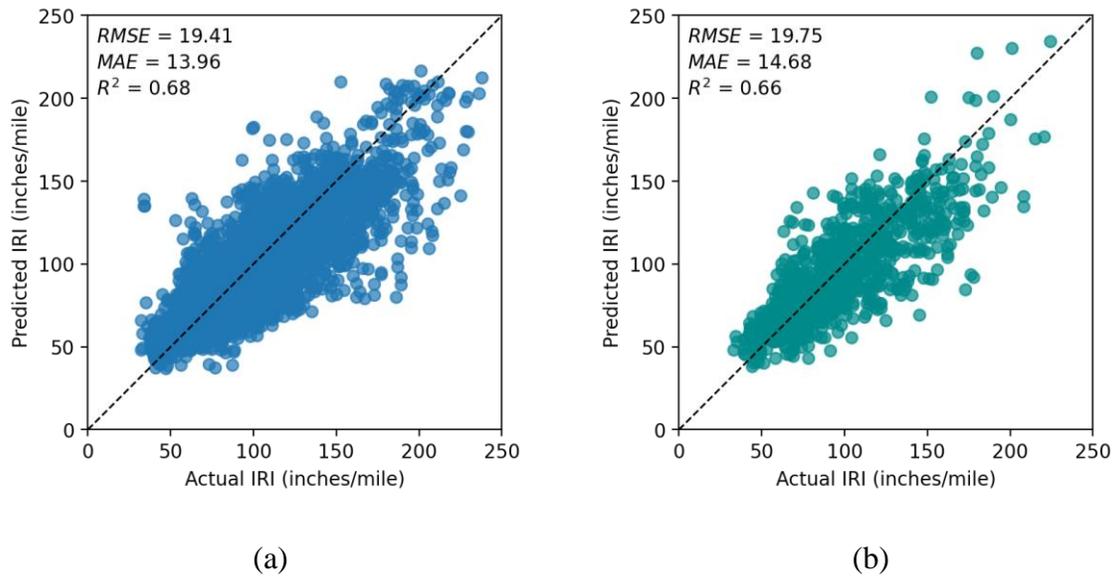


Figure 4.13 Performance of the model during (a) training, and (b) testing.

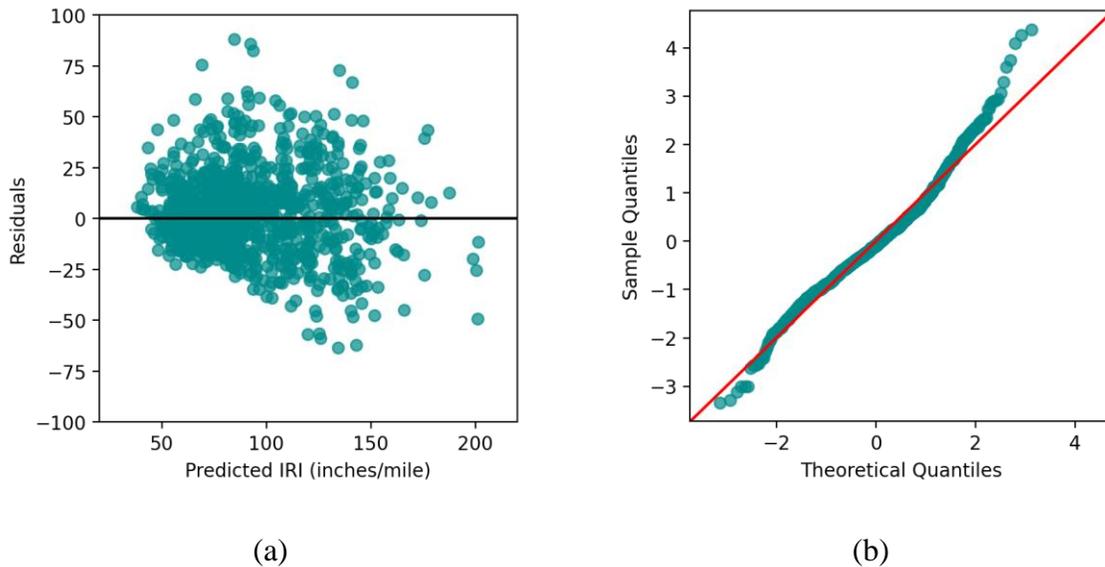


Figure 4.14 (a) Residual plot, and (b) normal Q-Q plot showing the distribution of residuals.

The value added by the deep learning approach can be assessed when the performance of Deep Neural model is compared traditional regression models. A simple linear regression model performance for the same training set is shown Figure 4.15(a), where IRI is predicted using the γ_0 values extracted from the SAR imagery. The multiple linear regression model, as shown in Figure

4.15(b), is trained with all the features in the dataset. While the multiple linear regression model results in a slightly higher correlation between the actual and predicted IRI values, the Deep Neural Network model captures significantly higher amount of variability in data and results in smaller errors in predictions. A similar outcome is observed when the performance of the Deep Neural Network model is compared with the exponential regression model presented in Meyer et al. (Meyer et al. 2020), which results in very high errors values (>30 inches/mile) for IRI values lower than 100 inches/mile.

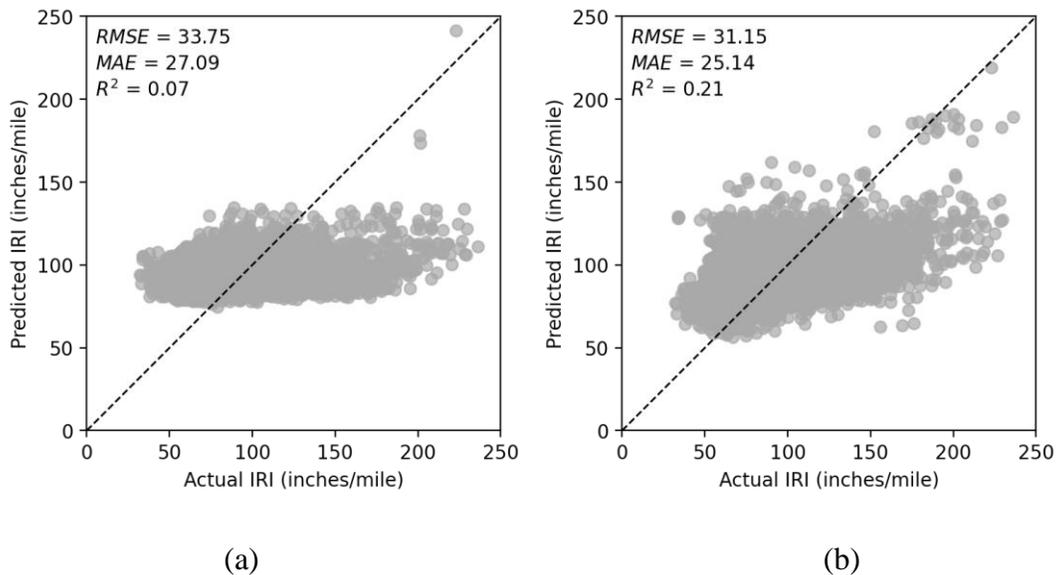


Figure 4.15 Performance of (a) simple linear regression model based on γ_0 , and (b) multiple linear regression based on all the features.

Prediction Intervals

The prediction intervals estimated from the Gradient Boosting Machine model were observed to capture 81% of the actual IRI values within their upper and lower limits. Figure 4.16 shows the estimated prediction intervals for 50 randomly sampled IRI predictions. This figure indicates that the prediction intervals can efficiently capture trends in actual IRI data. Higher values of the

prediction intervals were associated with the most erroneous predictions. These examples are observed for the red dots located way outside of the interval limits in Figure 4.16. The uncertainties captured by these intervals largely stem from the coarser resolution of the SAR pixels. High resolution SAR images with smaller pixel sizes will help filtering out the noises originating from the objects along the side of the roads and can be expected to result in more accurate predictions and smaller prediction intervals.

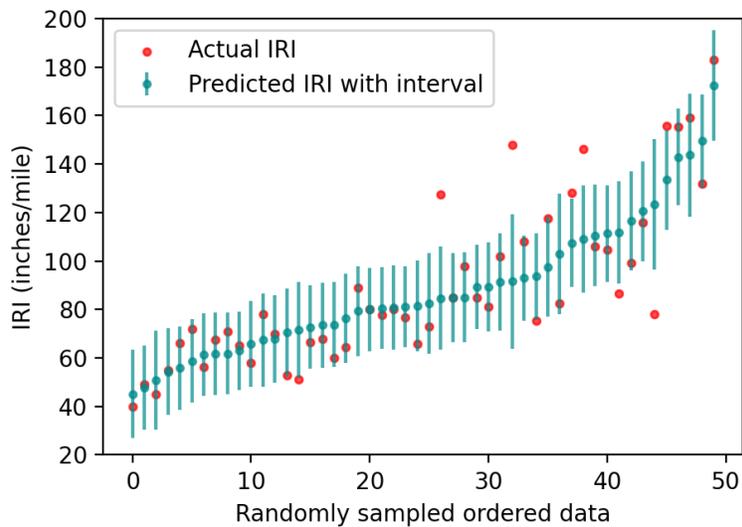


Figure 4.16 Prediction intervals associated with point estimations in comparison to actual IRI values

Classification Accuracy

RQI classes estimated based on the predicted IRI resulted in an overall accuracy of 83%. As illustrated in Figure 4.17, the model performs significantly better for the pavements in Good and Fair condition. When compared to the classification accuracy of 87% as reported for the L-band SAR data based binary logit model presented in Suanpaga and Yoshikazu (Suanpaga and Yoshikazu 2010), the Deep Neural Network model underperforms for the extreme categories. This performance was observed to be highly influenced by the sample size of the corresponding

categories. Classification accuracy sharply dropped to 31% for the Poor RQI class, as the representation of this class is only 1.4% in the dataset. The extreme classes constituted less than 1% of dataset and, as a result, the model rarely classifies a segment as very poor or very good. While the model performs satisfactorily for the common range of IRI values, a more balanced dataset will improve the model performance over a greater range of RQI classes.

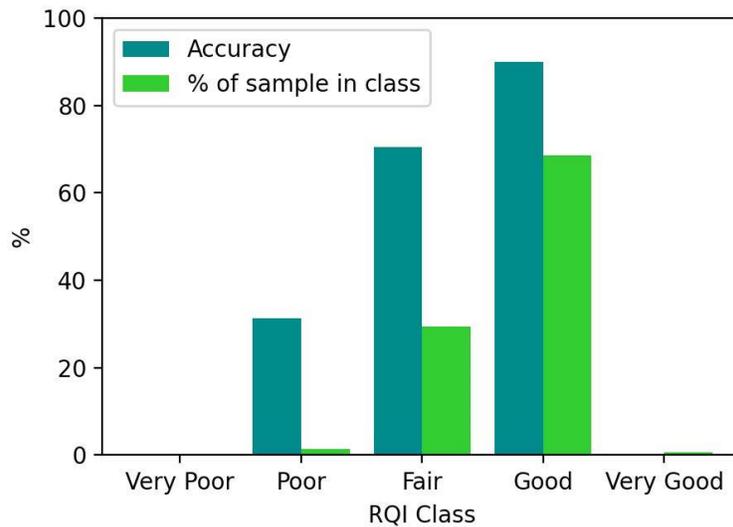


Figure 4.17 Classification accuracy of the model for different RQI classes

4.6.3 Model Deployment

To facilitate an easy deployment of the developed models by transportation agencies worldwide, a program with a graphical user interface was developed using Python’s Tkinter library. Given a properly processed SAR image and pavement features, the SAR based Condition (SAR-C) evaluation tool (Figure 4.18) estimates IRI, associated prediction intervals, and RQI class for the road segments of interest. The user manual of the program describes in detail the steps of processing SAR images with an example following the proposed framework. The user manual can be accessed here: <https://github.com/infra-health/sar-c>

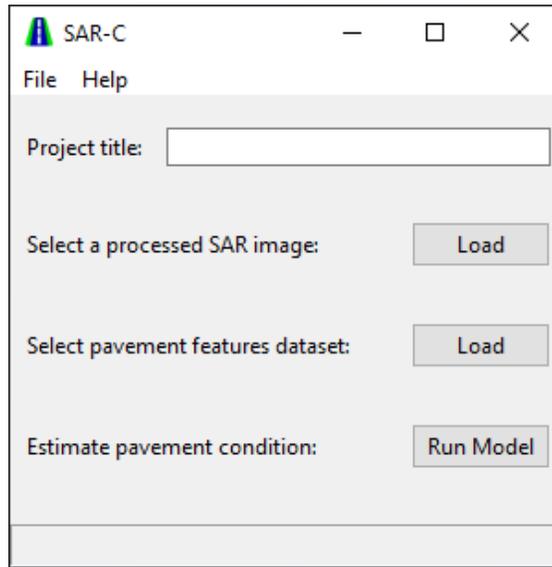


Figure 4.18 SAR-C user interface

4.7 Conclusions and Recommendations

This paper introduces a novel framework to estimate pavement IRI using deep learning and spaceborne SAR imagery. A case study analyzing the trunk highway network in Minnesota was undertaken to identify the improvements in SAR image processing for pavement applications as well as to demonstrate the predictive performance of the developed deep learning tool. Specific conclusions and recommendations derived from this project are summarized below.

4.7.1 Conclusions

- Sentinel-1 SAR images were found to have a strong potential in quantifying pavement roughness. While it is not as highly accurate as the IRI measured by digital inspection vehicles, it can be used to evaluate the condition of local, ancillary, or low priority roads which are not typically monitored, and where a very accuracy is not necessarily needed.

- The proposed framework is highly capable in improving SAR image processing for pavement applications as it effectively addresses the challenges of removing traffic noises from pavements, suppressing speckles without comprising the road features, and eliminating the effects of terrain on SAR backscatters.
- The deep learning tool can predict IRI with an *MAE* ranging from 13.9 to 14.6 inches/mile. The associated prediction intervals were found to capture 81% of the actual IRI values within their upper and lower limits. The tool is also effective at classifying RQI classes, with an overall classification accuracy of 83%.

4.7.2 Recommendations

- The VV polarization image was found to be more sensitive to pavement roughness as compared to the VH polarization.
- Refined Lee filter is recommended to remove speckles, as it preserves the edges and texture of linear road features.
- The analysis of SAR images should include a radiometric terrain correction to remove the effect of slopes on SAR backscatters.
- Identifying an appropriate time window for collecting SAR images over a specific region is critical to avoid the effects of weather on SAR backscatters.
- The generation of a minimum intensity image from a stack of SAR images is an effective solution to eliminate traffic noises from the pavement pixels.

4.7.3 Limitations and Future Research

The proposed framework is currently limited by the resolution of Sentinel-1 images as in many cases the width of the roads can be less than the size of the pixels. This raises an interesting future avenue for research using high resolution X-band SAR images captured by the Cosmo-SkyMed satellite.

The limitations of the deep learning tool in predicting higher IRI values can also be addressed by including examples in the dataset from a wider range of road classes. It will be particularly important to include examples of pavement in Very Good and Very Poor condition to have a more balanced dataset.

Finally, calibrating and testing the model for roads with different physical attributes (e.g., wider highways, narrower ancillary roads) and geographic locations using transfer learning will enhance the scale of implementation of the SAR-C software developed in this project.

CHAPTER 5. CONCLUSIONS, CONTRIBUTIONS, AND LIMITATIONS

This research aimed at leveraging the capabilities of satellite data and machine learning in the management of infrastructure assets. Two types of satellite data (i.e., high-resolution multispectral and Synthetic Aperture Radar (SAR) imagery) and different machine learning techniques, including partially observable Markov decision process and deep learning, were investigated in the context of evaluating pavement condition and making maintenance decisions. This chapter summarizes the outcomes and contributions resulting from each of the studies included in this dissertation. It also sheds light on avenues for future research based on the limitations of the work presented in this dissertation.

5.1 Conclusions and Recommendations

The conclusions and recommendations derived from each of the chapters are discussed below in the context of specific objectives of this dissertation:

ROI. Quantify the value of including satellite imagery in optimal inspection and maintenance strategies over the pavement life cycle.

The research presented in Chapter 2 concluded that satellite inspections are valuable when the observation accuracy is greater than 70%. This value is significantly greater than the accuracy obtained using the histogram-based method presented in Chapter 2 of this dissertation (i.e., approximately 11% to 57%) accuracy, highlighting the need to improve this accuracy for satellite-based observations to be a competitive alternative to monitor pavements. Satellite information is found to be significantly more valuable for roads which are not typically monitored as compared to roads monitored using highly accurate distress surveys. For non-monitored roads, satellite data results in about 6.5% reductions in expected lifecycle cost of the pavements, as compared to 0.75%

reductions for roads monitored annually with distress surveys. The value of a satellite-based monitoring system is estimated to range from 0.2 to 4.0% of the total cost of a monitoring system when used for monitored highways, and 10 to 22% when used for non-monitored (i.e., local and ancillary) roads. This outcome is critical for transportation agencies, as the investment to adopt a satellite-based inspection system needs to be lower than these amounts for the whole system to be cost-effective. While the negative value indicates that satellite-based inspections may not result in economic benefits in some cases because of its accuracy, it is important to point out that this accuracy is not necessarily a limitation of the satellite data itself, rather how it has been used to characterize the pavement condition states (e.g., the pixel brightness-based histogram approach). Therefore, improvements in satellite data processing and frameworks to estimate pavement condition using satellite data will add more value to these observations without increasing the cost and make the system more affordable and appealing to the transportation authorities.

RO2. Evaluate the performance of machine learning algorithms in predicting pavement condition as compared to traditional techniques.

The work presented in Chapter 3 concluded that machine learning algorithms outperformed traditional regression-based approaches in predicting pavement condition (i.e., measured in terms of the International Roughness Index, IRI). Machine learning algorithms could capture, on average, 15.6% more variability in IRI than traditional techniques. Random Forest was found to be the most accurate machine learning technique with an overall performance of 0.995. Artificial Neural Networks and Support Vector Machine were also found to be highly accurate in predicting IRI, with performance values of 0.930 and 0.916 respectively, although the Support Vector Machine models showed a very high degree of variability. Artificial Neural Networks showed a

consistent performance over different studies with varying sample sizes and data sources. Based on these results, Artificial Neural Networks are recommended to model IRI. For these models, a single layer with nodes equal to 0.3 to 2 times the number of input features, and a sample size greater than 50 times the number of input features are recommended to obtain a satisfactory predictive performance.

RQ3. Develop a machine-learning based approach to assess pavement condition using publicly available Synthetic Aperture Radar (SAR) data.

The paper presented in Chapter 4 explores the capabilities of using machine learning in the analysis of satellite-based data to evaluate pavement condition. This study introduced a framework to improve SAR data processing that effectively addresses the challenges of removing traffic noises from pavements, suppressing speckles without comprising the road features, and eliminating the effects of terrain on SAR backscatters. The proposed deep learning model resulted in accurate IRI predictions with a mean absolute error (*MAE*) ranging from 13.9 to 14.6 inches/mile. The associated prediction intervals were found to capture 81% of the actual IRI values within their upper and lower limits. The tool is also effective at classifying Road Quality Index (RQI) classes, with an overall classification accuracy of 83%. Specific recommendations to apply SAR data for pavement applications derived from this project include: use vertical-vertical (VV) polarization image to model pavement roughness, apply Refined Lee filter to remove speckles, apply radiometric terrain correction to remove the effect of slopes on SAR backscatters, and generate a stack-based minimum intensity image to remove traffic and other temporary noises from the pavement pixels.

5.2 Contributions

The contributions of this research hinge around the two foci of this dissertation: application of uncertain satellite data and machine learning to pavement asset management. The theoretical and practical contributions resulting from this research are summarized below.

5.2.1 Contributions to the Theory

Chapter 2 contributes to the body of knowledge of using uncertain sensor data in pavement monitoring and builds upon previous research studying the application of optical satellite imagery to pavements (Herold et al. 2004; Mohammadi 2012; Noronha et al. 2002; Pan et al. 2016; Shahi et al. 2015). While all these studies agree that the accuracy of optical satellite imagery is the major limitation hindering the deployment of this technology for pavement condition monitoring, no studies have inspected how these uncertainties would influence decision-making if satellite-based data were to be included in the management process. The research presented in chapter addresses this knowledge gap by establishing theoretical relationships between observational uncertainty and life-cycle value of optimal policies in the context of a partially observable stochastic environment. The proposed approach allows to quantify the value of satellite-based information and how this value changes with the accuracy of satellite-based data. This study also introduces a histogram-based approach to establish levels of outcomes from any uncertain sensor data, which in turn, determines the accuracy that can be achieved for each of outcome level. Another major contribution of this research is that it provides an objective process to quantify the value of the information provided by uncertain satellite data, which in turn, gives the cost of adopting a satellite-based pavement monitoring system. This process will allow future work involving novel

sensor types to determine if adopting those sensor-based monitoring systems would add value to the pavement maintenance decision-making process.

Chapter 3 contributes to the body of knowledge of applying machine learning techniques to model pavement performance. It paints an overall picture of the current state of machine learning in pavement monitoring by bringing together previous research efforts that applied machine learning to estimate pavement condition. While machine learning techniques have become a common practice in a wide variety of fields, the lack of quantitative evidence supporting the effectiveness of these techniques for pavement application limits the adoption of these techniques by infrastructure managers and transportation agencies. This study addresses this gap in knowledge by quantitatively establishing that machine learning algorithms capture, on average, 15.6% more variability than traditional techniques.

Chapter 4 contributes to the body of knowledge of using SAR data for pavement applications. It develops a framework to leverage publicly available SAR data in estimating pavement roughness. This study reviews common practices in SAR data processing and modifies the standard SAR data processing routine to better address the challenges associated with pavement applications. The primary limitation of using radar signals to model pavement roughness is cancelling out the return signals originating from the vehicles on roads. This research addresses this challenge by introducing a stacking-based minimum intensity image generation process to effectively remove traffic and other temporary noises from pavement pixels. It also advances the knowledge on sensitivity of radar signal polarization to pavement roughness, filters to effectively suppress speckles without compromising the texture and edge of linear road features, and radiometric terrain correction of SAR images to remove the effects of slopes on radar backscatters.

5.2.2 Contributions to the Practice

Chapter 2 contributes to the current practice of determining optimal maintenance strategies over pavement life cycle considering the observational uncertainties. Current methods for pavement maintenance decision-making assume the condition of pavements are known with certainty, while in reality the observations are not completely accurate. As a result, decisions made based on these observations may not be optimal. The research presented in this chapter addresses this gap by demonstrating the application of POMDP framework to account for uncertainties in decision-making. Furthermore, the values estimated for satellite-based monitoring system will help the transportation agencies determine if adding satellite inspections to their existing pavement monitoring system will be cost effective and how much they can practically invest to acquire, install, operate, and maintain such a system.

Chapter 3 contributes the current practice of using machine learning algorithms for pavement applications. The recommendations derived from the meta-analysis presented in this chapter will serve as a state-of-the-art guide and reference for practitioners, researchers, and highway agencies on the use of machine learning techniques as a tool to predict pavement performance.

Chapter 4 contributes to the current practice of evaluating pavement condition. The deep learning tool developed in this research can be used to estimate pavement IRI using publicly available SAR data. The agencies which already have a sophisticated condition assessment system in place will be able to use this tool to quantitatively evaluate the condition of their non-monitored part of their road network in a rapid and cost-effective manner. Minnesota Department of Transportation is currently testing the feasibility of the SAR-C software developed in this research to evaluate the condition of their ancillary roads. This software will especially benefit countries where the concept

of pavement monitoring is still emerging, and the agencies are operating at limited capacity with the need for a cost-effective system to evaluate the condition of their roads.

5.3 Limitations

The major limitation of the work presented in Chapter 2 arises from our assumption of stationarity in deterioration transitions. This results in somewhat unrealistic optimal policies where the decision-maker can bring the pavements back to a good state with a minor repair at the end of the pavements' service life. In reality, the transitions are non-stationary as the pavements deteriorate at a faster rate with age (Bashar et al. 2019). This would require the decision maker to perform a major repair to cost-effectively keep the pavement in a good state. Also, characterizing pavement condition with 3 discrete condition states do not address the Markovian property of independence from history which states that given the present, the future is independent of the past (Papakonstantinou and Shinozuka 2014b).

This study presented in Chapter 3 does not consider the effects of data pre-processing on the performance of the predictive models. While the quality of data pre-processing significantly influences the model performance, measuring it based on the descriptions from a paper is highly subjective. The study is also limited by the number of Random Forest models included in the analysis, as compared to the other machine learning techniques (i.e., artificial neural network and support vector machine) which have been more widely applied by pavement researchers.

The framework presented in Chapter 4 is currently limited by the resolution of Sentinel-1 satellite, as in many cases, the width of the pavements is smaller than the size of the pixels. This problem is particularly important when analyzing local roads, as they are usually narrower than highways. This low resolution causes noise in the return signal, as it captures information from objects on the

side of the roads, thus affecting the performance of the model. The model was also observed to suffer from a loss in performance when predicting high IRI values. This limitation primarily stemmed from the unbalanced dataset used to train the models, as 95% of the data represents pavements in Fair and Good, and only 5% of the dataset corresponded to pavements in Poor condition.

5.4 Future Research

The limitations in current research raise interesting avenues for future research. Potential extensions of current work would include modeling time dependent POMDPs to capture the true effect of observational uncertainty on maintenance decisions over the pavement life cycle. To simulate a more dynamic system with continuous states (e.g., IRI values), we plan to augment the number of states by combining levels of pavement condition with varying rates of deterioration. Another priority for future research relies on improving the accuracy of the histogram-based approach by considering other texture metrics. This analysis would result in a higher accuracy of the satellite observations without increasing the cost of satellite inspections. Another long-term goal would be to investigate the applicability of a pansharpened image (i.e., combination of multispectral and panchromatic) in deriving the histograms to leverage both high-resolution texture and rich spectral information in establishing a more accurate observation matrix.

The predictive performance of the deep learning tool can be substantially improved by using a more balanced dataset for model training. Including more examples from a wider range of road classes will help the models to generalize over a greater range of IRI values. Calibrating the models using transfer learning for different road types (e.g., ancillary, unpaved) and geographic regions will further improve the model performance for specific applications.

Exploring the capabilities of higher resolution SAR images is another potential avenue for future research, as it can significantly reduce the uncertainties associated with coarser SAR pixels. X-band SAR data captured by the Cosmo-SkyMed satellite has a spatial resolution of 3 m, which is significantly better than the resolution of Sentinel-1A used in this study (i.e., 10 m). Higher spatial resolution is particularly important for ancillary roads, as they are narrower than highways. While this data is commercial, it offers several advantages over the C-band Sentinel-1 images. Their smaller sensor wavelength (i.e., ~3.1 cm), results in better surface scatters, as no signal is lost from penetration into the ground, when compared to other sensors with higher wavelengths. Higher spatial resolution will also allow averaging scatters over every a smaller (i.e., 3 m) section and capture higher variations in pavement roughness than the C-band images. Studies using X-band SAR data to estimate pavement condition (Karimzadeh and Matsuoka 2020; Meyer et al. 2020) were able to achieve a very high degree of accuracy. Furthermore, pavement texture derived from optical imagery has been found to be highly correlated with surface condition (Bashar and Torres-Machi 2022). Therefore, combining SAR responses with the textural and spectral information derived from optical imagery can help develop a robust system that would leverage the capabilities of satellite data to estimate pavement condition.

REFERENCES

- Abaza, K. A., S. A. Ashur, S. A. Abu-Eisheh, and A. Rabay'a. 2001. "Macroscopic Optimum System for Management of Pavement Rehabilitation." *J. Transp. Eng.*, 127 (6): 493–500. American Society of Civil Engineers. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2001\)127:6\(493\)](https://doi.org/10.1061/(ASCE)0733-947X(2001)127:6(493)).
- Abd El-Hakim, R., and S. El-Badawy. 2013. "International roughness index prediction for rigid pavements: An Artificial Neural Network Application." *Adv. Mater. Res.*, 723: 854–860. <https://doi.org/10.4028/www.scientific.net/AMR.723.854>.
- Abdelaziz, N., R. T. Abd El-Hakim, S. M. El-Badawy, and H. A. Afify. 2020. "International Roughness Index prediction model for flexible pavements." *Int. J. Pavement Eng.*, 21 (1): 88–99. Taylor & Francis. <https://doi.org/10.1080/10298436.2018.1441414>.
- Abduljabbar, R., H. Dia, S. Liyanage, and S. A. Bagloee. 2019. "Applications of artificial intelligence in transport: An overview." *Sustain.*, 11 (1). <https://doi.org/10.3390/su11010189>.
- Al-Qadi, I. L., and S. Lahouar. 2005. "Measuring layer thicknesses with GPR – Theory to practice." *Constr. Build. Mater.*, 19 (10): 763–772. Elsevier. <https://doi.org/10.1016/J.CONBUILDMAT.2005.06.005>.
- Alruqi, W. M., and M. R. Hallowell. 2019. "Critical Success Factors for Construction Safety: Review and Meta-Analysis of Safety Leading Indicators." *J. Constr. Eng. Manag.*, 145 (3). [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001626](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001626).
- Amin, S. R., and L. E. Amador-Jiménez. 2016. "Backpropagation Neural Network to estimate pavement performance: dealing with measurement errors." <http://dx.doi.org/10.1080/14680629.2016.1202129>, 18 (5): 1218–1238. Taylor & Francis. <https://doi.org/10.1080/14680629.2016.1202129>.
- Andriotis, C. P., K. G. Papakonstantinou, and E. N. Chatzi. 2020. "Value of structural health information in partially observable stochastic environments."
- Andriotis, C. P., K. G. Papakonstantinou, and E. N. Chatzi. 2021. "Value of structural health information in partially observable stochastic environments." *Struct. Saf.*, 93 (August): 102072. Elsevier Ltd. <https://doi.org/10.1016/j.strusafe.2020.102072>.
- Anys, H., A. Bannari, D. C. He, and D. Morin. 1994. "Texture analysis for the mapping of urban areas using airborne MEIS-II images." *1st Int. Airborne Remote Sens. Conf. Exhib.*, 231–245.
- Arhin, S. A., E. C. Noel, and A. Ribbiso. 2015. "Acceptable International Roughness Index Thresholds based on Present Serviceability Rating." *J. Civ. Eng. Res.*, 5 (4): 90–96. <https://doi.org/10.5923/j.jce.20150504.03>.
- ASCE. 2021. *2021 Report Card for America's Infrastructure*.

- ASF DAAC. 2021. “Contains modified Copernicus Sentinel data, processed by ESA.”
- Barnes, G., and P. Langworthy. 2003. *The per-mile costs of operating automobiles and trucks*.
- Barua, L., B. Zou, M. Noruzoliaee, and S. Derrible. 2021. “A gradient boosting approach to understanding airport runway and taxiway pavement deterioration.” *Int. J. Pavement Eng.*, 22 (13): 1673–1687. Taylor and Francis Ltd. <https://doi.org/10.1080/10298436.2020.1714616/FORMAT/EPUB>.
- Bashar, M. Z., M. A. Elseifi, M. R. Mousa, Z. Zhang, and K. Gaspard. 2019. “Optimizing the Performance of Microsurfacing Treatments in Flexible Pavements and Assessing Its Effects on Moisture Damage.” *Transp. Res. Rec.* <https://doi.org/10.1177/0361198118821874>.
- Bashar, M. Z., and C. Torres-Machi. 2021. “Performance of Machine Learning Algorithms in Predicting the Pavement International Roughness Index.” *Transp. Res. Rec. J. Transp. Res. Board*.
- Bashar, M. Z., and C. Torres-Machi. 2022. “Exploring the Capabilities of Optical Satellite Imagery in Evaluating Pavement Condition.” *Constr. Res. Congr. 2022*, 108–115.
- Batratkov, D. O., A. G. Batratkova, S. N. Urdzik, and V. R. Danielyan. 2021. “Nondestructive Diagnostics and Detection Of Subsurface Cracks In Non-Rigid Pavements With GPR.” *DIAGNOSTYKA*, 22 (2). <https://doi.org/10.29354/diag/137915>.
- Batratkova, A. G., D. O. Batratkov, and M. S. Antyufeyeva. 2018. “Pavement deterioration model based on GPR datasets.” *Roads Bridg. - Drog. i Most.*, 17 (1): 55–71. Instytut Badawczy Drog i Mostow. <https://doi.org/10.7409/RABDIM.018.004>.
- Bayrak, M. B., E. Teomete, and M. Agarwal. 2004. “Use of artificial neural networks for predicting rigid pavement roughness.” *Midwest Transp. Consort.*, 1–18. Ames, Iowa.
- Bekkari, A., S. Idbraim, A. Elhassouny, D. Mammass, M. El Yassa, and D. Ducrot. 2012. “SVM and haralick features for classification of high resolution satellite images from urban areas.” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 17–26. Springer, Berlin, Heidelberg.
- Bellman, R. E. 1957. *Dynamic Programming*. Princeton, NJ: Princeton University Press.
- Berke, L., and P. Hajela. 1993. “Application of Neural Nets in Structural Optimization.” *Optim. Large Struct. Syst.*, 731–745. Springer Netherlands.
- Borenstein, M., L. V Hedfes, J. Higgins, and H. R. Rothstein. 2009. *Introduction to Meta-Analysis*. Wiley.
- Breiman, L. 2001. “Random forests.” *Mach. Learn.*, 45 (1): 5–32. Springer. <https://doi.org/10.1023/A:1010933404324>.

- Brownlee, J. 2019a. “How to Configure the Learning Rate When Training Deep Learning Neural Networks.” *machinelearningmastery*. Accessed July 29, 2020. <https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks/>.
- Brownlee, J. 2019b. “How Much Training Data is Required for Machine Learning?” *Mach. Learn. Mastery*. Accessed July 26, 2020. <https://machinelearningmastery.com/much-training-data-required-machine-learning/>.
- Bruckno, B., A. Vaccari, E. Hoppe, W. Niemann, and E. Campbell. 2013. “Validation of Interferometric Synthetic Aperture Radar as a tool for identification of geohazards and at-risk transportation infrastructure.” *Proc. Highw. Geol. Symp.*, 1–19.
- Butt, A., M. Shahin, K. Feighan, and S. Carpenter. 1987. *Pavement performance prediction model using the Markov process*.
- Carnahan, J. V. 1988. “Analytical framework for optimizing pavement maintenance.” *J. Transp. Eng.*, 114 (3): 307–322. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1988\)114:3\(307\)](https://doi.org/10.1061/(ASCE)0733-947X(1988)114:3(307)).
- Ceylan, H., M. B. Bayrak, and K. Gopalakrishnan. 2014. “Neural networks applications in pavement engineering: A recent survey.” *Int. J. Pavement Res. Technol.*, 7 (6): 434–444. <https://doi.org/10.6135/ijprt.org.tw/2014>.
- Chakraborty, D., H. Elhegazy, H. Elzarka, and L. Gutierrez. 2020. “A novel construction cost prediction model using hybrid natural and light gradient boosting.” *Adv. Eng. Informatics*, 46: 101201. Elsevier. <https://doi.org/10.1016/J.AEI.2020.101201>.
- Chakure, A. 2019. “Random Forest Regression.” *towardsdatascience*. Accessed July 29, 2020. <https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f>.
- Chandra, S., C. R. Sekhar, A. K. Bharti, and B. Kangadurai. 2013. “Relationship between pavement roughness and distress parameters for indian highways.” *J. Transp. Eng.*, 139 (5): 467–475. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000512](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000512).
- Chester, D. 1990. “Why two hidden layers are better than one.” *Proceeding IEEE Int. Jt. Conf. Neural Networks*, 265-268.
- Choi, J. H., T. M. Adams, and H. U. Bahia. 2004. “Pavement roughness modeling using back-propagation neural networks.” *Comput. Civ. Infrastruct. Eng.*, 19 (4): 295–303. <https://doi.org/10.1111/j.1467-8667.2004.00356.x>.
- Choi, S., and M. Do. 2020. “Development of the road pavement deterioration model based on the deep learning method.” *Electron.*, 9 (1). <https://doi.org/10.3390/electronics9010003>.
- “Copernicus Sentinel data. Retrieved from ASF DAAC. Processed by ESA.” 2019.
- Corotis, R. B., J. Hugh Ellis, and M. Jiang. 2005. “Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision processes.” *Struct. Infrastruct.*

- Eng.*, 1 (1): 75–84. <https://doi.org/10.1080/15732470412331289305>.
- Crimmins, T. R. 1985. “Geometric filter for speckle reduction.” *Appl. Opt. Vol. 24, Issue 10*, pp. 1438–1443, 24 (10): 1438–1443. Optica Publishing Group. <https://doi.org/10.1364/AO.24.001438>.
- Cybenko, G. 1989. “Approximation by superpositions of a sigmoidal function.” *Math. Control. Signals, Syst.*, 2 (4): 303–314. Springer-Verlag. <https://doi.org/10.1007/BF02551274>.
- Darko, A., A. P. C. Chan, M. A. Adabre, D. J. Edwards, M. R. Hosseini, and E. E. Ameyaw. 2020. “Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities.” *Autom. Constr.*, 112: 103081. Elsevier B.V. <https://doi.org/10.1016/j.autcon.2020.103081>.
- Doi, S. A. R., J. J. Barendregt, S. Khan, L. Thalib, and G. M. Williams. 2015. “Advances in the meta-analysis of heterogeneous clinical trials I: The inverse variance heterogeneity model.” *Contemp. Clin. Trials*, 45: 130–138. Elsevier Inc. <https://doi.org/10.1016/j.cct.2015.05.009>.
- Du, Y., C. Liu, D. Wu, and S. Jiang. 2014. “Measurement of international roughness index by using Z -axis accelerometers and GPS.” *Math. Probl. Eng.*, 2014. <https://doi.org/10.1155/2014/928980>.
- Earthdata. 2020. “What is Synthetic Aperture Radar?” NASA. Accessed March 16, 2021. <https://earthdata.nasa.gov/learn/backgrounders/what-is-sar>.
- Ellis, H., M. Jiang, and R. B. Corotis. 1995. “Inspection, Maintenance, and Repair with Partial Observability.” *J. Infrastruct. Syst.*, 1 (2): 92–99. American Society of Civil Engineers (ASCE). [https://doi.org/10.1061/\(asce\)1076-0342\(1995\)1:2\(92\)](https://doi.org/10.1061/(asce)1076-0342(1995)1:2(92)).
- Elvik, R. 2005. “Introductory guide to systematic reviews and meta-analysis.” *Transp. Res. Rec.*, (1908): 230–235. <https://doi.org/10.3141/1908-28>.
- Emery, W. 2014. *Large-Area Road-Surface Quality Using Very-High Spatial Resolution Aerial and Satellite Data*. U.S. Department of Transportation.
- Faggella, D. 2020. “What is Machine Learning?” *emerj*. Accessed July 28, 2020. <https://emerj.com/ai-glossary-terms/what-is-machine-learning/>.
- Fagrhi, A., and A. Ozden. 2015. *Satellite Assessment and Monitoring for Pavement Management*. Newark, Delaware.
- FHWA. 2015. “Public Road Length - 2015 Miles by Ownership Table HM-10.” *Off. Highw. Policy Inf.* Accessed April 23, 2022. <https://www.fhwa.dot.gov/policyinformation/statistics/2015/hm10.cfm>.
- FHWA. 2016. *Highway Performance Monitoring System Field Manual*.

- Flintsch, G. W., and C. Chen. 2004. "Soft Computing Applications in Infrastructure Management." *J. Infrastruct. Syst.*, 10 (4): 157–166. [https://doi.org/10.1061/\(asce\)1076-0342\(2004\)10:4\(157\)](https://doi.org/10.1061/(asce)1076-0342(2004)10:4(157)).
- Flood, I., and N. Kartam. 1994. "Neural Networks in Civil Engineering. I: Principles and Understanding." *J. Comput. Civ. Eng.*, 8 (2): 131–148. [https://doi.org/10.1061/\(ASCE\)0887-3801\(1994\)8:2\(131\)](https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(131)).
- Georgiou, P., C. Plati, and A. Loizos. 2018. "Soft Computing Models to Predict Pavement Roughness: A Comparative Study." *Adv. Civ. Eng.*, 2018. <https://doi.org/10.1155/2018/5939806>.
- Goel, K., and N. Adam. 2014. "A distributed scatterer interferometry approach for precision monitoring of known surface deformation phenomena." *IEEE Trans. Geosci. Remote Sens.*, 52 (9): 5454–5468. IEEE. <https://doi.org/10.1109/TGRS.2013.2289370>.
- Golabi, K., R. B. Kulkarni, and G. B. Way. 1982. "A Statewide Pavement Management System." *Interfaces (Providence)*, 12 (6): 5–21. <https://doi.org/10.1287/inte.12.6.5>.
- Gong, H., Y. Sun, X. Shu, and B. Huang. 2018. "Use of random forests regression for predicting IRI of asphalt pavements." *Constr. Build. Mater.*, 189: 890–897. Elsevier Ltd. <https://doi.org/10.1016/j.conbuildmat.2018.09.017>.
- Guo, F., G. Zhang, Q. Zhang, R. Zhao, M. Deng, and K. Xu. 2018. "Speckle Suppression by Weighted Euclidean Distance Anisotropic Diffusion." *Remote Sens. 2018, Vol. 10, Page 722*, 10 (5): 722. Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/RS10050722>.
- Haider, S. W., G. Y. Baladi, and K. Chatti. 2011. "Impact of pavement surface monitoring frequency on pavement management decision making." *T DI Congr. 2011 Integr. Transp. Dev. a Better Tomorrow - Proc. 1st Congr. Transp. Dev. Inst. ASCE*, 656–666. American Society of Civil Engineers.
- Halevy, A., P. Norvig, and F. Pereira. 2009. "The Unreasonable Effectiveness of Data." *IEEE Intell. Syst.*, B. Brannon, ed. IEEE Computer Society.
- Haralick, R. M., I. Dinstein, and K. Shanmugam. 1973. "Textural Features for Image Classification." *IEEE Trans. Syst. Man Cybern.*, SMC-3 (6): 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>.
- Hecht-Nielsen, R. 1987. "Kolmogorov's mapping neural network existence theorem." *1st IEEE Int. Jt. Conf. Neural Networks*. San Diego, CA.
- Herold, M., and D. Roberts. 2005. "Spectral characteristics of asphalt road aging and deterioration: Implications for remote-sensing applications." *Appl. Opt.*, 44 (20): 4327–4334. <https://doi.org/10.1364/AO.44.004327>.

- Herold, M., D. Roberts, V. Noronha, and O. Smadi. 2008. "Imaging spectrometry and asphalt road surveys." *Transp. Res. Part C Emerg. Technol.*, 16 (2): 153–166. <https://doi.org/10.1016/j.trc.2007.07.001>.
- Herold, M., D. Roberts, O. Smadi, and V. Noronha. 2004. "Road condition mapping with hyperspectral remote sensing." *Aviris 2004*, (August 2016): 1–15.
- Hicks, G., A. Simpson, and J. Groeger. 2011. *Pavement Management Practices in State Highway Agencies: Newington, Connecticut Peer Exchange Results*.
- Hoppe, E., B. Bruckno, E. Campbell, S. Acton, A. Vaccari, M. Stuecheli, A. Bohane, G. Falorni, and J. Morgan. 2014. "Transportation Infrastructure Monitoring Using Satellite Remote Sensing." *Transp. Res. Arena 5th Conf. Transp. Solut. from Res. to Deploy*.
- Hornik, K., M. Stinchcombe, and H. White. 1989. "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2 (5): 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- Hossain, M., L. Gopiseti, and S. Miah. 2020. "Artificial neural network modelling to predict international roughness index of rigid pavements." *Int. J. Pavement Res. Technol.*, 13: 229–239.
- ICEYE. 2021. "Satellite Data." Accessed March 17, 2021. <https://www.iceye.com/satellite-data#C2>.
- Inkoom, S., J. Sobanjo, A. Barbu, and X. Niu. 2019. "Pavement Crack Rating Using Machine Learning Frameworks: Partitioning, Bootstrap Forest, Boosted Trees, Naïve Bayes, and K - Nearest Neighbors." *J. Transp. Eng. Part B Pavements*, 145 (3): 04019031. American Society of Civil Engineers (ASCE). <https://doi.org/10.1061/JPEODX.0000126>.
- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning*. Springer.
- Jaybhay, J., and R. Shastri. 2015. "A Study of Speckle Noise Reduction Filters." *Signal Image Process. An Int. J.*, 6 (3). <https://doi.org/10.5121/sipij.2015.6306>.
- Kamalzadeh, H., and M. Hahsler. 2019. "POMDP: Introduction to Partially Observable Markov Decision Processes." Accessed February 7, 2021. <https://cran.r-project.org/web/packages/pomdp/vignettes/POMDP.html>.
- Kang, M., M. Kim, and J. H. Lee. 2010. "Analysis of Rigid Pavement Distresses on Interstate Highway Using Decision Tree Algorithms." *KSCE J. Civ. Eng.*, 14 (2): 123–130. <https://doi.org/10.1007/s12205-010-0123-7>.
- Kargah-Ostadi, N. 2014. "Comparison of Machine Learning Techniques for Developing Performance Prediction Models Nima." *Comput. Civ. Build. Eng.*, 1222–1229. <https://doi.org/10.1061/9780784413616.053>.

- Kargah-Ostadi, N., and S. M. Stoffels. 2015. "Framework for development and comprehensive comparison of empirical pavement performance models." *J. Transp. Eng.*, 141 (8). [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000779](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000779).
- Karimzadeh, S., and M. Matsuoka. 2020. "Remote Sensing X-Band SAR Data for Land Subsidence and Pavement Monitoring." *Sensors*, 20 (4751).
- Kaya, O., H. Ceylan, S. Kim, D. Waid, and B. P. Moore. 2020. "Statistics and Artificial Intelligence-Based Pavement Performance and Remaining Service Life Prediction Models for Flexible and Composite Pavement Systems." *Transp. Res. Rec. J. Transp. Res. Board*, 036119812091588. <https://doi.org/10.1177/0361198120915889>.
- Kim, J. W., G. B. Choi, and J. M. Lee. 2018. "A POMDP framework for integrated scheduling of infrastructure maintenance and inspection." *Comput. Chem. Eng.*, 112: 239–252. Elsevier Ltd. <https://doi.org/10.1016/j.compchemeng.2018.02.015>.
- Kobbacy, K. A. H. 2012. "Application of Artificial Intelligence in maintenance modelling and management." *IFAC Proc. Vol.*, 45 (31): 54–59. IFAC. <https://doi.org/10.3182/20121122-2-ES-4026.00046>.
- Koch, C., K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth. 2015. "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure." *Adv. Eng. Informatics*, 29 (2): 196–210. Elsevier Ltd. <https://doi.org/10.1016/j.aei.2015.01.008>.
- Kupidura, P. 2019. "The comparison of different methods of texture analysis for their efficacy for land use classification in satellite imagery." *Remote Sens.*, 11 (10). <https://doi.org/10.3390/rs11101233>.
- Kurniawati, H., D. H. Wee, and S. Lee. 2008. "SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces." *Robot. Sci. Syst.* Zurich, Switzerland.
- Lapedes, A., and R. Farber. 1988. *Abstract: How Neural Nets Work. papers.nips.cc*.
- Li, J., G. Yin, X. Wang, and W. Yan. 2022. "Automated decision making in highway pavement preventive maintenance based on deep learning." *Autom. Constr.*, 135: 104111. Elsevier. <https://doi.org/10.1016/J.AUTCON.2021.104111>.
- Li, M., A. Faghri, A. Ozden, and Y. Yue. 2017. "Economic feasibility study for pavement monitoring using synthetic aperture radar-based satellite remote sensing cost-benefit analysis." *Transp. Res. Rec.*, 2645: 1–11. <https://doi.org/10.3141/2645-01>.
- Li, N., R. Haas, and W.-C. Xie. 1997. "Investigation of Relationship Between Deterministic and Probabilistic Prediction Models in Pavement Management." *Transp. Res. Rec. J. Transp. Res. Board*, 1592 (1): 70–79. National Research Council. <https://doi.org/10.3141/1592-09>.

- Li, Z. 2018. *Transportation Asset Management Methodology and Applications*. Boca Raton, London: CRC Press.
- Lin, J., J.-T. Yau, and L.-H. Hsiao. 2003. "Correlation Analysis Between International Roughness Index (IRI) By." *Transp. Res. Board 82th Annu. Meet.*, (January): 1–21.
- Ma, Y., M. Elseifi, N. Dhakal, M. Bashar, and Z. Zhang. 2021. "Non-Destructive Detection of Asphalt Concrete Stripping Damage Using Ground Penetrating Radar." *Transp. Res. Rec.*
- Madanat, S., and M. Ben-Akiva. 1994. "Optimal Inspection and Repair Policies for Infrastructure Facilities." *Transp. Sci.*, 28 (1): 55–62. Institute for Operations Research and the Management Sciences (INFORMS). <https://doi.org/10.1287/trsc.28.1.55>.
- Marcelino, P., M. de Lurdes Antunes, E. Fortunato, and M. Castilho Gomes. 2019. "Machine learning approach for pavement performance prediction." *Int. J. Pavement Eng.*, 0 (0): 1–14. Taylor & Francis. <https://doi.org/10.1080/10298436.2019.1609673>.
- Marcelino, P., M. de Lurdes Antunes, E. Fortunato, and M. C. Gomes. 2020. "Transfer learning for pavement performance prediction." *Int. J. Pavement Res. Technol.*, 13 (2): 154–167. <https://doi.org/10.1007/s42947-019-0096-z>.
- Masters, T. 1993. *Practical Neural Networks Recipes in C++*. Book. Morgan Kaufmann.
- Mazari, M., and D. D. Rodriguez. 2016. "Prediction of pavement roughness using a hybrid gene expression programming-neural network technique." *J. Traffic Transp. Eng. (English Ed.)*, 3 (5): 448–455. Elsevier Ltd. <https://doi.org/10.1016/j.jtte.2016.09.007>.
- Medury, A., S. Madanat, and M. Asce. 2014. "Simultaneous Network Optimization Approach for Pavement Management Systems." *ascelibrary.org*, 20 (3): 04014010. American Society of Civil Engineers (ASCE). [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000149](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000149).
- Mei, A., R. Salvatori, N. Fiore, A. Allegrini, and A. D'Andrea. 2014. "Integration of field and laboratory spectral data with multi-resolution remote sensed imagery for asphalt surface differentiation." *Remote Sens.*, 6 (4): 2765–2781. <https://doi.org/10.3390/rs6042765>.
- Memarzadeh, M., and M. Pozzi. 2016. "Integrated Inspection Scheduling and Maintenance Planning for Infrastructure Systems." *Comput. Civ. Infrastruct. Eng.*, 31 (6): 403–415. Blackwell Publishing Inc. <https://doi.org/10.1111/mice.12178>.
- Memarzadeh, M., M. Pozzi, and J. Zico Kolter. 2015. "Optimal Planning and Learning in Uncertain Environments for the Management of Wind Farms." *J. Comput. Civ. Eng.*, 29 (5): 04014076. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000390](https://doi.org/10.1061/(asce)cp.1943-5487.0000390).
- Mettas, C., K. Themistocleous, K. Neocleous, A. Christofe, K. Pilakoutas, and D. Hadjimitsis. 2015. "Monitoring asphalt pavement damages using remote sensing techniques." *Third Int. Conf. Remote Sens. Geoinf. Environ.*, 9535 (June 2015): 95350S. <https://doi.org/10.1117/12.2195702>.

- Meyer, F. J. 2019. “Spaceborne Synthetic Aperture Radar – Principles, Data Access, and Basic Processing Techniques.” *SAR Handb. Compr. Methodol. For. Monit. Biomass Estim.*, A. Flores, K. Herndon, R. Thapa, and E. Cherrington, eds. NASA.
- Meyer, F. J., O. A. Ajadi, and E. Hoppe. 2020. “Studying the Applicability of X-Band SAR Data to the Network-Scale Mapping of Pavement Roughness on US Roads.” *Remote Sens.*, 12 (1507).
- Michigan Department of Transportation. 2017. *Asset Management Background International Roughness Index (IRI)*.
- MnDNR. 2022. “Minneapolis/St. Paul Climate Data Snow Data Resources.” *Minnesota Dep. Nat. Resour.* Accessed February 21, 2022. https://www.dnr.state.mn.us/climate/twin_cities/snowfall.html.
- MnDOT. 2011. *MnDOT Pavement Distress Identification Manual*.
- MnDOT. 2019. *2019 Pavement Condition Annual Report*.
- Mohammadi, M. 2012. “Road Classification and Condition Determination Using Hyperspectral Imagery.” *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, XXXIX-B7 (September): 141–146. <https://doi.org/10.5194/isprsarchives-XXXIX-B7-141-2012>.
- Munawar, H. S., A. W. A. Hammad, and S. T. Waller. 2021. “A review on flood management technologies related to image processing and machine learning.” *Autom. Constr.*, 132: 103916. Elsevier. <https://doi.org/10.1016/J.AUTCON.2021.103916>.
- Murdzek, R., H. Malik, and A. Leśniak. 2018. “The use of the DInSAR method in the monitoring of road damage caused by mining activities.” *E3S Web Conf.*, 36. <https://doi.org/10.1051/e3sconf/20183602005>.
- Nitsche, P., R. Stütz, M. Kammer, and P. Maurer. 2012. “Comparison of Machine Learning Methods for Evaluating Pavement Roughness Based on Vehicle Response.” *J. Comput. Civ. Eng.*, 28 (4): 04014015. [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000285](https://doi.org/10.1061/(asce)cp.1943-5487.0000285).
- Noronha, V., M. Herold, D. Roberts, and M. Gardner. 2002. “Spectrometry and Hyperspectral Remote Sensing for Road Centerline Extraction and Evaluation of Pavement Condition.” *Natl. Consort. Remote Sens. Transp. - Infrastruct. Manag.*, (August): 12.
- Nussbaum, S., I. Niemeyer, and M. J. Canty. 2006. “SEath - a New Tool for Automated Feature Extraction in the Context of Object-Based Image Analysis.” *1st Int. Conf. Object-based Image Anal. (OBIA 2006)*, XXXV (July 2006): 4/C42.
- Oliver, C., and S. Quegan. 1997. “Understanding synthetic aperture radar images.” *Norwood, MA Artech House, 1997.*, 53 (9): 1689–1699.
- Osorio-Lird, A., A. Chamorro, C. Videla, S. Tighe, and C. Torres-Machi. 2018. “Application of

- Markov chains and Monte Carlo simulations for developing pavement performance models for urban network management.” *Struct. Infrastruct. Eng.*, 14 (9): 1169–1181. Taylor and Francis Ltd. <https://doi.org/10.1080/15732479.2017.1402064>.
- Ozbay, K., and R. Laub. 2001. *Models for Pavement Deterioration Using LTPP. NJDOT*.
- Ozden, A., A. Faghri, M. Li, and K. Tabrizi. 2016. “Evaluation of Synthetic Aperture Radar Satellite Remote Sensing for Pavement and Infrastructure Monitoring.” *Procedia Eng.*, 145: 752–759. Elsevier B.V. <https://doi.org/10.1016/j.proeng.2016.04.098>.
- Pan, Y., X. Zhang, X. Jin, H. Yu, J. Rao, S. Tian, L. Luo, and C. Li. 2016. “Road pavement condition mapping and assessment using remote sensing data based on MESMA.” *IOP Conf. Ser. Earth Environ. Sci.*, 34 (1). <https://doi.org/10.1088/1755-1315/34/1/012023>.
- Papadimitriou, E., and A. Theofilatos. 2017. “Meta-analysis of crash-risk factors in freeway entrance and exit areas.” *J. Transp. Eng. Part A Syst.*, 143 (10): 1–10. <https://doi.org/10.1061/JTEPBS.0000082>.
- Papakonstantinou, K. G., C. P. Andriotis, and H. Gao. 2019. “Quantifying the value of structural monitoring for decision making.” *13th Int. Conf. Appl. Stat. Probab. Civ. Eng.*, 1–8. Seoul, South Korea.
- Papakonstantinou, K. G., C. P. Andriotis, and M. Shinozuka. 2018. “POMDP and MOMDP solutions for structural life-cycle cost minimization under partial and mixed observability.” *Struct. Infrastruct. Eng.*, 14 (7): 869–882. Taylor & Francis. <https://doi.org/10.1080/15732479.2018.1439973>.
- Papakonstantinou, K. G., and M. Shinozuka. 2014a. “Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part I: Theory.” *Reliab. Eng. Syst. Saf.*, 130: 202–213. Elsevier. <https://doi.org/10.1016/j.ress.2014.04.005>.
- Papakonstantinou, K. G., and M. Shinozuka. 2014b. “Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation.” *Reliab. Eng. Syst. Saf.*, 130: 214–224. Elsevier. <https://doi.org/10.1016/j.ress.2014.04.006>.
- Parrilli, S., M. Poderico, C. V. Angelino, and L. Verdoliva. 2012. “A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage.” *IEEE Trans. Geosci. Remote Sens.*, 50 (2): 606–616. <https://doi.org/10.1109/TGRS.2011.2161586>.
- Pell, R., R. Svoboda, R. Eagar, P. Ondko, and F. Kirschnick. 2015. *Effective Infrastructure Asset Management- A holistic approach to transformation. Prism*.
- Pierce, L., and N. D. Weitzel. 2019. *NCHRP SYNTHESIS 531 Automated Pavement Condition Surveys*. Transportation Research Board.
- Pineau, J., G. Gordon, and S. Thrun. 2003. “Point-based value iteration: An anytime algorithm for

- POMDPs.” *Int. Jt. Conf. Artif. Intell.*, 1025–1032.
- Plati, C., and A. Loizos. 2013. “Estimation of in-situ density and moisture content in HMA pavements based on GPR trace reflection amplitude using different frequencies.” *J. Appl. Geophys.*, 97: 3–10. Elsevier. <https://doi.org/10.1016/j.jappgeo.2013.04.007>.
- Ripley, B. D. 2007. *Pattern Recognition via Neural Networks*. Cambridge university press.
- Rodriguez-Lozano, F. J., F. León-García, J. C. Gámez-Granados, J. M. Palomares, and J. Olivares. 2020. “Benefits of ensemble models in road pavement cracking classification.” *Comput. Civ. Infrastruct. Eng.*, mice.12543. Wiley. <https://doi.org/10.1111/mice.12543>.
- Salchenberger, L. M., E. M. Cinar, and N. A. Lash. 1992. “Neural Networks: A New Tool for Predicting Thrift Failures.” *Decis. Sci.*, 23 (4): 899–916. <https://doi.org/10.1111/j.1540-5915.1992.tb00425.x>.
- Salehi, H., and R. Burgueño. 2018. “Emerging artificial intelligence methods in structural engineering.” *Eng. Struct.*, 171 (May): 170–189. Elsevier. <https://doi.org/10.1016/j.engstruct.2018.05.084>.
- Sayers, M. W. 1995. “On the calculation of International Roughness Index from Longitudinal Road Profile.” *Transp. Res. Rec.*, 1501.
- Schnebele, E., B. F. Tanyu, G. Cervone, and N. Waters. 2015. “Review of remote sensing methodologies for pavement management and assessment.” *Eur. Transp. Res. Rev.*, 7 (2). <https://doi.org/10.1007/s12544-015-0156-6>.
- Schöbi, R., and E. N. Chatzi. 2016. “Maintenance planning using continuous-state partially observable Markov decision processes and non-linear action models processes and non-linear action models.” *Struct. Infrastruct. Eng.*, 2479: 1–18. Taylor & Francis. <https://doi.org/10.1080/15732479.2015.1076485>.
- Seites-Rundlett, W., M. Z. Bashar, C. Torres-Machi, and R. B. Corotis. 2021. “Combined Evidence Model to Enhance Pavement Condition Prediction from Highly Uncertain Sensor Data.” *Reliab. Eng. Syst. Saf.*, 217 (May 2021): 108031. Elsevier Ltd. <https://doi.org/10.1016/j.res.2021.108031>.
- Shahi, K., H. Z. M. Shafri, and A. Hamedianfar. 2017. “Road condition assessment by OBIA and feature selection techniques using very high-resolution WorldView-2 imagery.” *Geocarto Int.*, 32 (12): 1389–1406. Taylor & Francis. <https://doi.org/10.1080/10106049.2016.1213888>.
- Shahi, K., H. Z. M. Shafri, E. Taherzadeh, S. Mansor, and R. Muniandy. 2015. “A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery.” *Egypt. J. Remote Sens. Sp. Sci.*, 18 (1): 27–33. Authority for Remote Sensing and Space Sciences. <https://doi.org/10.1016/j.ejrs.2014.12.003>.
- Smallwood, R. D., and E. J. Sondik. 1973. “The Optimal Control of Partially Observable Markov

- Processes Over a Finite Horizon on JSTOR.” *Oper. Res.*, 21 (5): 1071–1088.
- Smith, T., and R. Simmons. 2006. *Focused Real-Time Dynamic Programming for MDPs: Squeezing More Out of a Heuristic*.
- Smith, T., and R. Simmons. 2012. *Point-Based POMDP Algorithms: Improved Analysis and Implementation*.
- Sollazzo, G., T. F. Fwa, and G. Bosurgi. 2017. “An ANN model to correlate roughness and structural performance in asphalt pavements.” *Constr. Build. Mater.*, 134: 684–693. Elsevier Ltd. <https://doi.org/10.1016/j.conbuildmat.2016.12.186>.
- Sondik, E. J. 1978. “The Optimal Control of Partially Observable Markov Processes Over the Infinite Horizon: Discounted Costs.” *Oper. Res.*, 26 (2): 282–304.
- Spaan, M. T. J., and N. Vlassis. 2005. *Perseus: Randomized Point-based Value Iteration for POMDPs*. *J. Artif. Intell. Res.*
- Spencer, B. F., V. Hoskere, and Y. Narazaki. 2019. “Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring.” *Engineering*, 5 (2): 199–222. Chinese Academy of Engineering. <https://doi.org/10.1016/j.eng.2018.11.030>.
- Suanpaga, W., and K. Yoshikazu. 2010. “Riding quality model for asphalt pavement monitoring using phase array type L-band synthetic aperture radar (PALSAR).” *Remote Sens.*, 2 (11): 2531–2546. <https://doi.org/10.3390/rs2112531>.
- Tong, Z., J. Gao, A. Sha, L. Hu, and S. Li. 2018. “Convolutional Neural Network for Asphalt Pavement Surface Texture Analysis.” *Comput. Civ. Infrastruct. Eng.*, 33 (12): 1056–1072. John Wiley & Sons, Ltd. <https://doi.org/10.1111/MICE.12406>.
- Tong, Z., J. Gao, and H. Zhang. 2017. “Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural networks.” *Constr. Build. Mater.*, 146: 775–787. Elsevier Ltd. <https://doi.org/10.1016/j.conbuildmat.2017.04.097>.
- TRIP. 2018. “Key Facts About America’s Surface Transportation System and Federal Funding.” 2018.
- Tsai, Y.-C. J., and F. Li. 2012. “Critical Assessment of Detecting Asphalt Pavement Cracks under Different Lighting and Low Intensity Contrast Conditions Using Emerging 3D Laser Technology.” [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000353](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000353).
- Tsai, Y. C. (James), Y. Zhao, B. Pop-Stefanov, and A. Chatterjee. 2021. “Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning.” *Int. J. Pavement Res. Technol.*, 14 (4): 487–495. <https://doi.org/10.1007/s42947-020-0138-5>.
- Üstün, B., W. J. Melssen, and L. M. C. Buydens. 2006. “Facilitating the application of Support Vector Regression by using a universal Pearson VII function based kernel.” *Chemom. Intell.*

- Lab. Syst.*, 81 (1): 29–40. Elsevier. <https://doi.org/10.1016/j.chemolab.2005.09.003>.
- Wang, K. C. P., Q. J. Li, G. Yang, Y. Zhan, and Y. Qiu. 2015. “Network level pavement evaluation with 1 mm 3D survey system.” *J. Traffic Transp. Eng. (English Ed.)*, 2 (6): 391–398. Periodical Offices of Chang-an University. <https://doi.org/10.1016/j.jtte.2015.10.005>.
- Wang, K. C. P., A. Zhang, J. Q. Li, Y. Fei, C. Chen, and B. Li. 2017. “Deep Learning for Asphalt Pavement Cracking Recognition Using Convolutional Neural Network.” *Airf. Highw. Pavements*, 166–177. <https://doi.org/10.1061/9780784480922.015>.
- Wang, K., and O. Smadi. 2011. *Automated Imaging Technologies for Pavement Distress Surveys*.
- Warner, T. 2011. “Kernel-Based Texture in Remote Sensing Image Classification.” *Geogr. Compass*, 5 (10): 781–798. <https://doi.org/10.1111/j.1749-8198.2011.00451.x>.
- Yamany, M. S., T. U. Saeed, M. Volovski, and A. Ahmed. 2020. “Characterizing the Performance of Interstate Flexible Pavements Using Artificial Neural Networks and Random Parameters Regression.” *J. Infrastruct. Syst.*, 26 (2). [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000542](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000542).
- Yıldırım, S. 2020. “Hyperparameter Tuning for Support Vector Machines - C and Gamma Parameters.” *towardsdatascience*. Accessed July 26, 2020. <https://towardsdatascience.com/hyperparameter-tuning-for-support-vector-machines-c-and-gamma-parameters-6a5097416167>.
- Zeiada, W., S. A. Dabous, K. Hamad, R. Al-Ruzouq, and M. A. Khalil. 2020. “Machine Learning for Pavement Performance Modelling in Warm Climate Regions.” *Arab. J. Sci. Eng.*, 45 (5): 4091–4109. Springer Berlin Heidelberg. <https://doi.org/10.1007/s13369-020-04398-6>.
- Zhang, Z., C. Sun, R. Bridgelall, and M. Sun. 2018. “Application of a Machine Learning Method to Evaluate Road Roughness from Connected Vehicles.” *J. Transp. Eng. Part B Pavements*, 144 (4): 04018043. <https://doi.org/10.1061/jpeodx.0000074>.
- Ziari, H., M. Maghrebi, J. Ayoubinejad, and S. T. Waller. 2016a. “Prediction of pavement performance: Application of support vector regression with different kernels.” *Transp. Res. Rec.*, 2589 (2589): 135–145. <https://doi.org/10.3141/2589-15>.
- Ziari, H., J. Sobhani, J. Ayoubinejad, and T. Hartmann. 2016b. “Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods.” *Int. J. Pavement Eng.*, 17 (9): 776–788. Taylor & Francis. <https://doi.org/10.1080/10298436.2015.1019498>.