

**Harnessing Visual Analytics for Enhanced Winter Road Safety: A Framework for  
Continuous Road Surface Friction Estimation**

by

Qian Xie

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

TRANSPORTATION ENGINEERING

Department of Civil and Environmental Engineering  
University of Alberta

© Qian Xie, 2023

## **Abstract**

Winter Road Surface Condition (RSC) monitoring currently relies on qualitative descriptors such as "bare lane," "partially snow-covered," and "fully snow-covered. These descriptors present two inherent problems—their subjective nature, which gives rise to measurement inconsistencies, and the difficulty for road users to discern between safe and unsafe roads due to the broad spectrum of conditions intermediate RSC classes cover. Friction-based measurements, ranging from 0 to 1, alleviate these issues as they provide a much more objective measure. Despite this, the large-scale implementation of such measurements has long been impeded by high collection costs. Even with the incorporation of friction values, the existing RSC monitoring system is often plagued by limited spatial coverage and infrequent updates. This can result in potential discrepancies between reported and actual conditions, thus compromising the reliability and usability of the system.

To overcome these challenges, this thesis capitalizes on the advancements in computing capabilities and sophisticated machine learning techniques. The primary focus is to repurpose the information-rich image datasets and to enable continuous monitoring of dynamic winter RSCs. This strategy ensures a consistent flow of reliable and timely information for maintenance personnel and road users. At the core of this novel approach is a comprehensive framework designed to convert winter road surface images into friction values. This framework synergistically integrates three essential components—an image-based friction model, a friction interpolator, and a binary collision model. Together, they form a robust system aimed at addressing the prevalent issues in winter RSC monitoring for improved winter road safety.

For the development of the image-based friction prediction model, 128 friction measurements and their corresponding road surface images were collected over select roads in the city of Edmonton.

Since the images themselves could not serve as direct inputs, feature extraction techniques were used to summarize the information found within the images. These techniques encompass RSC labeling into four classes (bare lane, one-track bare, two-track bare, and fully snow-covered), image thresholding, Local Binary Pattern (LBP), and Gray Level Co-occurrence Matrix (GLCM). With these extracted features, three tree-based algorithms; namely, decision tree, Random Forest (RF), and Gradient Boosting (GB), were used to model the relationship between the friction values and the extracted feature. All three tree-based models displayed robust performance based on standard statistical measures such as RMSE and RMSPE.

In an effort to improve the spatial coverage of RSC monitoring, the friction interpolator was subsequently constructed. Continuous friction values were generated by feeding collected road surface images into the friction model. Ten datasets were then created by varying the distance between available observation points from 100 m to 1000 m per observation. A total of six interpolators were evaluated on these datasets, including Ordinary Kriging (OK), Regression Kriging (RK), Random Forest (RF), RFOK, Random Forest Spatial Interpolator (RFSI), and RFSIOK. Among these interpolators, OK yielded the highest accuracy and displayed the least sensitivity across nearly all observation distances. An additional case study was conducted using the City of Edmonton's traffic camera locations as a real-world scenario to demonstrate the applicability of the proposed model. This case study assesses the potential for using traffic camera data to interpolate friction measurements. Identical to the sensitivity analysis, OK was found to be the top-performing interpolator, achieving commendable accuracy with only five friction measurements as input.

To underscore the potential of the proposed framework in enhancing road safety, a binary collision likelihood model was developed by utilizing continuous friction measurements. After examining

segment lengths ranging from 500 m to 20 km, the 6.5 km model was determined to offer the optimal balance between interpretability and accuracy. The developed model demonstrated robust performance leveraging Annual Average Daily Traffic (AADT) and friction as predictors. Furthermore, an additional analysis was conducted assuming the availability of Connected Vehicles (CV) and appropriate sensing technologies. This analysis took a more pragmatic approach to assess potential safety benefits by evaluating the predictive capabilities of the generated continuous friction measurements in relation to collision events. The findings offered a tangible demonstration of the potential safety improvements that could be realized using the proposed framework.

## **Preface**

Some of the work presented in this thesis has been accepted, published, or is under review in various journals related to Transportation Engineering.

### **Journal Papers Accepted/Published**

1. **Xie, Q., & Kwon, T. J.** (2022). Development of a Highly Transferable Urban Winter Road Surface Classification Model: A Deep Learning Approach. *Transportation Research Record*, 2676(10), 445–459. <https://doi.org/10.1177/03611981221090235>
2. **Xie, Q., & Kwon, T. J.** (2023). Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images – A City of Edmonton, Canada, Case Study. *Canadian Journal of Civil Engineering*. Just-IN. <https://doi.org/10.1139/cjce-2023-0015>

### **Journal Papers Under Preparation**

1. **Xie, Q., & Kwon, T. J.** A Framework for Determining Collision Likelihood Using Continuous Friction Values in a Connected Vehicles Environment

## **Acknowledgment**

I would like to express my sincere gratitude to my research supervisor, Dr. Tae J. Kwon, for his invaluable guidance and support throughout my time as an MSc Student. He has been a constant source of inspiration and encouragement and has always challenged me to achieve higher standards of academic excellence. His expertise and feedback have greatly improved the quality and clarity of my work. I am also grateful to the professors who served as members of my thesis defense committee: Dr. Karim El-Basyouny, Dr. Stephen Wong, and Dr. Yuxiang Chen. They provided me with constructive feedback and insightful suggestions that helped me improve the quality of my work. Finally, I would like to thank my research lab mates Mingjian Wu, Simita Biswas, Xueru Ding, Zehua Shuai, Nancy Huynh, and Michael Urbiztondo, as well as my family for their support during this challenging period.

# Table of Contents

Abstract.....	ii
Preface.....	v
Acknowledgment.....	vi
Table of Contents.....	vii
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations.....	xi
Chapter 1 Introduction.....	1
1.1 Background.....	1
1.2 Problem Statement and Motivation.....	3
1.3 Research Objective.....	4
1.4 Thesis Structure.....	5
Chapter 2 Literature Review.....	7
2.1 Road Surface Friction.....	7
2.2 Spatial Interpolators.....	12
2.3 Relationship Between Friction and Collisions.....	14
2.4 Summary.....	16
Chapter 3 Methodology.....	18
3.1 Component One: Image-Based Friction Model.....	19
3.2 Component Two: Spatial Interpolator.....	31
3.3 Component Three: Binary Collision Likelihood Model.....	37
3.4 Summary.....	38
Chapter 4 Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images.....	40
4.1 Friction Model Development.....	40
4.2 Summary.....	49
Chapter 5 Assessing Interpolation Methods and Collision Models for Road Safety Enhancements.....	51

5.1	Evaluation and Development of Spatial Interpolators .....	51
5.2	Collision Model Development for Safety Assessment .....	57
5.3	Summary .....	63
Chapter 6	Conclusion and Future Work.....	65
6.1	Overview .....	65
6.2	Research Findings .....	66
6.3	Research Contributions .....	68
6.4	Limitations and Future Research.....	70
References	.....	72
Appendix	.....	77



## List of Tables

Table 2.1 Factors That Impact Road Friction (Adapted from Hall et al., 2009) .....	8
Table 3.1 RSC Classification System (Adapted From Wu et al. (2022)) .....	22
Table 3.2 Common Theoretical Semivariograms .....	34
Table 5.1 Potential Number of Collisions Reduced Over Four Days Within the Study Area .....	62
Table 5.2 Potential Collision Reduction Over Four Days City-wide .....	63

## List of Figures

Figure 2.1 Two Main Factors That Impact Road Friction (Hall et al., 2009).....	7
Figure 2.2 Relationship Between Friction and Tire Slip .....	9
Figure 3.1 Proposed Methodological Framework .....	18
Figure 3.2 Study Area and Route Selected.....	20
Figure 3.3 Original Image Versus Transformed Image .....	21
Figure 3.4 Number of Images in Each RSC Class.....	22
Figure 3.5 CNN Model Architecture .....	23
Figure 3.6 Original Image Versus Thresholded Image.....	25
Figure 3.7 Generated Continuous Friction Values and Selected Segment for Interpolation.....	31
Figure 3.8 Training Data With 500-Meter Separation Distance.....	32
Figure 3.9 Typical Semivariogram Structure.....	35
Figure 4.1 Decision Tree Model Prediction vs. Observed Friction Values .....	41
Figure 4.2 Internal Logic of Decision Tree Model.....	41
Figure 4.3 Feature Importance of Decision Tree Model.....	43
Figure 4.4 Friction Model Performance Comparison .....	44
Figure 4.5 Random Forest and Gradient Boosting Feature Importance .....	45
Figure 4.6 Selected CNN Architecture .....	47
Figure 4.7 Validation Accuracy Difference Between the Developed CNN Models .....	47
Figure 4.8 CNN Model Accuracy on Friction Dataset Images.....	48
Figure 4.9 Prediction Accuracy Before and After Inclusion of CNN Model .....	49
Figure 5.1 Interpolator Performance Comparison .....	51
Figure 5.2 Interpolator Performance at 300 m Separation Distance.....	53
Figure 5.3 Interpolator Performance at 900 m Separation Distance.....	53
Figure 5.4 ML Hybrid Model Performance at 300 m Separation Distance .....	54
Figure 5.5 Friction Measurements Obtained at Locations with Traffic Cameras.....	54
Figure 5.6 Traffic Camera Interpolation Curves.....	55
Figure 5.7 Traffic Camera Interpolator Error Comparison.....	56
Figure 5.8 Model Validation Accuracy at Various Road Segment Lengths.....	58
Figure 5.9 Decision Tree Model Structure With 6.5 km Segment Length .....	59
Figure 5.10 Proposed Framework Implementation in CV and ITS Environment .....	61

## List of Abbreviations

<i>1-D</i>	One Dimensional
<i>3-D</i>	Three Dimensional
<i>AADT</i>	Annual Average Daily Traffic
<i>ANN</i>	Artificial Neural Network
<i>ASM</i>	Angular Second Moment
<i>BPT</i>	British Pendulum Tester
<i>CNN</i>	Convolutional Neural Network
<i>CV</i>	Connected Vehicle
<i>DFT</i>	Dynamic Friction Tester
<i>FC</i>	Fully Connected
<i>GB</i>	Gradient Boosting
<i>GLCM</i>	Gray Level Co-occurrence Matrix
<i>IDW</i>	Inverse Distance Weight
<i>ITS</i>	Intelligent Transportation System
<i>LBP</i>	Local Binary Pattern
<i>MAE</i>	Mean Absolute Error
<i>ML</i>	Machine Learning
<i>MLR</i>	Multi Linear Regression
<i>OK</i>	Ordinary Kriging
<i>PI</i>	Polynomial Interpolator
<i>RF</i>	Random Forest
<i>RFOK</i>	Random Forest Ordinary Kriging
<i>RFOK</i>	Random Forest Ordinary Kriging
<i>RFSI</i>	Random Forest Spatial Interpolator
<i>RFSIOK</i>	Random Forest Spatial Interpolator Ordinary Kriging
<i>RK</i>	Regression Kriging
<i>RMSE</i>	Root Mean Squared Error
<i>RMSPE</i>	Root Mean Squared Percentage Error
<i>Rpart</i>	Regression Tree
<i>RSC</i>	Road Surface Condition
<i>RSI</i>	Road Surface Index

<i>RST</i>	Road Surface Temperature
<i>RWIS</i>	Road Weather Information System
<i>SEM</i>	Structural Equation Modeling
<i>SPF</i>	Safety Performance Function
<i>SS</i>	Semantic Segmentation
<i>SVM</i>	Support Vector Machines
<i>TPS</i>	Thin Plate Splines

# Chapter 1 Introduction

## 1.1 Background

Road Surface Condition (RSC) monitoring is essential for jurisdictions that experience winter weather. The information gathered can be used by maintenance personnel for targeted treatment and disseminated to the public for advance travel planning. The former provides better-conditioned roads, and the latter allows road users to avoid dangerous road sections, both of which have the potential to improve road safety.

Most commonly, RSC information is provided in the form of qualitative descriptions (e.g., Alberta, CA, and Iowa, U.S., use a three-category system of bare, partially snow-covered, and fully snow-covered to describe the condition of the road surface (*511 Alberta, 2023; Iowa DOT 511, 2023*)). The problem with labeling or describing pavement condition is that it is subjective in nature. Different people will have different opinions on what qualifies as a particular category, leading to measurement discrepancies that make the information disseminated less reliable. Moreover, generalizing all winter conditions into a limited number of categories creates situations where a single category simultaneously represents two conflicting safety levels. Such is the case with the partially snow-covered class mentioned above, where the amount of grip provided varies over a wide range (Wu & Kwon, 2022). When road users are confronted with this type of ambiguous condition, it becomes unclear what the appropriate response should be. A solution to this problem is changing the surrogate measure to friction. The advantage of this change is that it eliminates the need to make subjective interpretations, as friction is a measured variable made by a device. It also removes the safety ambiguity inherent in qualitative, category-based systems by using the direct relationship between friction and collisions to find friction thresholds where collision risk is most pronounced (Abohassan et al., 2021).

Although friction is beneficial in many ways, it is not widely used due to the financial constraints associated with the collection process. The technology to effortlessly collect friction value as the vehicle travels already exists; however, the equipment cost associated is too high for mass implementation. Comparatively cheaper options are also available, but these devices are much more labor-intensive due to their relatively time-consuming measuring procedure and the need for

traffic control (Hall et al., 2009). For friction to become more widely adopted, equipment costs must be reduced while minimizing labor demand in the collection process.

Changing the surrogate measure used to represent RSC is only half the problem; the spatial coverage of RSC must also be extended. From our observation, RSC is assumed to be constant over long distances. This assumption may need to be corrected according to the existing literature, where it has been found that conditions tend to vary even in short stretches (Gu, 2019). Additionally, there is also an issue with update frequency. Depending on the jurisdiction in question, the update frequency can be as low as once or twice per day. As a result, the information presented to road users may not reflect the actual conditions they may encounter.

Two conditions must be met in order to address the above shortcomings: first, an automated, low-cost method of friction measurement must be developed; and second, there must be a way to convert point measurements to continuous measurements without additional investment in increasing measurement density and patrol frequency. Researchers have tried to develop methods that satisfy one of these two conditions but never both. Attempts at developing friction models have shown that there are several viable ways of estimating friction values. Cameras can be used to capture images of the road surface, which can then be converted to friction using a Machine Learning (ML) model (Du et al., 2020). Vehicle tires can be embedded with force sensors to measure parameters needed to estimate friction using mathematical friction models (Matilainen & Tuononen, 2011). There are also methods that use the vehicle's existing sensors to estimate friction through mathematical formulation (Khaleghian et al., 2017). Among these methods, the use of road surface images to estimate friction is particularly promising because it has several advantages over other methods; first, it is much easier to implement than tire sensors, and second, its implementation is not limited to vehicles; it can be applied to any infrastructure that has imaging capabilities. However, there is a significant research gap associated with image-based friction models: existing studies have focused solely on bare conditions; it is unknown whether friction values can be accurately estimated for winter road conditions.

In the field of developing interpolators to estimate friction, existing efforts revolve around using Regression Kriging (RK) to interpolate a friction-like measure called Road Surface Index (RSI). Although these studies produced encouraging results, several environmental studies have shown

that Random Forest (RF) and RFOK (hybrid model that combines RF and Ordinary Kriging) have shown exceptional performance in interpolating variables like mud sea content, solar flare, etc. (Leirvik & Yuan, 2021; Li et al., 2011; Sekulić et al., 2020). In these studies, the accuracy of RF and RFOK exceeds geostatistical interpolators like RK and other ML methods. The fact that RF and RFOK have not been explored in friction interpolation and road weather interpolation is a missed opportunity, i.e., a gap that needs to be addressed.

## 1.2 Problem Statement and Motivation

Based on the above discussions, it is clear that both a low-cost friction model and a friction interpolator are needed to solve the problems found in the existing winter RSC monitoring system. In order to develop effective components that will enable this improvement, three research questions need to be answered.

The first research question is, can friction values be accurately predicted using winter road surface images? As mentioned previously, friction models have only been developed for bare asphalt conditions. These methods used image textures and RSC descriptions as predictors of friction values. Whether these features can be used to accurately predict winter friction values is unknown. Nevertheless, there are papers that propose a framework for estimating winter friction values, but the authors were unable to validate their proposed method due to a lack of actual friction measurements (Roychowdhury et al., 2018; Wu & Kwon, 2022). As a result, the question of whether friction values can be estimated from winter road surface images remains unanswered.

The second research question is, what is the best interpolator for friction? RK is the only interpolator that has been evaluated in the existing literature. It is unknown how it compares to Ordinary Kriging (OK), which has been shown to outperform RK when there is no trend in the data (Zhu & Lin, 2010). It is also unclear whether the superior performance shown by RF models persists when the target variable changes to friction. Likewise, what happens when RF is combined with OK? Overall, there is a need to compare the performance of RK with other high-performing interpolators.

Finally, can continuous friction values be used to identify the occurrence of winter collisions? Assuming that both an image-based friction model and a friction interpolator have been

successfully developed, presenting the output continuous friction values without modification may confuse road users and maintenance personnel unfamiliar with the concept of friction. Therefore, a simpler way to present this information is needed. One solution is to develop a binary collision model—using friction as a predictor—to predict whether a collision is likely or unlikely. Such a model has the ability to improve road safety by providing advance warnings to road users and informing maintenance personnel where treatment is needed without additional interpretation. The development of a collision model with friction as an input has been explored in the existing literature. However, similar to image-based friction models, most efforts have focused on bare asphalt conditions, where the consensus is that lower friction values are associated with more collisions. In terms of winter collision modeling, it appears that only one author has attempted to model winter collisions using friction as a predictor (Abohassan et al., 2021, 2022). Furthermore, the current approach in collision modeling relies on parametric models that make a distribution assumption. It is possible that using non-parametric machine learning methods could enhance the performance of the collision model. Ultimately, additional research is required to substantiate the correlation between friction and winter collisions through modeling and to investigate the feasibility of utilizing machine learning for friction-collision modeling.

### **1.3 Research Objective**

This thesis contains a primary and a secondary objective. The primary objective is to develop a framework for generating continuous road friction values from winter road surface images. Then, to demonstrate how the proposed framework can be used to improve road safety, a secondary objective was formulated: to construct a binary collision likelihood model that provides advance warnings to road users regarding dangerous road sections. Three components are required to achieve these two objectives; the first two are designed to meet the primary objective and answer the first two research questions, and the third component answers the third research question while satisfying the secondary objective. The specifics of each component and what they aim to accomplish are listed below:

#### Component One: Image-based Friction Model

- Produce a friction model that maintenance personnel can apply to estimate friction coefficients from winter road surface images.



- Quantify and rank the viability of different image features on their ability to predict road friction values.
- Evaluate and validate the relationship between winter road surface conditions and road friction values.
- Compare the performance of three different tree-based algorithms (decision tree, random forest, and gradient boosting) in their ability to estimate friction values from road surface images.

#### Component Two: Friction Interpolator

- Evaluate the performance of Machine Learning (ML) models for interpolating friction coefficients.
- Compare the interpolation performance between ML models, geostatistical methods, and hybrid models that combines ML with geostatistical methods.
- Determine the feasibility of using traffic cameras to generate continuous friction measurements.

#### Component Three: Binary Collision Model

- Provide a novel collision likelihood modeling framework that uses interpolators to determine the optimal aggregation length.
- Validate the relationship between friction coefficient and winter collision occurrences through modeling.
- Evaluate the viability of using ML to model the friction-collision relationship.
- Quantify the expected savings of the proposed framework in a connected vehicle and intelligent transportation system environment.

## 1.4 Thesis Structure

This thesis contains six chapters. Excluding the introduction chapter, it is organized as follows:

In chapter two, the literature review section, readers will find a summary of the physics of friction, an overview of measurement methods commonly used in industry, and prototypes developed in previous research. Additionally, a summary of current research efforts in the interpolation of

friction and environmental variables is also provided. This section concludes with a discussion of the relationship between friction and collision occurrence.

Chapter Three is the Methodology section. The organization of this chapter follows the order of the three components. This chapter first explains the data and methods used to develop the image-based friction model, then the friction interpolator, and finally, the binary collision model.

Chapter Four covers the development of an automated image-based friction model. The prediction accuracy of the model, the internal logic of the model, and the relative importance of the features are presented in this chapter.

Chapter Five compares the performance of six interpolators on datasets with increasing distance between measurements to determine the best-performing friction interpolator. Also included in this chapter is the development of a binary collision model based on segment length calibration and its associated safety benefits.

The last chapter, chapter six, highlights this thesis's main findings and research contributions, along with its limitations and potential future research directions.

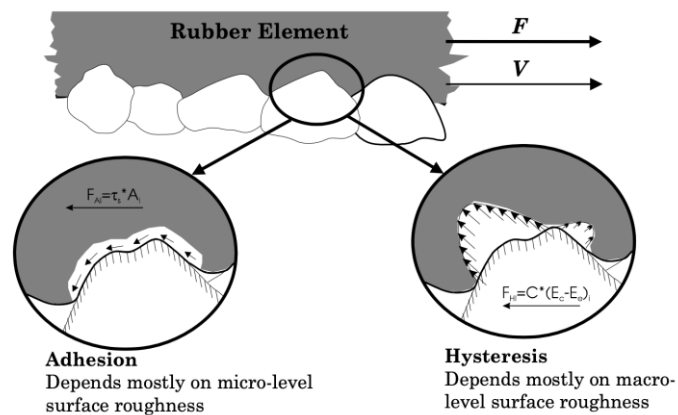
## Chapter 2 Literature Review

Chapter two aims to provide an in-depth overview of the theory behind friction and how it can be measured. The measurement methods discussed are divided into two categories: commonly used industry-accepted methods and prototype methods proposed in the literature. In addition, this chapter includes a comprehensive review of existing studies in spatial interpolation and friction-based collision modeling. A summary of the key points and existing research limitations can be found at the end of this chapter.

### 2.1 Road Surface Friction

#### Friction Overview

Road surface friction is the amount of resistant force between the tires and the road surface; it is what allows road users to manipulate and control their vehicles (Hall et al., 2009). The specific amount of friction force generated by the tires is a product of two mechanisms: adhesion and hysteresis. Adhesion is the force generated from the tires interlocking with the road surface, and hysteresis occurs as a result of the tires compressing after engaging with the road surface, resulting in energy being stored within the tires. When the tire relaxes, some of this stored energy is lost as heat, which produces a frictional force. Other factors also impact pavement friction; however, they are considered insignificant compared to adhesion and hysteresis (Hall et al., 2009). **Figure 2.1** illustrates the main mechanisms that impact road friction.



**Figure 2.1 Two Main Factors That Impact Road Friction (Hall et al., 2009)**

These two forces act as two different kinds of road textures. Adhesion depends on micro textures, whereas hysteresis depends on macro textures. As their name suggests, micro and macro textures operate at different scales. Macro textures can be thought of as the uneven pavement texture from the material (gravel, sand, etc.) that make up the pavement surface, and micro textures are the unevenness that exists on a single material element.

**Factors Affecting Pavement Friction**

Other than macro and micro textures, several other factors influence the two friction forces. The most impactful parameters are listed in **Table 2.1**.

**Table 2.1 Factors That Impact Road Friction (Adapted from Hall et al., 2009)**

<b>Road Surface Characteristics</b>	<b>Tire Parameters</b>	<b>Environmental Factors</b>	<b>Vehicle Parameters</b>
<ul style="list-style-type: none"> <li>• Micro texture</li> <li>• Macro texture</li> </ul>	<ul style="list-style-type: none"> <li>• Tread design and condition</li> <li>• Tire pressure</li> </ul>	<ul style="list-style-type: none"> <li>• Snow and ice</li> <li>• Water</li> </ul>	<ul style="list-style-type: none"> <li>• Vehicle speed</li> <li>• Braking method</li> </ul>

***Tire Parameters***

Tire tread patterns and material usage affect the amount of friction provided (Hall et al., 2009). In wet conditions, certain tread patterns can expel water trapped at the tire/road interface, allowing the tires to have more contact with the pavement. Similarly, tires designed for winter conditions have deeper and wider grooves to dissipate snow and slush. As the tire ages, the tread depth decreases, resulting in poorer performance. As for tire pressure, a poorly inflated tire is prone to concaving. Instead of expelling water through the treads, water is now trapped under the tire. (Hall et al., 2009).

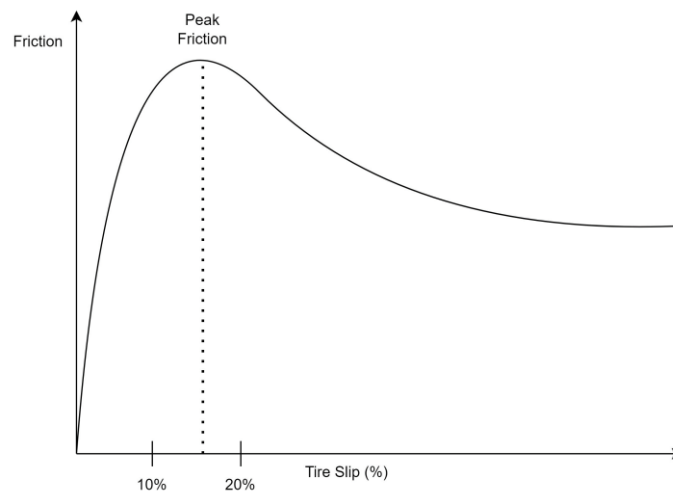
***Environmental Factors***

Water, snow, and ice all act as lubricants in the tire-pavement interface (Hall et al., 2009). The effect of water is most pronounced at higher speeds during hydroplaning, where the vehicle’s tires float on the layer of water. Since there is no contact between the tires and the road, friction is close to zero. Snow and ice have the same effect—they prevent contact between the tires and the road

surface (Hall et al., 2009). In general, the greater the presence of snow and ice, the lower the coefficient of friction (Wallman & Åström, 2001).

### ***Vehicle Parameters***

The speed of the vehicle and the braking method control the slip ratio (Hall et al., 2009). When the vehicle is rolling freely, the tire and vehicle speeds are the same. In this case, the slip ratio is 0. During braking, the tire speed can rotate at a slower speed than the vehicle speed. Depending on the difference in speed, the coefficient of friction will vary. Peak friction is reached between 10 and 20% slip ratio, beyond which friction decreases until the sliding friction coefficient is reached (achieved at 100% slip ratio) (Hall et al., 2009). **Figure 2.2** illustrates this relationship.



**Figure 2.2 Relationship Between Friction and Tire Slip**

### **Methods for Measuring Friction**

Friction measurement methods fall into two categories: generally accepted methods that have been rigorously tested, and prototype methods developed by researchers that have only been validated in the confine of the research itself.

### ***Existing Testing Methods***

Friction measurement devices currently used by jurisdictions can be split into four categories: locked-wheel, deceleration-based, slip-based, and portable. Locked-wheel and slip-based methods require the attachment of additional equipment to the vehicle in the form of a small trailer or a

third tire (Al-Qadi et al., 2002). In the locked-wheel method, the vehicle travels at a constant speed of 40 mph (64 km/h), at which point the driver can choose to lock the device wheel to collect point measurements (Al-Qadi et al., 2002). In comparison, the slip-based methods collect friction values continuously by measuring the rotational resistances of smooth tires at a specific slip ratio (Lee et al., 2019). However, its ability to collect continuous measurements comes at the cost of complex data processing and high maintenance fees (Lee et al., 2019).

Deceleration-based methods measure changes in velocity via an in-vehicle device. The procedure involves the vehicle reaching a certain speed, then fully initiating the breaks (Al-Qadi et al., 2002). During the breaking process, the initial velocity and braking time are recorded, from which friction can be calculated by dividing the deceleration rate by gravity. This method has been shown to be simple and accurate; however, it only provides point measurements and requires traffic control measures (Al-Qadi et al., 2002).

The last type of method is portable friction testers. Devices in this category include the British Pendulum Tester (BPT) and Dynamic Friction Tester (DFT) (Hall et al., 2009). BPT measures friction based on how high the pendulum swings after its rubber shoe slides across the pavement surface. The results generated using this method are said to be highly dependent on operator procedure and wind effects. On the other hand, DFT is an automated device that measures friction through a spinning disk that hovers over the road surface. The spinning disk is spun at a constant tangential speed of 55 mph (88 km/h); it is then lowered to make contact with the road surface. The force generated between the disk and the pavement is measured and used to determine friction (Hall et al., 2009).

### ***Prototype Methods***

Researchers over the years have attempted to develop more efficient and cost-effective methods of estimating friction. According to Khaleghian et al. (2017), existing research can be either experimental or model-based. Experimental-based methods attempt to find the correlation between sensor data and friction, whereas model-based methods attempt to mathematically model friction.

In the field of experimental-based methods, researchers have attempted to use optical sensors to estimate friction. One researcher, Du et al. (2020), collected closeup images of bare asphalt

surfaces and classified them into two friction levels (qualified and unqualified resistance). They extracted features using Local Binary Pattern (LBP), Gray Level Co-occurrence Matrix (GLCM), color histograms, and Convolutional Neural Network (CNN), then built models using RF and Artificial Neural Network (ANN) algorithms to predict the friction level. The ANN model achieved 90.67% accuracy, while the RF model achieved 76%. Another study was done by Roychowdhury et al. (2018), who proposed a two-stage framework to label winter road images by friction. Stage one used a CNN to filter out dry asphalt images (assumed to be high friction). Stage two divided the image into quadrants and applied a logistic regression model to estimate snow likelihood. Then, based on the snow likelihood of each column, the image was labeled as high, medium, or low friction. Using this approach, the authors reported an 89.5% road friction estimation accuracy. Note that it is unknown how this accuracy value was obtained as the author did not have actual friction values.

Yang et al. (2021) differed from the previous two studies by attempting to predict friction using regression rather than categorization. They used texture analysis and Multi-Linear Regression (MLR) on 3D scanned images of bare asphalt at different depths. The models with the highest R-squared values of 0.87 and 0.88 were for 60 km/h and 10 km/h braking speeds, respectively, at 2.5 mm depth. Wu et al. (2022) also tried to predict friction coefficients. They made a system to turn winter road images into a Road Surface Index (RSI)—a surrogate of friction. Images were inputted into a CNN RSC model to label images as bare, partially snow-covered, and fully snow-covered. An upper and lower RSI limit would restrict the RSI value depending on which RSC category the image belongs to. Then, using image thresholding to separate pixels into foreground and background values, the precise RSI can be determined using the formula  $(RSI_{max} - (RSI_{max} - RSI_{min}) \cdot \% \text{ background values})$ . In terms of prediction accuracy, because the author did not collect any friction values, it is not possible to examine the validity of their proposed method.

Excluding optical sensors, tire sensors have also been used to estimate friction. Erdogan et al. (2011) installed a piezoelectric sensor to measure tread deflections; the estimated deflection was then inputted into the brush model (a mathematical model) to estimate friction. Likewise, Matilainen et al. (2011) installed force sensors on the vehicle tie rods to measure the tire moment conditions needed in the brush model for friction estimation. Researchers have also embedded tires

with accelerometers to estimate tire load, which is necessary to estimate friction through formulation and track vibration differences resulting from slippery surfaces (Khaleghian et al., 2016; Niskanen & Tuononen, 2015). Overall, the tire sensor approach tends to focus on estimating the necessary parameters found in a mathematical model instead of directly using it to predict friction.

In terms of pure model-based methods, they do not require additional sensors that the vehicle does not already have (Khaleghian et al., 2017). One group of models in this category is the vehicle dynamic based-methods; these models contain two types of parameters—measurable and unmeasurable; through formulation, the measurable are used to estimate the unmeasurable. An example of a model that belongs in this class is the wheel dynamic model, which focuses on the motion of a single wheel. Another is the bicycle model that combines the rear and front wheels to create a two-wheel model. There are also other model categories like the tire-based and slip-slope-based methods, where the former centers around tire forces and slip ratios/angles, and the latter estimate friction using the slip curve (Khaleghian et al., 2017).

## **2.2 Spatial Interpolators**

The spatial interpolator aims to reduce the required data collection efforts and optimize the use of collected friction values. Although certain friction estimation methods can gather continuous measurements, employing workers to drive through a jurisdiction's entire road network to collect data regularly is not financially feasible. A solution is to use spatial interpolators to generate continuous friction readings from limited point measurements. In the existing literature, three types of methods display the most promising performance: geostatistical-based, Machine Learning-based (ML-based), and hybrid models that combine geostatistical with ML-based methods.

### **Road Weather Interpolators**

Wu et al. (2022) used Kriging, a geostatistical method that interpolates data based on spatial correlation, to estimate Road Surface Temperature (RST) and Road Surface Index (RSI) from point measurements collected by Road Weather Information System (RWIS) stations. The objective of the study was to evaluate whether Kriging can be used to estimate the values between point measurements collected by RWIS stations. The input data for RST interpolation were



collected by mobile RWIS stations, with portions removed to mimic point measurements collected by stationary RWIS. By using Regression Kriging (RK), the study reported RMSE values ranging from 0.237 to 0.805, depending on the number of measurements available and the time of collection. Similarly, the study applied the same data processing to simulate stationary RWIS measurements for RSI, resulting in an RMSE value of approximately 0.15.

Gu et al. (2019) conducted a similar investigation using geostatistical methods to interpolate RSI and RST. The input data used in this project were also collected from mobile RWIS. However, unlike Wu et al., who focused solely on Kriging, multiple interpolation methods were evaluated. These include Inverse Distance Weighting (IDW), Polynomial Interpolation (PI), and Thin Plate Splines (TPS). Performance was evaluated by randomly excluding 30% of the available data for testing. Among these methods, Regression Kriging (RK) produced the lowest average RMSE error of 0.231. For RSI, only RK was used for interpolation and produced RMSE between 0.004 and 0.169, depending on the time of measurement and the highway segment in question. After determining that RK was the best-performing interpolator, stationary RWIS point measurements were artificially created by hiding a portion of the mobile RWIS's RST data for validation. An RMSE between 0.52 and 0.88 was obtained via RK. This evaluation was performed only for RST, as stationary RWIS stations did not collect RSI data.

### **ML and Hybrid Interpolators**

Outside of road weather-related research, there has been a growing interest in using machine learning (ML) for interpolations. Li et al. (2011) compared several ML methods with traditional statistical interpolators for predicting mud sea content. They tested 23 models, including Random Forest (RF), Support Vector Machine (SVM), Regression Tree (RPart), Ordinary Kriging (OK), IDW, and hybrid methods combining two algorithms. The top five models were RFOK, RF, RFIDW (2nd power), RpartOK, and IDW (3rd power). They concluded that RF captured the spatial trend well and performed better than Kriging and IDW.

Leivik et al. (2021) also looked into using ML algorithms in making spatial interpolations. Their study evaluated the performance of RF, RK, and OK at interpolating solar flare radiation across the globe. The results show that for virtually all continents, RF produced the lowest error; in most cases, the next most accurate model had double the error of RF. For example, in North America,

RF had an estimation error of 9.69 Mean Absolute Error (MAE), whereas the next accurate model—a form of RK—had an MAE of 12.91. Hence, the author described RF as this study’s most accurate spatial interpolator.

Recognizing the strength of RF, Sekulic et al. (2020) created a modified RF model called Random Forest Spatial Interpolator (RFSI). The difference between this and traditional RF is that two additional covariates were added: “values at nearby locations” and “distance to these measurements.” The reasoning is that these two features allow RF to learn the similarity between neighboring values. Using precipitation and temperature as the target variables, the author compared the performance of RFSI with RF, RK, and IDW. Overall, it was found that RFSI was the superior interpolator by a small margin.

### **2.3 Relationship Between Friction and Collisions**

The relationship between friction and collisions is complicated, as friction is not the only factor that influences road safety. Researchers have attempted to shed light on this relationship through empirical evidence by comparing the proportion of collisions that occurred on low-friction road surfaces. The Road Research Laboratory in London (1965) compared the friction level between sites where collisions occurred and a series of randomly selected locations. It was found that sites linked to collisions had a lower mean friction value than the randomly selected ones; 0.45 versus 0.60. In addition, the number of sites with friction below 0.45 could only be found in sites linked to collisions. A similar study was conducted by Schulze et al. (1977), where they observed that the friction distribution of collision sites differed from randomly selected sites; it was found that the majority of sites found at friction value below 0.40 were from collision sites. Another observation was made in the same study that accident rates increased as friction decreased.

Researchers have also made efforts to statistically model the relationship between friction and collisions. Kuttesch (2004) developed a crash rate model using data collected in Virginia, U.S. The author developed several models using an aggregation length of one mile with friction only, AADT only, and friction and AADT together. In all three models, friction was found to be significant; however, by itself, it could only explain 0.4% of the total variation. AADT, on the other hand, was able to explain 29.9% while also being significant. The combined model posted the best results with an R-squared value of 0.3122, with both parameters being significant. In all three models, the

coefficient of friction was negative, which supports the idea that lower frictions result in more collisions. McCarthy et al. (2016) also developed a model using AADT and friction, where both AADT and friction were determined to be significant parameters that improved model performance. The authors developed three models for three different types of roads, and in all three models, friction had a negative coefficient. In addition to these two studies, many other researchers have examined the friction-collision relationship and concluded that the lower the friction, the higher the number of collisions (Hall et al., 2009; Lee et al., 2019).

Up to this point, all of the studies discussed focused exclusively on non-winter conditions. The number of studies examining the relationship between friction and collisions during the winter is far fewer than in bare/wet conditions. According to Wallman et al. (2001), these studies generally compare the collision rate on various winter road conditions or are maintenance studies investigating the effect of anti-icing treatments. The results showed that the barer the road surface, the lower the accident rate. One notable Norwegian study found that the collision rate was eight times higher at friction coefficients below 0.15 than at 0.35 or higher (Wallman & Åström, 2001).

In terms of researchers trying to model the relationship between winter collisions and friction, Abohassan et al. (2021) performed one of the few studies that used friction coefficients as a predictor of winter collision frequency. In their study, weather and maintenance variables, road surface friction, and road and traffic characteristics were assumed to be predictors of collision frequency. Using Structural Equation Modeling (SEM), the authors found that friction has a direct causal relationship with collision frequency; collisions decrease when friction increases. On the other hand, maintenance operations and weather data affect collision frequency through friction as a medium, i.e., maintenance operations and weather are a predictor of friction, which is a predictor of collisions. In a subsequent study, Abohassan et al. (2022) attempted to determine a safety threshold for when collisions are more likely to occur. They created three separate collision frequency models with three different ranges: less than 0.35, between 0.35 and 0.60, and greater than 0.60. They found that friction values less than 0.35 resulted in a significant increase in collisions, values between 0.35 and 0.60 resulted in an insignificant increase in collisions, and values greater than 0.60 resulted in a significant decrease in collisions. Based on these results, the authors recommended 0.35 as a safety threshold. Zhao et al. (2022) also used friction for collision modeling based on year-round data; however, the target variable was collision severity instead of

collision frequency. Other variables like roadway, traffic, and driver characteristics were included as predictors. A total of four models were evaluated: logit model, SVM, ANN, and XGBoost, among which logit, ANN, and XGBoost had near-identical accuracy of around 70%.

## 2.4 Summary

This section has provided an overview of the mechanics behind road friction and methods of measuring friction. In general, road friction is affected by several factors, including road surface texture, vehicle and tire parameters, and environmental characteristics. Existing methods that account for these factors require cumbersome measurement procedures that provide only point measurements. While more complex methods can measure friction continuously, they involve using devices that are bulky, more complicated to operate, and costly to maintain. Because of the limitations associated with existing methods, it is difficult to collect friction data on a large scale.

In order to make the collection process more efficient and cost-effective, researchers have developed prototype methods that convert data collected by sensors into friction values using mathematical, statistical, and ML models. Among these proposed methods, optical sensors in the form of cameras have shown high potential for friction estimation when used in combination with ML models. The advantage of using cameras over other methods is that they are non-intrusive; they do not require embedding sensors to collect vehicle dynamic data. Furthermore, they are not limited to in-vehicle implementation; any roadside infrastructure equipped with a camera gains the ability to collect friction values, a property that makes optical sensors the most suitable for large-scale implementation. Nevertheless, most of the studies that revolve around image-based models focus solely on bare conditions. The few that have incorporated winter images have done so without any validation data; hence, whether an accurate friction model can be developed for winter conditions is unknown.

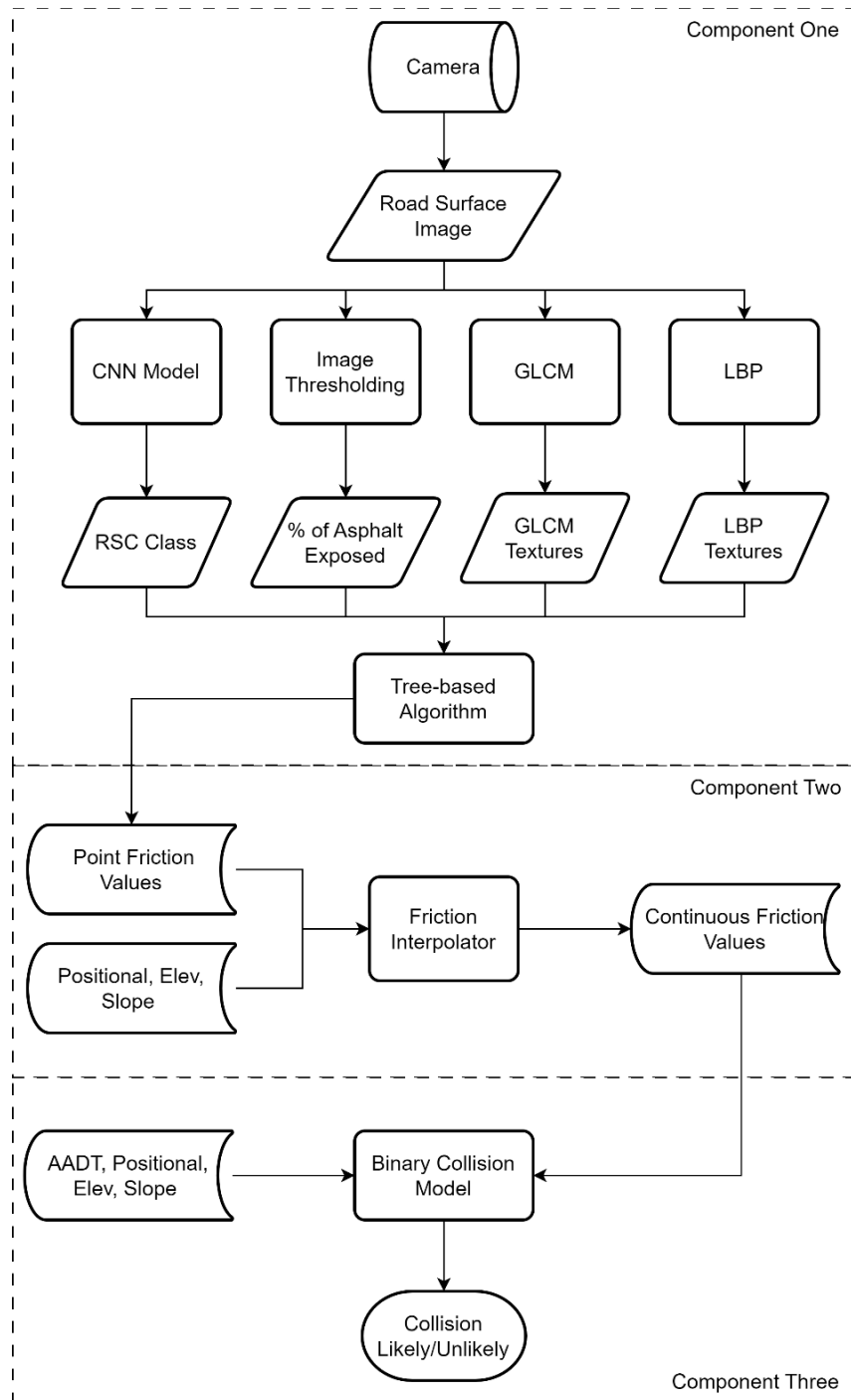
On top of providing a summary of existing friction collection methods, the necessity of spatial interpolation and the relationship between friction and collisions were also discussed. Spatial interpolators complement the friction collection process by allowing point measurements to be converted to continuous measurements, which significantly reduces the data collection effort required to collect continuous measurements. Researchers have attempted to investigate the performance of interpolators for road-related variables, which has shown Kriging to be the most

accurate method. For non-road-related variables, RF, an ML method, has demonstrated higher estimation accuracy than Kriging. However, its performance on road weather variables is unknown.

Regarding the relationship between friction and collisions, the consensus from examining empirical evidence is that collisions are linked to low-friction road surfaces. A similar observation has also been made through modeling, where friction has a negative coefficient, meaning that lower friction leads to more collisions. That being said, most of the studies examining this relationship have been conducted under bare asphalt conditions; the number of studies examining winter conditions is somewhat lacking. Furthermore, the development of these collision models uses the parametric Safety Performance Function (SPF) approach, which makes an assumption about the data distribution that can potentially negatively affect model performance. It is possible that by using ML to model the relationship between friction and collisions, higher performance can be achieved as ML models are non-parametric and, therefore, better able to model the relationship between friction and collisions.

## Chapter 3 Methodology

This chapter explains the specifics of the techniques involved in our proposed framework, an illustration of which is shown in **Figure 3.1**.



**Figure 3.1 Proposed Methodological Framework**

As illustrated in **Figure 3.1**, the proposed framework consists of three components. The first component, the image-based friction model, focuses on the conversion of road surface images into point friction values. Road surface images collected from cameras are fed into a series of feature extraction techniques to convert the images into predictive features. These techniques include a CNN model to label the RSC class, image thresholding to determine the degree of asphalt exposure, and Local Binary Pattern (LBP) and Gray Level Co-occurrence Matrix (GLCM) to extract generic image textures that describe the pixel patterns found in the image. The extracted features are then fed into a friction model built using a tree-based algorithm to output a point friction value. In the second component, the output point measurements and supporting auxiliary variables (position, elevation, and slope data) are fed into a friction interpolator to convert the point measurements into continuous ones. The final component of the framework uses the friction-collision relationship to determine whether a collision is likely or unlikely using Average Annual Daily Traffic (AADT) and other supporting variables.

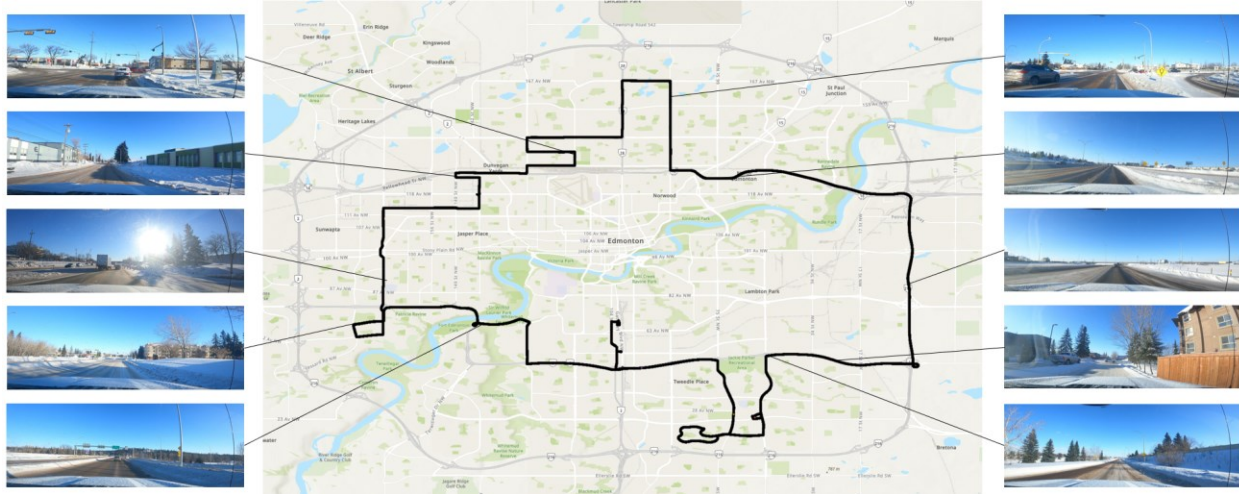
### **3.1 Component One: Image-Based Friction Model<sup>1</sup>**

#### ***Friction Model Data***

Data collection for this work occurred in the City of Edmonton, Alberta, Canada, during its 2021/2022 winter season. In total, 128 friction tests were performed during daytime over three days: Jan 18, 19, and 25<sup>th</sup>, with temperatures of -21.2°C, -21.6 °C, and -1.2°C, respectively. These tests were taken after plowing and anti-icing treatments in response to snowfall events on Jan 17 (7.3 mm) and 24<sup>th</sup> (2.4 mm). Nighttime data were not collected as city personnel does not perform traction testing at night. **Figure 3.2** depicts the study area and routes visited for data collection.

---

<sup>1</sup> Contents in Section 3.1 has been published in Xie, Q., & Kwon, T. J. (2023). Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images – A City of Edmonton, Canada, Case Study. Canadian Journal of Civil Engineering. Just-IN. <https://doi.org/10.1139/cjce-2023-0015> and Xie, Q., & Kwon, T. J. (2022) and Development of a Highly Transferable Urban Winter Road Surface Classification Model: A Deep Learning Approach. Transportation Research Record, 2676(10), 445–459. <https://doi.org/10.1177/03611981221090235>



**Figure 3.2 Study Area and Route Selected**

Road friction values were collected using Vericom VC4000 (a deceleration-based device) installed within a Ford F150 (City of Edmonton fleet vehicle). The testing procedure requires the driver to reach a speed of 30 km/h, at which point the driver fully initiates the brakes until the vehicle comes to a complete stop. Based on the device's internal accelerometer, the device records the average longitudinal G force experienced by the vehicle during the breaking process as the friction coefficient. In addition, to provide road surface imagery alongside friction tests, a GoPro Hero 9 camera was mounted inside the vehicle to record road surface footage. The advantage of using a GoPro is that each recorded video file is attached with metadata containing GPS and speed information, which are critical in linking the road surface image with its corresponding friction value. We need these two pieces of information because the friction measuring device records time to the nearest minute. Instead of physically watching the recorded footage to locate the exact time of the friction test, speed changes can be used to pinpoint the precise testing time. In essence, within the minute interval, if the speed drops significantly by 2 m/s or greater (signifying that the driver slammed the brakes), this point in time would be when the friction testing started. Using this method, 128 images that correspond with the 128 friction tests were extracted.

With the extracted images, a technique called perspective transformation (Roychowdhury et al., 2018) was applied to remove irrelevant environmental features. Like cropping, perspective transformation removes everything but the area of interest; however, it is unique in that it shifts the perspective of the road surface to a top-down view, which improves the visibility of the road surface located farther away in the image. In this study, we used OpenCV (Bradski, 2000) to derive



a transformation matrix that warps the cropped image into a 300 x 300 square image. An example of this transformation is shown in **Figure 3.3**.



Original Image



Processed Image

**Figure 3.3 Original Image Versus Transformed Image**

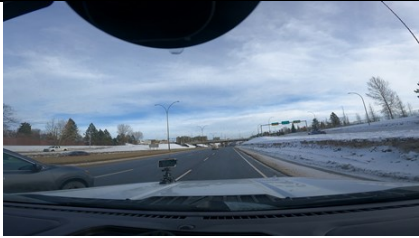



### **Image Feature Extraction**

#### ***Road Surface Condition (RSC) Classification***

The processed image cannot be used directly for friction modeling. Instead, images must first be translated into numerical features that summarize the image's content. One way to accomplish this task is to classify the images using an RSC classification system. RSC classes that represent a lesser degree of precipitation can be interpreted to have higher friction than one with more precipitation due to greater asphalt exposure. Researchers have also suggested a connection between RSC and road surface friction but could not validate the relationship due to the lack of friction data (Roychowdhury et al., 2018; Wu & Kwon, 2022). As a result of the need to validate said relationship and the intuitive nature between the two measures, the RSC category was chosen as one of the predictor features.

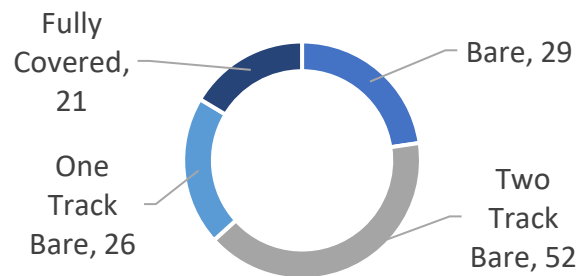
In this study, a four-class system described in **Table 3.1** was selected. This classification system is based on what Wu et al. (2022) used in their study with one modification: partially snow-covered was further segregated into two-track and one-track bare under the assumption that the location and extent of asphalt exposure affect friction values. Since asphalt provides more friction than snow/ice, having exposed asphalt under both tires offers more grip than only one tire. Likewise, there are also cases where the exposed asphalt is under neither tire, or the width of the exposed asphalt is too narrow to allow for sufficient contact between tire and asphalt.

**Table 3.1 RSC Classification System (Adapted From Wu et al. (2022))**

Sample Image	RSC Description	RSC Category
	Driving lane is free of contaminants. Asphalt is completely exposed	Bare
	Both tire tracks in the driving lane are visible and free of contaminants	Two-track Bare
	Tire tracks are not entirely free of contaminants. This is an intermediate case between two-track bare and fully snow-covered	One-track Bare
	Driving lane is completely covered in snow. Asphalt is not visible	Fully snow covered

In accordance with this four-class system, the extracted images were manually classified into the four classes during model development to prevent misclassifications from affecting performance.

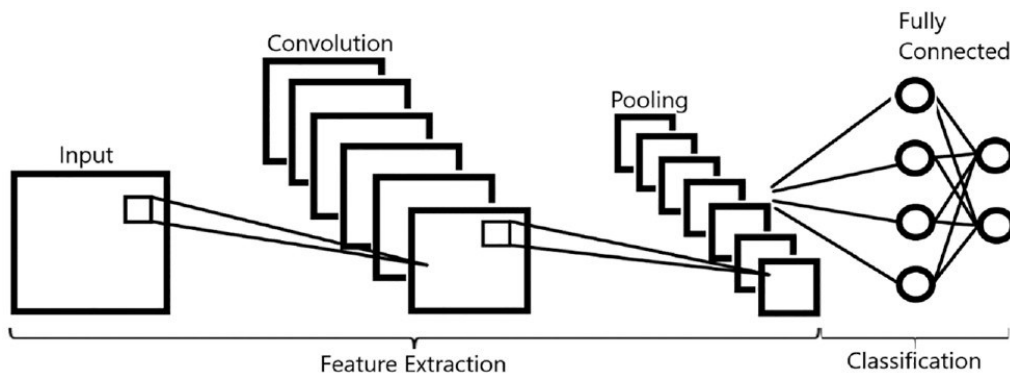
**Figure 3.4** provides a breakdown of the number of images in each class.



**Figure 3.4 Number of Images in Each RSC Class**

For model implementation, a Convolutional Neural Network was selected to automate the RSC classification process due to its high performance in existing literature (Wu & Kwon, 2022; Xie & Kwon, 2022). A typical CNN model can be broken down into two components—feature extraction and classification. In the feature extraction step, images are first processed into three-dimensional (3-D) image arrays, where each dimension represents a channel of the RGB color spectrum. Following this step, image arrays will go through a series of convolutional and pooling layers. Convolutional layers use filters to extract feature maps such as edge location, color patterns, and so forth, whereas pooling layers reduce the matrix size (height and width) of the feature maps. Images will typically go through the convolutional layer first, and then be passed through the pooling layer to reduce matrix size. In most cases, the outputted feature maps will continue to go through additional convolutional and pooling layers until the matrix size is sufficiently small, at which point, it is converted into a one-dimensional (1-D) array and passed into a Fully Connected (FC) layer.

The FC layer is where classification occurs; it is typically the final step in most CNN models. Unlike the feature extraction layers, there are no convolutional or pooling layers. FC is entirely made up of dense layers, with each dense layer composed of smaller units called neurons. The dense layer receives the 1-D feature map and converts it into a probability vector with the same number of entries as the number of classes. By looking for the highest probability in the output vector, one can determine which class the image belongs to. An example of a generic CNN network is illustrated in **Figure 3.5**.



**Figure 3.5 CNN Model Architecture**

There are two ways of training a CNN model, one is to train the model from scratch, and the other is to use transfer learning. The benefit of training a model from scratch is that the model is tailored

to a specific purpose. Given that there are a sufficient amount of training data, the model can become highly accurate. However, the training speed can be slow depending on the number of parameters in the model. In comparison, transfer learning builds upon previously learned knowledge. This process is based on the idea that features captured in the initial layers are generic features, and features captured deeper in the CNN are more dataset-specific (Kensert et al., 2019). Hence, the training process involves taking a pre-trained model and training only parameters found in the deeper layers. Since fewer parameters need to be calibrated, the model can be developed faster and with a smaller dataset than what is normally required.

Before training the model, the image dataset was split 80/20, with 80% of data allocated to training and 20% allocated to validation. In addition, the image array was normalized to be within [0,1], and its matrix dimension was reshaped to [224,224] to align with the input requirements of the pre-trained model. When evaluating the performance of the CNN network, two variables were tracked in this study: accuracy and loss. Accuracy is based on whether the model classification is equivalent to the actual class. Loss, on the other hand, is determined using cross-entropy loss, which is the difference between the true distribution and the predicted distribution. **Equation 1** gives the formula for cross-entropy loss where the true distribution is a one-hot vector—a vector containing all 0s except for one entry set to 1.

$$Loss = -\log (p) \tag{1}$$

Where  $p$  is the probability assigned to the actual class in the output probability vector.

During training, a stopping criterion is placed to prevent overfitting. If the loss of the model does not improve in 20 epochs (one pass through the entire dataset), training is concluded automatically. All CNN models were trained using the ADAM optimizer.

### ***Image Thresholding***

Image thresholding was chosen similarly to why RSC was chosen as a predictor: the greater presence of snow means lower friction values due to less contact area between tire and asphalt. Thresholding involves converting the pixel value of the image to either 1 (white) or 0 (black). Suppose we assume that the frequencies of 1s and 0s in a thresholded image represent the number of pixels classified as snow and asphalt, respectively. In that case, the output from thresholding

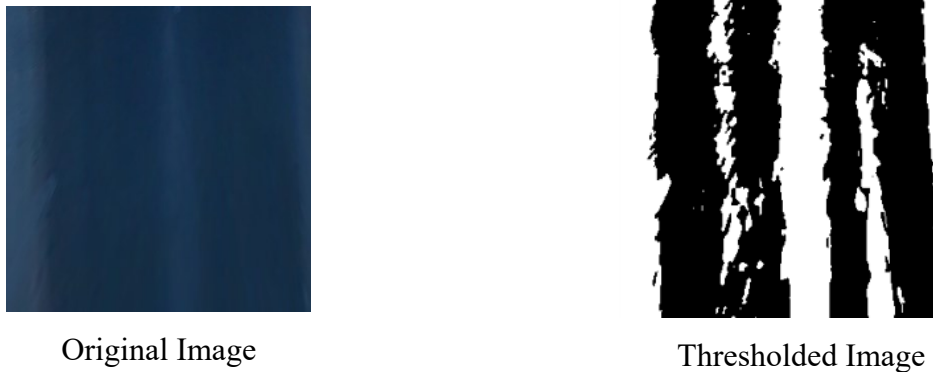
can be used to represent the degree of snow coverage, making it a potential input feature. In this study, the chosen thresholding method is mean adaptive thresholding, a regionalized thresholding technique where the threshold value varies depending on the image region. For each pixel value, the mean of its neighborhood is used as the threshold. If the pixel exceeds the local threshold, it is assigned 1. If not, it is given 0. This process is summarized in **Equation 2**.

$$Threshold = \frac{\sum_{i=1}^n \sum_{j=1}^n x_{ij}}{n^2} \quad (2)$$

$$x_{ij} = \begin{cases} 1, & x_{ij} > Threshold \\ 0, & x_{ij} \leq Threshold \end{cases}$$

Where  $x_{ij}$  is the pixel value at row  $i$  and column  $j$ , and  $n$  is the width of the square neighborhood region.

The benefit of adaptive thresholding is its ability to accommodate changing lighting conditions (Davies, 2005). Depending on the surrounding environment and time of day, certain portions of the road surface image may appear more illuminated than others. For these images, adaptive thresholding can account for image contrast differences. It can assign a lower threshold value to shaded areas and a higher value to illuminated regions, thereby allowing for a more accurate separation between snow and asphalt pixels (an example of this process is depicted in **Figure 3.6**).



**Figure 3.6 Original Image Versus Thresholded Image**

### ***Local Binary Pattern (LBP)***

Unlike RSC and image thresholding, which is based on prior knowledge of factors that influence friction, LBP is a texture analysis method that does not require any previous understanding of the data. LBP characterizes local spatial structures and contrast patterns. It is simple to implement and produces high-performance accuracy (Armi & Fekri-Ershad, 2019). The extraction process is similar to adaptive thresholding, where the value of a pixel is based on its surrounding neighbors. This study applied the uniform non-rotational variant of LBP to reduce dimensionality. Uniform means that the output of LBP is classified into either uniform or non-uniform features, and non-rotational implies that LBP is not affected by image rotation. Standard LBP outputs 256 features for a neighborhood of eight pixels, whereas the variant we use only outputs ten but preserves approximately 90% of the variation found (Ojala et al., 2002). Since we only have 128 samples, using all 256 features would result in poor model performance as we need more data to fully explore the relationship between the predictor and output variables, i.e., the curse of dimensionality (Chen, 2009).

The LBP process implemented herein can be summarized as follows: if the neighbor pixel is smaller than the center pixel, it is assigned 0; otherwise, it is set to 1. The eight neighboring values are then organized into a binary sequence. Within each sequence, if there are greater than two transitions from 0 to 1, it would be labeled as a non-uniform sequence; all other cases are uniform (Tchangou Toudjeu & Tapamo, 2019). Among the uniform sequences, the binary sequences are reorganized into their lowest value form, making them rotational invariant. Then, the center value of these sequences is replaced by the sum of the binary sequence. In contrast, the center value of non-uniform sequences is automatically assigned the value 9. Once the image has been processed, each unique LBP value's frequency is counted and considered an input feature. The basic formulation of LBP is depicted below in **Equation 3**.

$$LBP = \sum_0^{n-1} f(x_n - x_c) \quad (3)$$
$$f(x) = \begin{cases} 1 & (x_n - x_c) \geq 0 \\ 0 & (x_n - x_c) < 0 \end{cases}$$

Where  $x_n$  is the neighbor pixel,  $x_c$  is the center pixel, and  $n$  is the number of neighbors.

It is important to point out that features extracted from LBP do not have any inherent meaning; they merely represent image patterns that resemble edges, corners, and flat surfaces. Due to the generic nature of these features, it is incredibly challenging to interpret how the frequencies of these patterns influence friction. In this project, the purpose of applying LBP is to capture the macro-textures of the road surface as well as any road surface details missed by the human eye, like the presence of sand and ice.

### ***Gray Level Co-occurrence Matrix (GLCM)***

Like LBP, GLCM is also a texture analysis method that examines the spatial relationship between pixels. GLCM generates a matrix that tracks the number of unique combination pairs of pixel values in a specified direction. This generated matrix is symmetrical, meaning a GLCM matrix in the west direction is equivalent to one in the east direction. Hence, to track the pixel value changes in all directions of the road image, GLCM only needs to be calculated in four directions: 0, 45, 90, and 135°. Subsequently, texture measures can be calculated using the GLCM matrix generated in each direction that describes the image. This study calculated the following measures: contrast, dissimilarity, homogeneity, Angular Second Moment (ASM), energy, entropy, and correlation. Contrast, dissimilarity, and homogeneity are all contrast measures; they represent how similar or different neighboring pixels are to one another. In comparison, ASM, energy, and entropy are orderliness measures; they quantify the randomness within an image. Lastly, correlation is a descriptive statistics measure that describes the linear dependency between neighboring pixels. Ultimately, all features extracted by GLCM are non-task specific. They do not have direct meaning tied to any output variable, making them difficult to interpret. Nevertheless, like LBP, it is a way to capture road surface texture information and details that are not easily labeled without physically examining the road surface.

**Equations 4 to 11** detail how each texture measure was calculated (Hall-Beyer, 2017).

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}} \quad (4)$$

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad (5)$$

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} P_{i,j}|i-j| \quad (6)$$

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (7)$$

$$\text{ASM} = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (8)$$

$$\text{Energy} = \sqrt{\text{ASM}} \quad (9)$$

$$\text{Entropy} = \sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j}) \quad (10)$$

$$\text{Correlation} = \sum_{i,j=0}^{N-1} P_{i,j} \left[ \frac{(i-u_i)^2}{\sigma^2} \right]$$

$$u = \sum_{i,j=0}^{N-1} i(P_{i,j}) \quad (11)$$

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i-u_i)^2$$

Where  $V_{i,j}$  is the pixel pair frequency for reference pixel value  $i$  and neighbor pixel value  $j$ ,  $P_{i,j}$  is the probability of  $V_{i,j}$ ,  $u$  is the GLCM mean,  $\sigma^2$  is the GLCM variance, and  $N$  is 255, which is the max pixel value in 8-bit gray-scale images.



### *Tree-based Machine Learning Algorithms*

Using the methods discussed above, candidate predictor features can be extracted from the collected road surface images. But to formulate the relationship between the extracted features and friction values, an appropriate modeling technique must be used. In this study, tree-based machine learning methods were applied over traditional statistical methods and deep learning methods due to their non-parametric nature and high interpretability. The issue with deep learning is that it creates black-box models, and in the case of traditional statistical methods like generalized linear models, assumptions are made about the relationship between the dependent and independent variables, which may hinder performance (Gaur et al., 2021; Pichler et al., 2020). Tree-based methods address both of these issues; they offer visibility into their decision-making process while also being non-parametric. These non-parametric algorithms use decision trees to predict the desired output variable from a set of predictor features. During the construction of a decision tree, each predictor feature will be evaluated on its ability to split the data in a way that generates the lowest squared error. The feature that produces the lowest squared error will be chosen as the tree's root. The benefit of this selection process is that redundant variables are automatically filtered out since they cause an insignificant reduction in error. Following the creation of the root node, the same predictor selection process will be carried out for each subset of the divided dataset, where the best predictor becomes the direct child of the root node. This process repeats until the minimum sample size requirement for further splitting is not met or if the tree has grown to its maximum allowable size. These restrictions aim to prevent the tree from overfitting to the training dataset. To determine the two parameters that govern tree growth, a parameter search must be performed to find the configuration that results in the highest model performance.

Three types of tree-based machine learning algorithms were explored in this study: decision tree (Quinlan, 1986), Random Forest (RF) (Breiman, 2001), and Gradient Boosting (GB) (Friedman, 2001). Unlike the decision tree model that uses only one tree, RF and GB employ multiple trees in their prediction process. RF constructs its decision trees by bootstrapping the dataset. Because each tree is constructed on a slightly different dataset, each tree would offer its own unique explanation, and when the constructed trees are grouped into a forest, it is capable of accounting for more variation than a single decision tree. Furthermore, the bootstrapping process also makes it more difficult to overfit the data, as the dataset changes each time a tree is constructed. Similarly,

GB also employs multiple decision trees. However, instead of bootstrapping the dataset, GB constructs trees to account for the residual between its predictions and observed values. The goal of the GB algorithm is to slightly reduce the model's prediction error with each additional tree. Among the three different algorithms, the general consensus is that RF and GB outperform the decision tree model. However, the improved performance comes at the cost of interpretability. With the decision tree model, there is only one tree to interpret; the decision path taken for each prediction can be clearly understood. In contrast, GB and RF models are composed of tens of thousands of trees in some instances. For a user to understand their internal workings, they would have to examine each individual tree model. Although possible, it makes interpreting the decision-making process incredibly challenging as each prediction is an amalgamation of tens of thousands of decision paths—a tradeoff that one must balance between accuracy and interpretability when developing models. Furthermore, there may be conflicting decision paths in the RF model as the dataset changes due to bootstrapping, and for GB, the decision trees predict residuals, not actual friction values, all of which further complicate the interpretation process. One method to get around this is through permutation feature importance. A feature's importance is the decrease in performance a model experiences when the feature of interest is perturbed through random shuffling. If no change in performance is observed after perturbation, that feature would have zero importance (Molnar, 2022).

Prior to model development, the dataset was split into 80% training and 20% testing. K-fold (Hastie et al., 2009) was then applied to the 80% training dataset to divide it into five folds for cross-validation. These five folds were put through 500 random search trials to determine the optimal parameter values for the tree-based models. From our observation, model performance typically plateaus after 300 trials; nevertheless, an additional 200 trials were added to explore the parameter space more conservatively. The best configuration found—the one with the lowest RMSE—was used to construct the final model, followed by a final performance evaluation using the testing dataset via RMSE and Root Mean Squared Percentage Error (RMSPE); the formulas for these two error metrics are shown in **Equation 12 and 13**. This final performance evaluation is the basis for comparing the three tree-based algorithms.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N E_i^2}{N}} \quad (12)$$

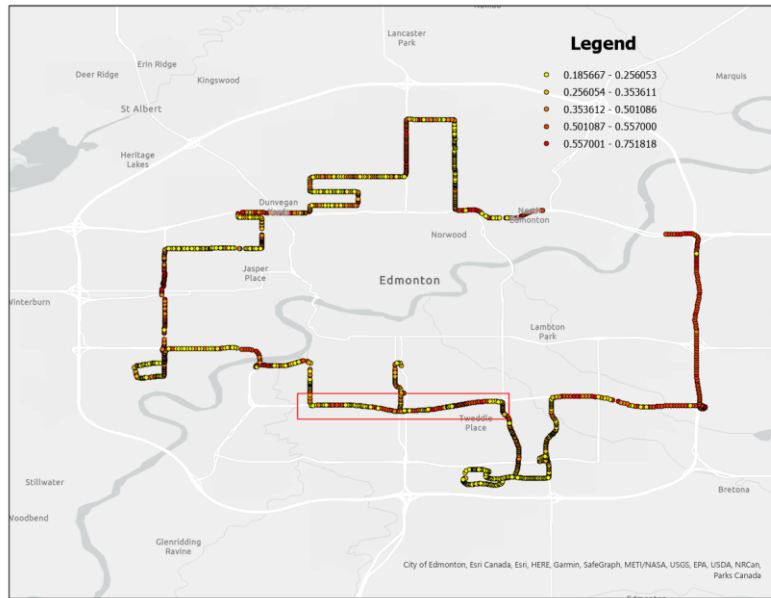
$$RMSPE = \sqrt{\frac{\sum_{i=1}^N \left(\frac{E_i}{Y_i}\right)^2}{N}} \quad (13)$$

Where  $E_i$  is the difference between predicted and observed,  $N$  is the sample size, and  $Y_i$  is the actual value.

### 3.2 Component Two: Spatial Interpolator

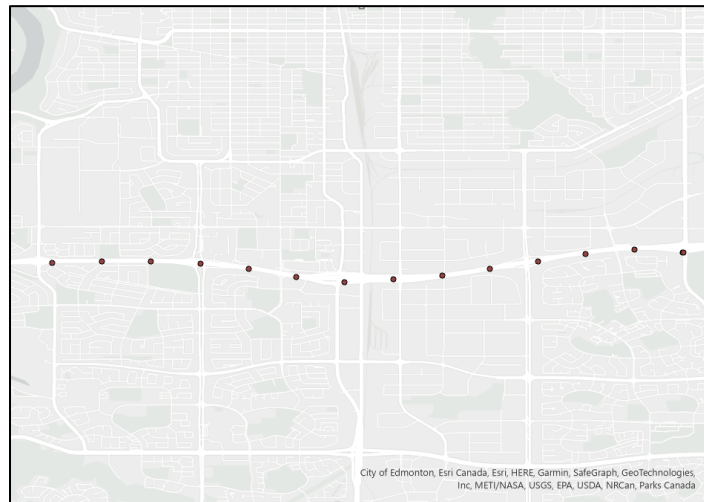
#### Interpolator Data

Evaluating interpolator performance requires sufficient data density to allow the chosen interpolator to learn the observed spatial variations. This condition prevented us from using the collected friction values because they were extremely sparse. To overcome this problem, road surface images were extracted from Jan 18, 2022, footage at a rate of one image every five seconds. These images were fed into the image-based friction model to generate spatially dense friction values. Within the friction dataset generated, the Whitemud Drive section was selected as the focus of this study due to the high friction variation present. After extracting data from this section, each friction measurement was averaged with neighboring values (50 m radius) to make the spatial patterns more distinct (*Smoothing Time Series*, 2023). The generated friction values and the road section selected for this study are shown in **Figure 3.7**.



**Figure 3.7 Generated Continuous Friction Values and Selected Segment for Interpolation**

The generated data were then separated into training and validation data. This process involved keeping only measurements spaced by a certain distance, i.e., a measure is kept every “x meters.” The measurements kept were used for training, whereas the measurements removed were reserved for validation. A total of 10 datasets were created in this manner with increasing distance—from 100 m to 1000 m per observation (step size of 100 m)—to evaluate the impact separation distance has on interpolation accuracy. An example of the training data is depicted in **Figure 3.8**.



**Figure 3.8 Training Data With 500-Meter Separation Distance**

The function of an interpolator is to estimate unmeasured values using known values. From the perspective of this study, the interpolator was used to estimate the removed friction values using the retained values. Three different types of interpolators were compared to perform this operation: Kriging, machine learning, and hybrid models that combine two algorithms. The two Kriging methods selected are OK and RK. The reason for choosing RK is that it has been shown to be the highest-performing interpolator for road weather variables (Gu, 2019; Wu & Kwon, 2022). In order for ML methods to be considered superior, they must exceed the performance of the state-of-the-art RK, i.e., RK serves as the comparison baseline. In terms of OK, it was selected in case of a lack of linear trend in the data. In this situation, OK replaces RK as the comparison baseline as it has shown higher performance (Zhu & Lin, 2010). With respect to the ML methods, the two interpolators selected were RF and RFSI, as these two models have shown superior performance than RK and OK in several environmental studies (Leirvik & Yuan, 2021; Li et al., 2011; Sekulić et al., 2020). We would like to know if their better performance extends into road-related variables, which is one of the objectives of component two. The last category of methods is hybrid methods, which are RSIOK and RFOK. The reason for combining ML with OK is the same as why

regression was combined with OK to form RK—to detrend the dataset. However, unlike RK, which assumes the trend to be linear, RF and RFSI may do a better job detrending the dataset as they are non-parametric, leading to overall better performance. The specifics of each of the three categories of interpolators are discussed below.

### **Ordinary Kriging (OK)**

Kriging has been shown to be the most accurate interpolation method for road weather variables (Gu, 2019; Wu & Kwon, 2022). The unique property of this method is that it considers the spatial covariance structure of the measurements and uses it to estimate unmeasured values. **Equation 14** below depicts the general Kriging formula.

$$\hat{Z}(x) = m(x) + \sum_{i=1}^m \lambda_i [Z(x_i) - m(x_i)] \quad (14)$$

Where  $\hat{Z}(x)$  is the interpolated value at an unknown  $x$  location,  $m$  is the number of observation points,  $m(x)$  is the expected value at unknown location  $x$ ,  $\lambda_i$  is the Kriging weight for observation  $i$ ,  $Z(x_i)$  is the observed value at location  $i$ , and  $m(x_i)$  is the expected value at location  $i$ .

There are several forms of Kriging, such as Simple Kriging (SK), Ordinary Kriging (OK), and Regression Kriging (RK). The main difference between these models is how the  $m(x)$  or trend component is modeled. The form we will focus on here is OK because, unlike SK, OK does not assume that the global mean is known, which leads to more accurate results. Furthermore, OK is required in RK and hybrid models to model residuals after detrending. The formula for OK (Wackernagel, 1995) is derived from **Equation 14** above by forcing the weights to sum to one, i.e.,  $\sum_{i=1}^m \lambda_i = 1$ . This change allows us to rewrite the Kriging formula as follows:

$$\hat{Z}(x) = \sum_{i=1}^m \lambda_i(x) Z(x_i) \quad (15)$$

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_2 \\ u \end{bmatrix} = \begin{bmatrix} C(x_1, x_1) & \cdots & C(x_1, x_m) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C(x_m, x_1) & \cdots & C(x_m, x_m) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} C(x_1, x) \\ \vdots \\ C(x_m, x) \\ 1 \end{bmatrix} \quad (16)$$

$$\sum_{i=1}^m \lambda_i(x) = 1 \quad (17)$$

Where  $C(x_n, x_m)$  is the covariance between the observations at location  $n$  and  $m$ .

In order to determine the covariance structure, a semivariogram (Hohn, 1999) must first be constructed to model the degree of dissimilarity between two measurements based on their separation distance. The dataset used for this purpose must be trend free to meet the modeling assumption of mean stationarity (Olea, 2006). Upon removing the trend, an empirical semivariogram can be assembled using the following formula:

$$\hat{\gamma}(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x_i + h) - Z(x_i)]^2 \quad (18)$$

Where  $\hat{\gamma}(h)$  is the semivariance at lag distance  $h$ ,  $n$  is the number of samples, and  $Z(x_i + h)$  is the observed value at  $h$  lag distance away from observation at  $x_i$ .

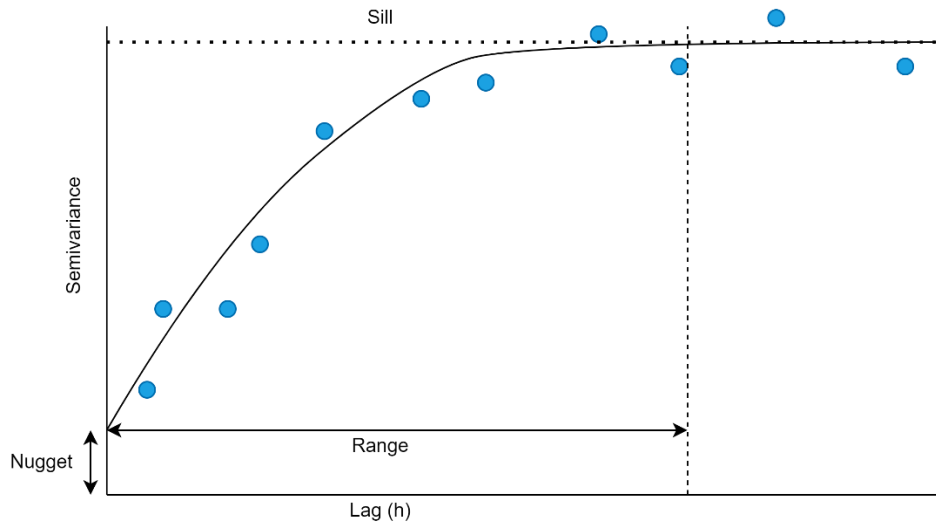
Based on the developed empirical semivariogram structure, a theoretical variogram with a similar structure must be fitted, from which the covariance values used in Kriging are obtained. There are several options for theoretical semivariogram models; the most common ones are shown in **Table 3.2**. The spherical model was selected for this study as it had the best fit.

**Table 3.2 Common Theoretical Semivariograms**

Theoretical Model	Formula:
Spherical	$C \left( \frac{3h}{3a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right)$
Exponential	$C \left( 1 - e^{-\frac{3h}{a}} \right)$
Gaussian	$C \left( 1 - e^{-3 \left( \frac{h}{a} \right)^2} \right)$

Where  $C$  is the sill,  $h$  is the distance, and  $a$  is the range.

Regardless of the model chosen, all semivariogram models contain three parameters: nugget, sill, and range. The nugget is the amount of semivariance at a lag distance of 0, typically due to measurement error. In comparison, the sill is the point at which semivariance plateaus, beyond which observations are no longer considered spatially correlated. Lastly, the range is the lag distance where the sill is reached. In other words, it is the maximum distance at which spatial autocorrelation is present. A stereotypical semivariogram is illustrated in **Figure 3.9**.



**Figure 3.9 Typical Semivariogram Structure**

**Random Forest (RF) and Random Forest Spatial Interpolator (RFSI)**

Among the many Machine Learning (ML) methods, RF has shown the most promising results as an interpolator (Leirvik & Yuan, 2021; Li et al., 2011; Sekulić et al., 2020). When used in this manner, it functions in the same way as traditional RF. The dataset is randomly sampled with replacement to create a new dataset, which is then used to create a decision tree model. This process of sampling the dataset and constructing a decision tree model repeats until the maximum number of trees has been reached, at which point, training concludes. The main difference is that position data (x and y coordinates) must be included to allow the model to capture spatial covariance.

However, using coordinate information as the sole spatial data may not be enough to capture the target variable’s spatial structure. Hence, the Random Forest Spatial Interpolator (RFSI) model was conceptualized to compensate for this deficiency. RFSI attempts to mimic how spatial structure is modeled in traditional statistical interpolators, where it is assumed that variables

decrease in similarity as the distance between them increases. By adding information regarding observation values at nearby observations and the distance to these neighbors, the RF algorithm can better capture the spatial structure of the target variable, leading to better performance (Sekulić et al., 2020). **Equation 19** shows the basic formulation of the RFSI model.

$$\hat{Z}(x) = f(Z(x_1), \dots, Z(x_m), d(x, x_1), \dots, d(x, x_m), a_1, \dots, a_n) \quad (19)$$

Where  $\hat{Z}(x)$  is the RFSI interpolator,  $Z(x_m)$  is the observed value at  $m^{\text{st}}$  nearest neighbor,  $d(x, x_m)$  is the distance between estimation location  $x$  and observed value at  $m^{\text{st}}$  nearest neighbor, and  $a_n$  is the  $n^{\text{th}}$  auxiliary variable.

This study considers only the five closest neighbors because when the observation distance is set to 1 km, there are not enough measurements to use a larger neighborhood size.

### Hybrid Models

Besides Kriging and ML, there is a third type of interpolator that combines OK with another algorithm. Compared to pure models, hybrid models have shown higher accuracy in existing research (Li et al., 2011). The principle behind this is that instead of assuming the mean to be constant within a local region, the mean is a function that depends on position coordinates and other auxiliary variables. To predict this mean, a separate model is developed using any algorithm capable of regression. This model would be used to predict the mean or trend in the observations. The mean would then be subtracted from the observation to determine the residual. This process of removing the mean from the observations is called detrending. With the detrended data, OK is used to model the spatial relationship between the residuals. At the end of this process, there are two models: a regression model that predicts the mean and an OK model that predicts the residual. By summing the two predictions, the final output is obtained via **Equation 20**.

$$\text{Regression Model} + \text{OK Model} = \text{Target Variable Prediction} \quad (20)$$

Three hybrid models are evaluated in this thesis: Regression Kriging (RK), which combines Linear Regression with OK, RFOK, and RFSIOK. Note that RK can be considered a Kriging model because the trend component can be formulated directly into the Kriging algorithm. In the context



of this thesis, RK is grouped with RFOK and RFSIOK because all three models are developed in the same manner.

### **3.3 Component Three: Binary Collision Likelihood Model**

#### **Collision Likelihood Dataset**

Similar to the dataset used for interpolator evaluation, the dataset used to develop the collision likelihood model was also based on the aforementioned friction model. However, in this case, all the road footage available (Jan 18, 19, 25, and Feb 04, 2022) was converted into friction coefficients to maximize the training data available. Following this conversion process, the road sections with friction values were divided into equal-length segments, where the friction value of a segment is the average of all overlapping friction values. In addition to friction, other road segment characteristics were also assigned, including elevation, AADT, slope, and x and y coordinates. Next, a binary classification was given to each segment identifying whether a collision had occurred based on traffic safety records provided by the City of Edmonton. Note that this process was done independently for each day.

#### **Decision Tree Classifier**

The decision tree classifier was chosen to model the relationship between collision occurrence and the selected independent variables. As mentioned earlier, decision trees are highly interpretable compared to other ML models. Post-training, the decision tree model offers unique transparency by allowing users to follow its internal logic and understand how predictions are made. This level of insight is in stark contrast to more complex algorithms such as RF, GB, and deep learning models. Although these advanced models possess strong predictive power, they often function as “black boxes” that often fail to provide any rationale for their predictions. This lack of transparency can impede targeted safety improvements as the underlying causes remain unclear. In contrast, while traditional parametric SPF functions are transparent in their decision-making, they assume that crashes follow a particular distribution—a strong assumption that undermines their performance. Past studies, including Das et al. (2021), have shown that prediction error for SPF is contingent on the chosen distribution and can be outperformed by regression tree models that do not make such assumptions. In addition, as a non-parametric model, decision tree algorithms can

manage high-dimensional datasets, which align with the multidimensional nature of collision data due to the multitude of contributing variables.

Hence, our choice of the decision tree algorithm hinges on its dual benefits: one, its proven performance advantage as a regression tree model compared to traditional SPF, and two, its transparency that empowers users to discern why a road section is deemed risky, enabling them to implement precise mitigation measures. Furthermore, from a friction research perspective, it allows us to see the underlying relationship between friction and collisions, identify safety thresholds, and check the model’s overall intuitiveness.

As for the model algorithm, the decision tree classifier works in the same way as a normal decision tree, except that the Gini Impurity Index (**Equation 21**) is used as the error metric. Of the available input variables, the one with the lowest Gini impurity is selected as the decision node. The stopping condition used to develop the collision model is based on the number of decision nodes. Training and validation accuracies are evaluated simultaneously as the number of tree nodes increases. The optimal number of nodes is the point where the validation accuracy peaks. In cases where multiple configurations produce the same accuracy, the structure with the least number of nodes was selected.

$$\text{Gini Impurity} = 1 - \sum_{i=1}^k p_i^2 \quad (21)$$

Where  $k$  is the number of classes, and  $p_i$  is the proportion of *the*  $i$  label.

A dataset split of 80% training and 20% validation was used during model development.

### 3.4 Summary

This chapter provides an overview of the proposed framework and explains the methodology behind each of the three components: image-based friction model, friction interpolator, and binary collision likelihood model.

The image-based friction model is developed based on traction testing data collected by the City of Edmonton and their associated road surface image. Because the images are not usable as inputs on their own, a series of feature extraction techniques are used to convert them into predictor

features. These methods include RSC categorization, image thresholding, LBP, and GLCM. Then, tree-based algorithms are employed to model the relationship between friction values and the extracted image features.

Component two focuses on the construction of the friction interpolator. Because collected traction data is too sparse, the dataset used to develop the interpolator is generated using the image-based friction model. Three types of interpolators are selected based on high performance in the existing literature: Kriging, RF-based, and hybrid models. Each interpolator is evaluated based on its ability to interpolate friction values as the separation distance between available input data increases (from 100 to 1000 m). The goal is to find the interpolator with the least sensitivity to change in separation distance while maintaining high accuracy.

The final component centers around the development of a binary collision model. Continuous friction values are inputted into the collision model, along with AADT and supporting variables, to produce a collision likelihood value that indicates whether a collision is likely or unlikely for a given road segment. The data used for this model is also generated from the image-based friction model for the same reason as component two, data sparsity. The algorithm used for this purpose is the decision tree model. Its high interpretability allows us to validate the friction-collision relationship, determine unsafe thresholds, and judge the intuitiveness of the model.

# Chapter 4 Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images<sup>2</sup>

As mentioned earlier, there are three components to the proposed framework. This chapter focuses on the first component: the image-based friction model, which converts images into point friction values. An accurate friction model is critical to the proposed framework as the values it generates are needed to develop the two remaining components. In this chapter, we first assess the accuracy of models built with tree-based algorithms and the importance of each predictor feature. Next, we demonstrate how the friction model can be automated using the CNN deep learning algorithm. Finally, we will summarize the key results at the end of this chapter.

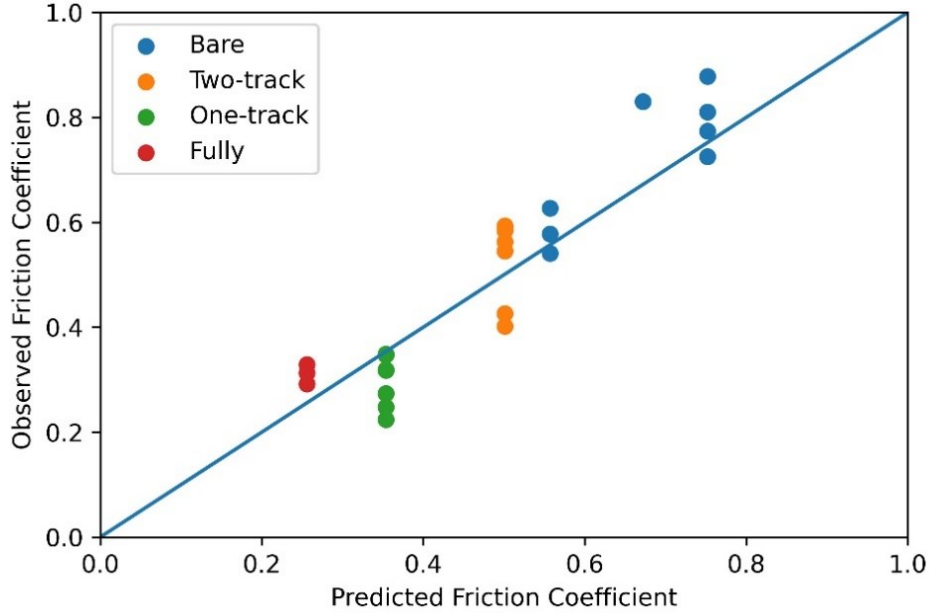
## 4.1 Friction Model Development

Following the techniques described in the methodology section, a total of 41 features were extracted: RSC category, frequencies of black and white pixels from image thresholding, 10 LBP features, and 28 GLCM features (7 in each of the four directions mentioned). With all 41 features as input, a parameter search was performed to determine the optimal configuration for the decision tree model. The results show a minimum split sample size of 8 and a max tree depth of 3 yielded the lowest RMSE.

When evaluated using the testing dataset, the decision tree model displayed a performance of 0.0759 RMSE and 19.6% RMSPE. In other words, 80.4% accuracy. This level of performance is reflected in the model's predictions vs. observation graph (**Figure 4.1**), where a linear relationship is observed between predictions and observations.

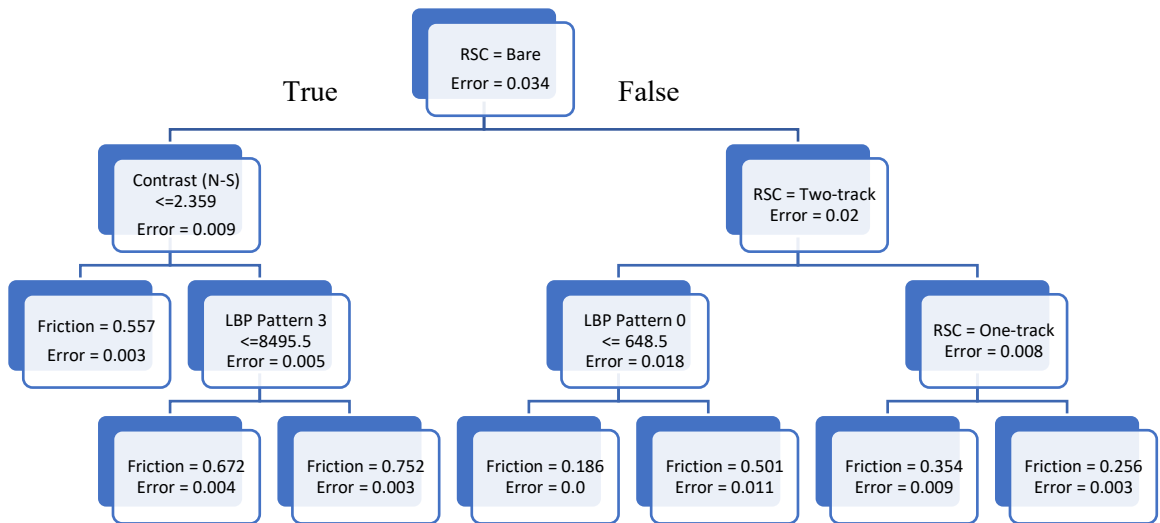
---

<sup>2</sup> Contents in Chapter 4 has been published in Xie, Q., & Kwon, T. J. (2023). Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images – A City of Edmonton, Canada, Case Study. Canadian Journal of Civil Engineering. Just-IN. <https://doi.org/10.1139/cjce-2023-0015>



**Figure 4.1 Decision Tree Model Prediction vs. Observed Friction Values**

Based on the decision tree model's low RMSE value, it is evident that the developed model is accurate. However, whether the predictions are based on intuitive interpretations of the extracted features is unknown. Fortunately, because the decision tree model is highly interpretable, the internal workings of the model can be visualized. By investigating the decision paths illustrated in **Figure 4.2**, we can further evaluate the model's reliability.

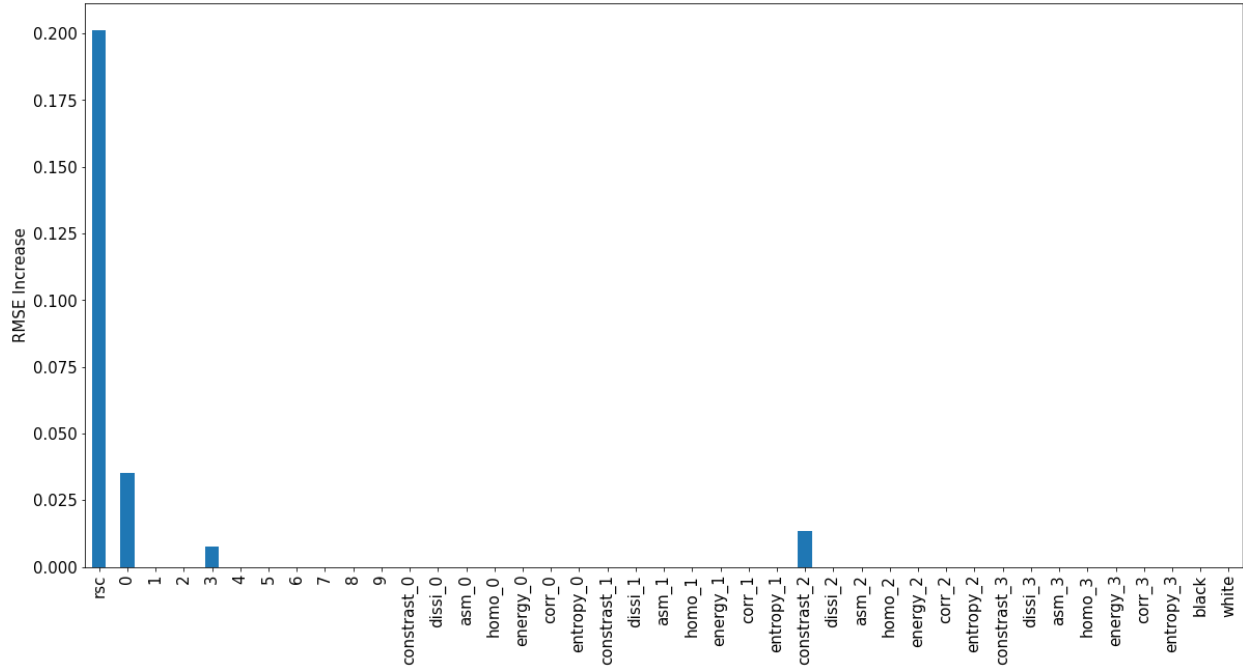


**Figure 4.2 Internal Logic of Decision Tree Model**

From the depicted tree model shown above, the user can see the logic behind the friction predictions and the relative importance of the predictor features. By examining the tree diagram, we can see that friction can be explained by four predictors: RSC category, GLCM contrast in the N-S direction, and LBP patterns 3 and 0. Among these features, their relative importance can be gauged by the change in squared error caused by the node's decision. In our developed model, RSC is the most critical feature because of its error-reduction capability and the number of times it appears. The decision tree we constructed has a depth of three, and at all three depths, RSC is used as the decision rule that reduces error by a significant amount: from 0.034 to 0.02, 0.02 to 0.008, and 0.008 to 0.003 at the top, middle, and bottom of the tree, respectively. This level of error reduction was not observed from any other input features, making the RSC the most vital predictor feature. Furthermore, how RSC splits the dataset is also intuitive. RSC value of 0 (bare lane) is assigned a high friction estimate of 0.679 because of the high level of exposed asphalt. Then, as the amount of road surface precipitation increases—equivalent to an increase in RSC value—the estimated friction value decreases due to more and more contaminants on the road surface.

In terms of the remaining three features, their relative importance in estimating friction is unclear as they appear the same number of times as decision nodes and cause similar levels of error reduction. They are also difficult to interpret as they are generic features that describe pixel patterns. Based on the decision rules, having low contrast in the N-S direction and less LBP pattern 0 and 3 decreases friction value, which could indicate less exposed asphalt or the presence of ice. Unfortunately, due to few studies in this subject area and the fact that these features are non-task-specific, we cannot confirm whether this is a common observation nor if it is intuitive, only that such a relationship may exist.

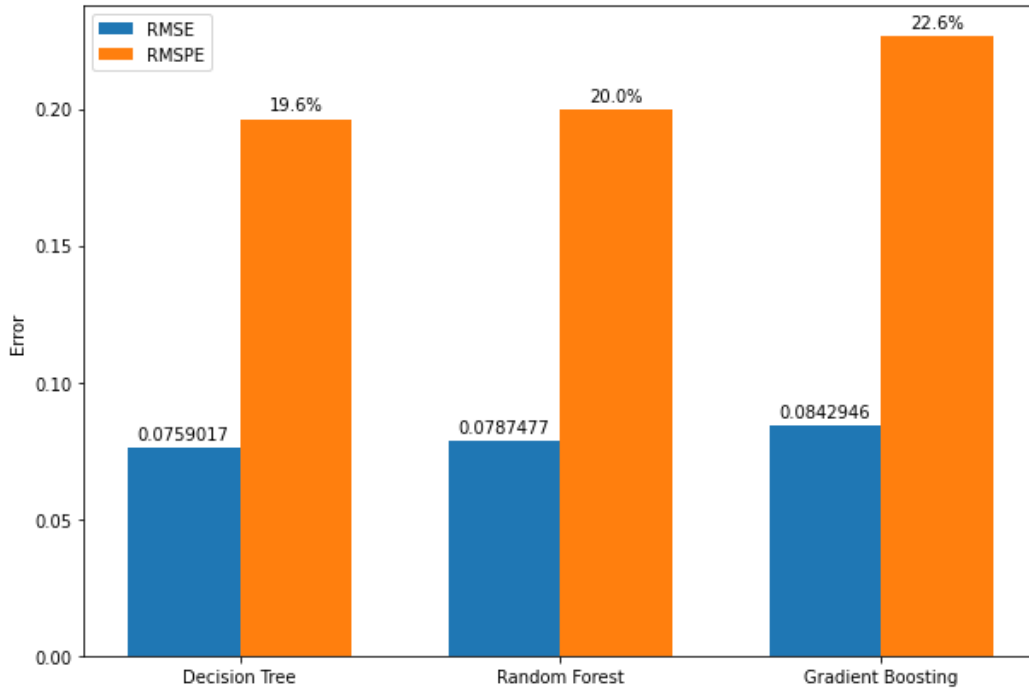
Other than focusing on the changes in squared error to evaluate feature importance, as stated in the methodology section, a more systematic way of assessing feature importance is by performing permutation feature importance, as shown in **Figure 4.3**.



**Figure 4.3 Feature Importance of Decision Tree Model**

The advantage of permutation feature importance is that it can quantify the difference between features that induce similar reductions in squared error. By examining **Figure 4.3**, we can see that RSC is undoubtedly the most important feature in our model. When we perturbed the RSC feature, it led to a 0.2 increase in RMSE, seven times more than the next most important feature. Excluding RSC, Pattern 0 is about twice as important as contrast and four times more important than pattern 3.

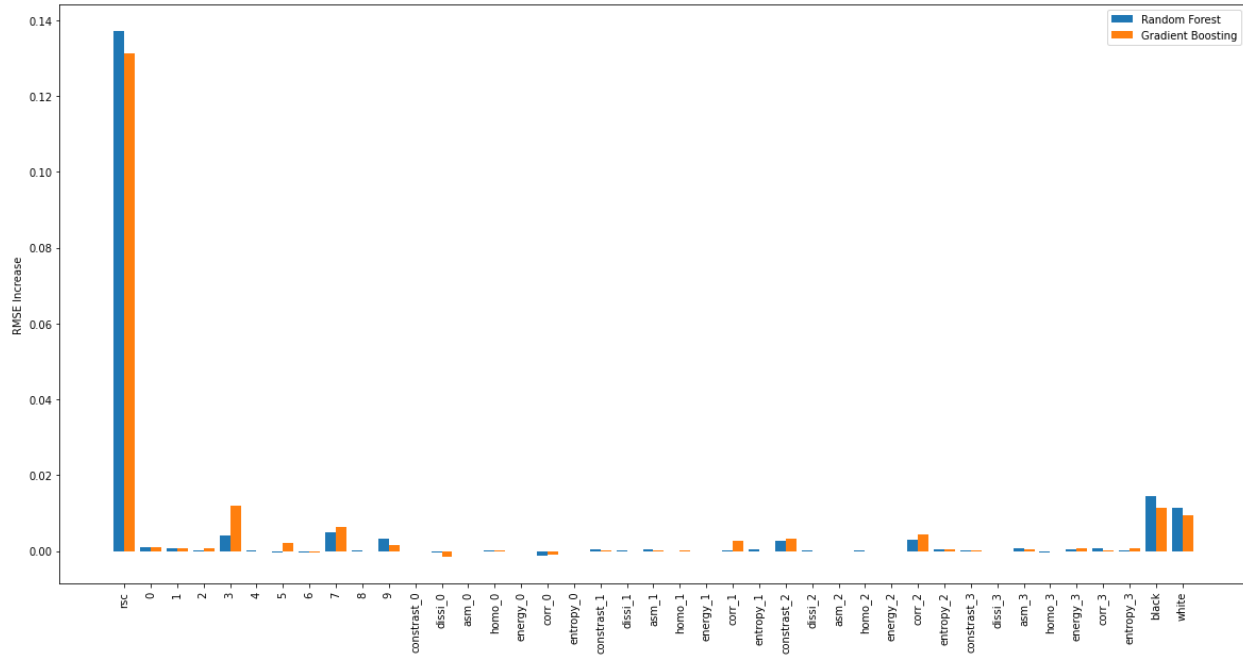
As mentioned, two other tree-based algorithms; namely, RF and GB were also investigated in this study as both aim to improve model performance by including multiple decision trees in the prediction process. From the results of our parameter search, we found the optimal configuration of our RF model to contain 75 trees with a minimum sample split of 7 and a max depth of 4. In the case of GB, 250 trees, minimum sample split of 2, max depth of 3, and a learning rate of 0.01 (dictates the contribution of each tree's prediction toward the outputted prediction). When these configurations are evaluated on the testing dataset, the results shown in **Figure 4.4** were obtained.



**Figure 4.4 Friction Model Performance Comparison**

As shown in **Figure 4.4**, both RF and GB obtained low errors. When considered in unison with the decision tree model, all three models displayed high performance on the testing dataset with an average RMSE and RMSPE of 0.0796 and 20.7%, respectively. This means that our developed models operate at approximately 80% accuracy, proving the proposed framework's feasibility. Although the decision tree model appears to outperform the other models, the difference in performance is relatively small: between 0.4 and 2.6%. One plausible cause could be the lack of data. Our dataset did not contain enough variation that required more complex algorithms. Another possibility is that the parameter search performed in this study did not find the most optimal configuration. Through additional trials and larger search space, we may be able to identify a higher-performing configuration. However, it must be pointed out that even if a slightly better model could be obtained, RF and GB are less interpretable due to a more complex estimation process. Unlike the decision tree model, where the user can clearly understand the decision-making process, RF and GB's feature importance can only be evaluated through permutation feature importance, which only shows the impact a feature has on overall model accuracy, as depicted in **Figure 4.5**.





**Figure 4.5 Random Forest and Gradient Boosting Feature Importance**

The feature importance from our developed RF and GB models show nearly identical results. Both figures depict RSC as the most important predictor feature, followed by the frequency of black and white pixels extracted from image thresholding. The importance of the remaining features varies between the two models (e.g., LBP pattern 3 is the second most important feature in GB but fourth in the RF model). Nevertheless, regardless of their order of importance, LBP patterns 3 and 7 and GLCM contrast and correlation in the N-S direction are relatively more critical than the remaining feature. Intuitively, the inclusion of black and white pixel frequencies makes sense as they represent the amount of bare asphalt and precipitation found in an image. The issue is that it is difficult to determine how they are used by the two models, as there are 75 trees in the RF model and 250 in the GB model. The same applies to the remaining features identified to be important; we only know that they are necessary for estimating friction but do not know why, owing to the low interpretability nature of RF and GB models.

From our analysis of feature importance in the tree-based models, it is evident that RSC is undoubtedly the most important feature in predicting friction. In all three developed models, RSC is seven times more important than the next feature and causes a significant increase in prediction error if removed—a minimum of 0.12 increase in RMSE. This observation validates the proposed connection between RSC and road friction suggested by previous researchers (Roychowdhury et

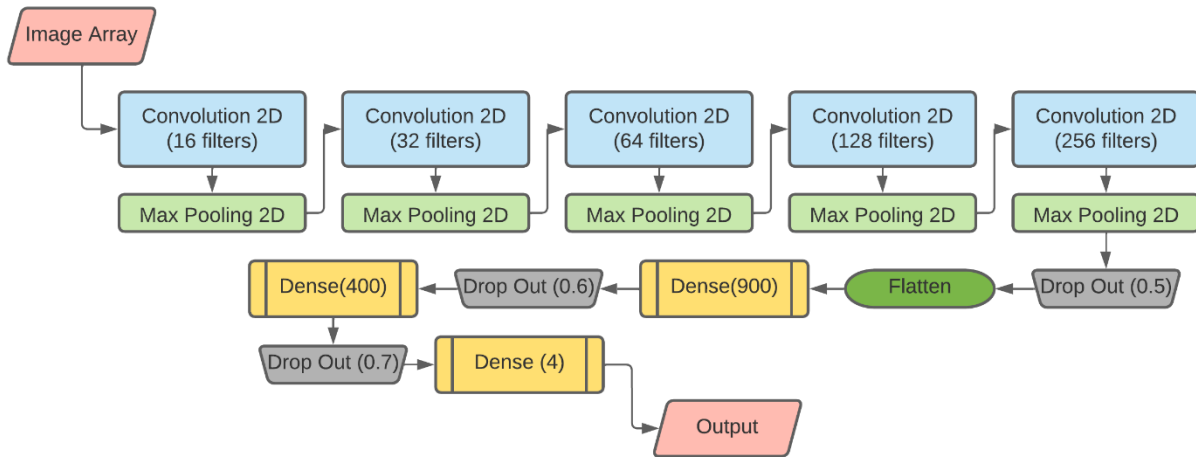
al. 2018; Wu and Kwon 2022). Other than RSC, LBP pattern 3 and GLCM contrast in the N-S direction were found to be important in all three models. Although they are not as impactful as RSC, it is speculated that they assist in finetuning the friction predictions. In contrast, the remaining features' importance depends on the modeling technique used. For example, frequencies of black and white pixels were insignificant in the decision tree model but important in the RF and GB models.

### **The Effect of Automating RSC Labeling**

Up to this point, the friction model performance shown is based on manual RSC labeling. Although it is possible to automate the RSC labeling process, it was not done during model development to prevent potential misclassifications from compromising model performance and feature interpretation. However, when it comes to model implementation, the RSC labeling process must be automated, or the need for manual labeling will severely limit update frequency and spatial coverage. Therefore, to maximize the cost-effectiveness and usability of the proposed work, the RSC labeling process was automated using a Convolutional Neural Network (CNN), thereby fully automating the image-based friction model.

In order to develop the CNN model, 5300 images were extracted from the GoPro footage recorded during traction testing. The number of images labeled in each category were as follows: 2281 bare, 1884 two-track, 582 one-track, and 553 fully snow-covered. Since the number of images in each category was unequal, the splitting of training and validation data was done separately. This ensured that 80% of the images in each category were in the training dataset. We also created two versions of this data, cropped and non-cropped. The benefit of cropping is that it eliminates environmental variables such as trees and nearby vehicles that may confuse the model, which may improve model performance.

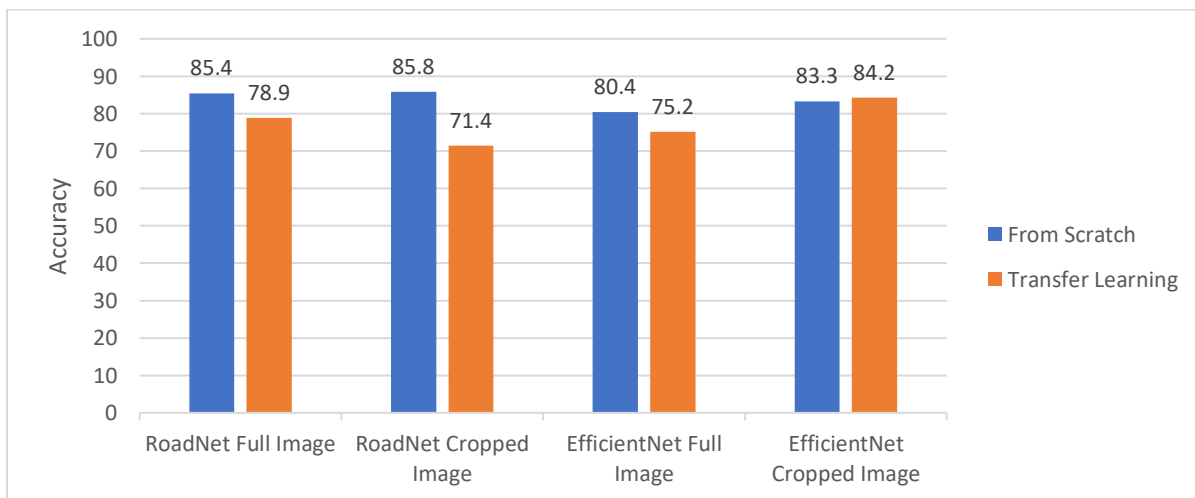
In terms of CNN architecture, the chosen architecture was initially designed by Carrillo et al. (2019) for RSC labeling. We and other researchers have optimized this specific architecture to develop models that label RSC into three categories (bare, partially snow-covered, and fully snow-covered) with >90% accuracy (Wu & Kwon, 2022; Xie & Kwon, 2022). **Figure 4.6** depicts the optimized CNN architecture found in our previous study (Xie & Kwon, 2022).



**Figure 4.6 Selected CNN Architecture**

In addition to the architecture depicted in **Figure 4.6**, a state-of-the-art general-purpose model—EfficientNet—one of the most accurate CNN models for image classification was included to show how our proposed model compares to the state-of-the-art. To help distinguish between the two designs, the architecture shown above will be referred to as "RoadNet" from this point forward.

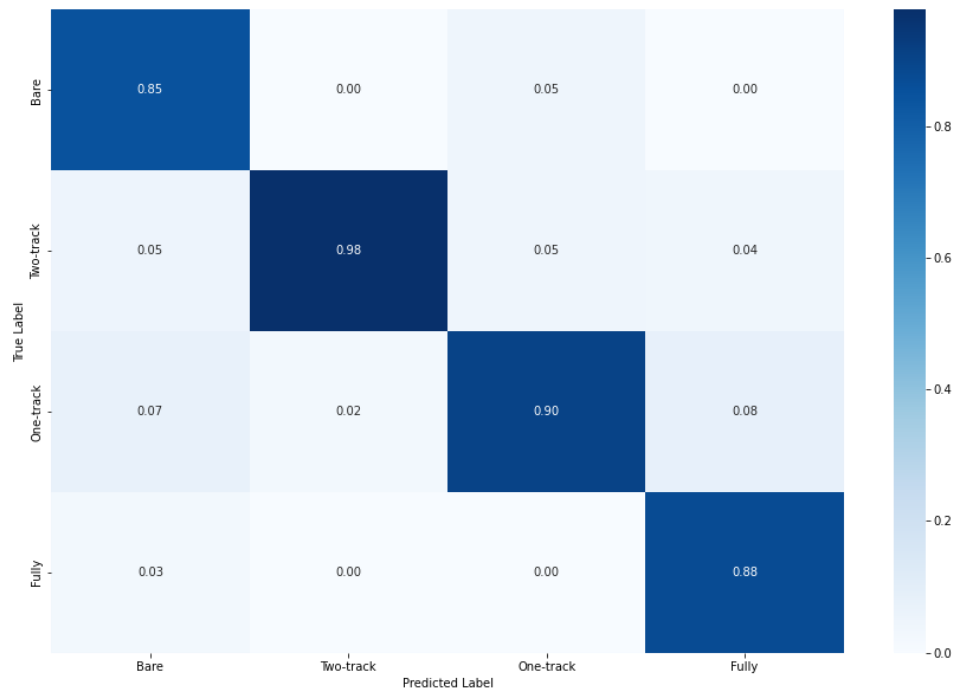
Based on these two architectures, two types of models were developed: models trained from scratch and models built using transfer learning. The required pre-trained models for transfer learning were taken from previous studies. These models include a three-category RSC model trained on 3,914 road images and an EfficientNet Model trained for everyday object identification. The accuracy results obtained are shown in **Figure 4.7**.



**Figure 4.7 Validation Accuracy Difference Between the Developed CNN Models**

From **Figure 4.7**, it is observed that training from scratch resulted in higher accuracy than transfer learning, which can be explained by dataset size. In a previous study, we performed a sensitivity analysis that compared the performance accuracy as the dataset size increased, from which we found that transfer learning outperformed training from scratch only when the dataset size was small (Xie & Kwon, 2022). As the dataset size increased, there was a point at which training from scratch became better than transfer learning. This is the case here, where the dataset size has exceeded the threshold. Another observation is that cropped images tended to have higher accuracy than full images, which was expected since removing the environmental variables allows the model to focus only on the road segment, simplifying the classification task. Overall, the most accurate model found was the RoadNet cropped image model trained from scratch, with an accuracy of 85.8%, 2 % higher than the state-of-the-art EfficientNet Model.

With the developed CNN RSC classifier, we evaluated the feasibility of automating the RSC labeling process by first comparing predictions made by the CNN model versus the actual labels. The result from this analysis is shown in **Figure 4.8**.

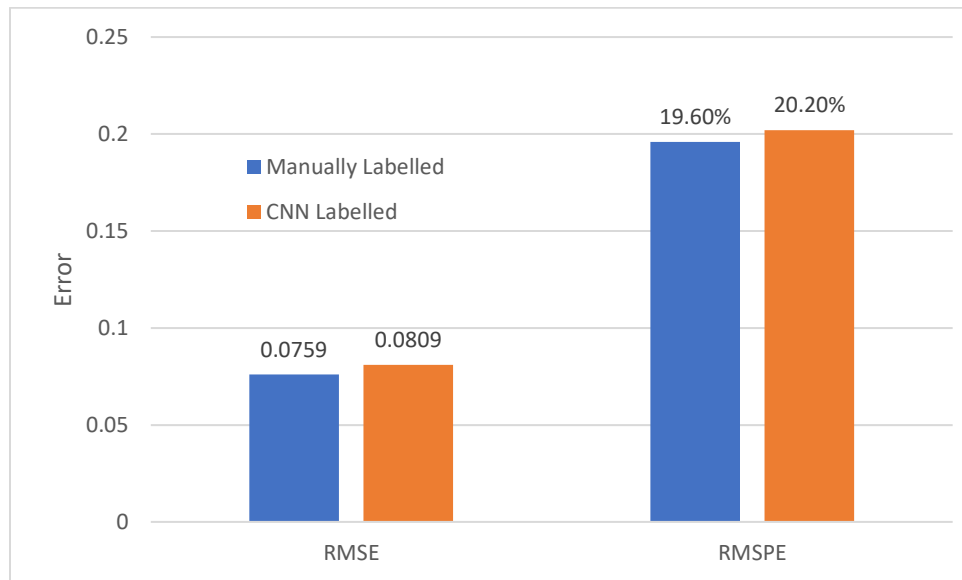


**Figure 4.8 CNN Model Accuracy on Friction Dataset Images**

According to **Figure 4.8**, it is evident that the developed CNN model is capable of assigning RSC labels accurately. All classes were labeled with an above 85% accuracy rate, signifying the high

performance of the developed model. Furthermore, the RSC categories commonly associated with low friction values had an even higher 90% accuracy, which is beneficial from a safety perspective, as it would mean the vast majority of potentially risky road conditions can be identified.

Next, the RSC labels in the validation dataset (friction model) were replaced with the CNN-predicted labels. The modified dataset was then used to re-generate the friction predictions using the friction model with the lowest error—the decision tree model shown in **Figure 4.2**. A performance comparison was then made between the prediction generated using CNN labels and actual labels with RMSE and RMSPE as metrics. The results generated from this comparison are shown in **Figure 4.9**.



**Figure 4.9 Prediction Accuracy Before and After Inclusion of CNN Model**

As shown in **Figure 4.9**, automating the RSC labeling process had minimal effect on friction prediction accuracy. RMSE increased by 0.05, and RMSPE increased by 0.6%. Based on these results, it is evident that the image-based friction model can be fully automated without impacting performance accuracy.

## 4.2 Summary

This chapter presents the results generated from the development of the image-based friction interpolator. There are two sections in this chapter, the first section focuses on the friction model

itself, and the second section explains how the model can be fully automated and the implications of doing so.

The developed model converts road surface images into 41 features through four feature extraction techniques, RSC classification, adaptive thresholding, Local Binary Pattern (LBP), and Gray Level Co-occurrence Matrix (GLCM). Using the extracted image features as input and friction values as output, three friction prediction models were developed via three tree-based algorithms: decision tree, random forest, and gradient boosting. All three models displayed high performance with an average RMSE of 0.0796 and an accuracy of 79.3% based on RMSPE. In addition, it was discovered through a method called permutation feature importance that RSC is the most critical predictor for friction. RSC was found to be seven times more important than the next most important measure and would cause at least a 0.12 RMSE increase had it not been included as a predictor. Other features like LBP pattern 3 and GLCM contrast in the N-S direction were also found to be necessary, but because they are generic image features without inherent meaning, we could only determine that they assist in estimating friction value without knowing the exact mechanism. Nevertheless, we speculate that these features may be proportional to asphalt exposure.

With an image-based friction model developed, the next step was to automate the RSC labeling process to fully automate the friction prediction process. For this purpose, eight CNN models were developed that differed in model architecture (RoadNet and EfficientNet), training method (from scratch and transfer learning), and input image transformation (cropped and full). By evaluating the performance of these eight models, it was observed that training from scratch outperformed transfer learning, using cropped images produced better results than full images, and our RoadNet model outperformed the state-of-the-art EfficientNet model. In an effort to determine the effect of adding the CNN model to the prediction process, friction predictions were made using the CNN-generated RSC labels and compared to the actual friction values to calculate RMSE and RMSPE. Compared to the manually labeled friction values, there was only a 0.05 and 0.6% increase in RMSE and RMSPE, respectively, demonstrating that the developed image-based friction model can be fully automated.

# Chapter 5 Assessing Interpolation Methods and Collision Models for Road Safety Enhancements

The focus of Chapter 5 is on components two and three of the proposed framework. Component two is the friction interpolator, which converts the point friction values generated by the image-based friction model into continuous values without any additional data collection effort, and component three is the binary collision model, which identifies the dangerous road sections. There are three sections in this chapter: the first section discusses the interpolation of the accuracy of the six chosen interpolators; the second section evaluates the possibility of using existing infrastructure (traffic cameras) to generate continuous friction values; the third section focuses on the development of the binary collision model via aggregation distance calibration; and the last section summarizes the main findings from the previous three sections.

## 5.1 Evaluation and Development of Spatial Interpolators

The interpolator evaluation process has two components: first, the RMSE of each interpolator was calculated to quantify the difference between the measured and the predicted, and then the ability of each interpolator to capture the spatial pattern was examined to verify the credibility of the obtained RMSE. **Figure 5.1** depicts the results generated from the first part of this analysis.

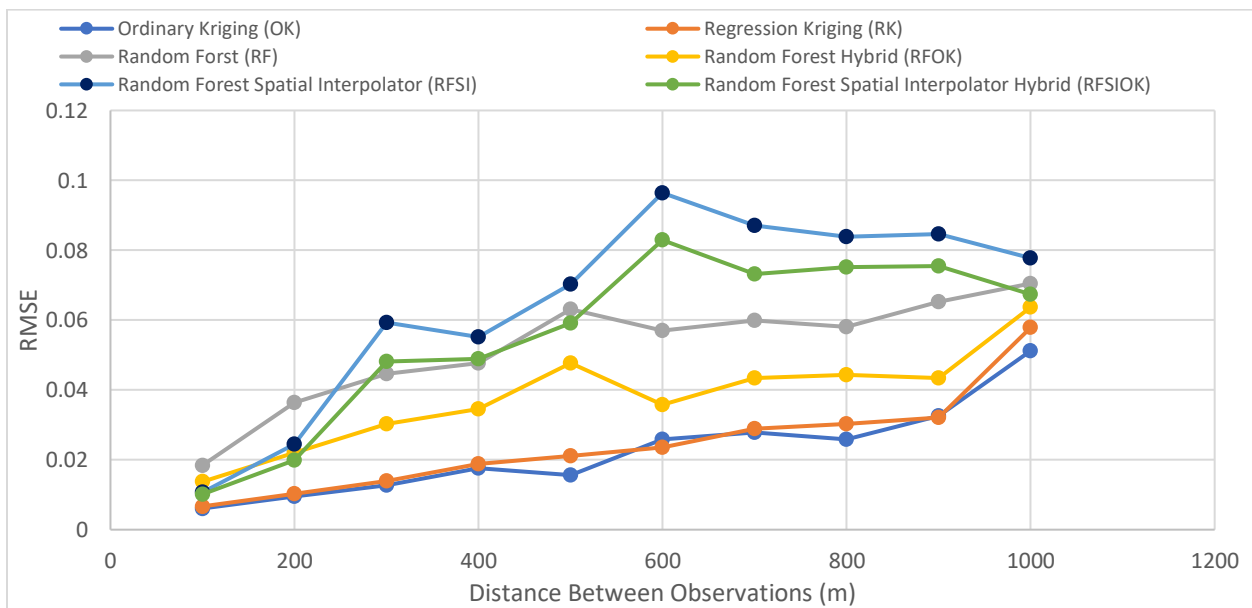


Figure 5.1 Interpolator Performance Comparison

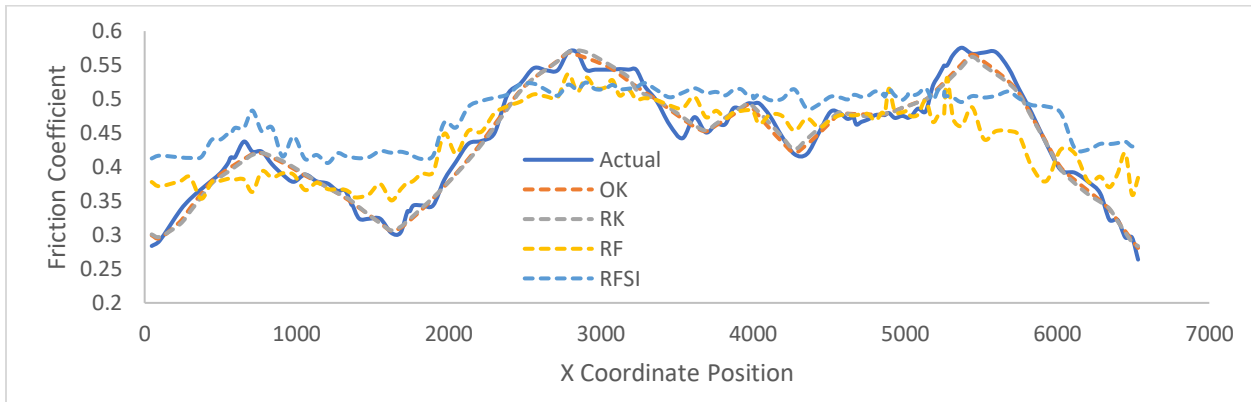
According to **Figure 5.1**, the interpolator accuracy decreases as the separation distance increases, which is expected due to the decrease in the information available to the model to capture the true spatial variation. However, the amount by which interpolator accuracy decreased with increasing distance varied. OK and RK had the slowest rate of performance degradation, with interpolation error increasing linearly between 100 and 900 m. Similarly, RF and RFOK had a relatively gradual rate of change but had a steeper slope and more instances of sharp increases in error. The two remaining interpolators, RFSI and RFSIOK, were the least stable. After the separation distance increased to above 200 m, significant drops in performance were observed. Between 200 and 600m, the error increased by ten folds, which was significantly larger than what was observed in the four other interpolators. Overall, RK and OK showed the least sensitivity to changes in separation distance, followed by RF and RFOK, then RFSI and RFSIOK.

In terms of interpolation accuracy, OK was observed to be the most accurate interpolator; it had the lowest error at all separation distances except 600 m. The next best performers were the RF models, followed by the RFSI models, with the hybrid versions of these models having lower errors. These findings are somewhat contrary to what was found in previous literature, where RF, RFOK, and RFSI produced higher accuracy than OK and RK, and RK performed better than OK (Li et al., 2011). The RF-based models' relatively poor performance could be attributed to differences in the validation approach and interpolation task. While previous studies used random sampling to divide the data into training and validation sets, we manually removed data to create equally spaced gaps. This may have increased the difficulty of the interpolation task because the validation points were further away from the locations with measurements. Another factor is the difference in the interpolation task itself. For variables such as solar flare and mud sea content, the interpolator used measurements from all directions, whereas for friction, interpolation was limited to only two directions. Regarding the observation that OK outperformed RK, one possible explanation is that the implementation of a regression model was unnecessary; the dataset was already trend free, i.e., the additional step of trend removal did not improve performance since all it did was reposition the data.

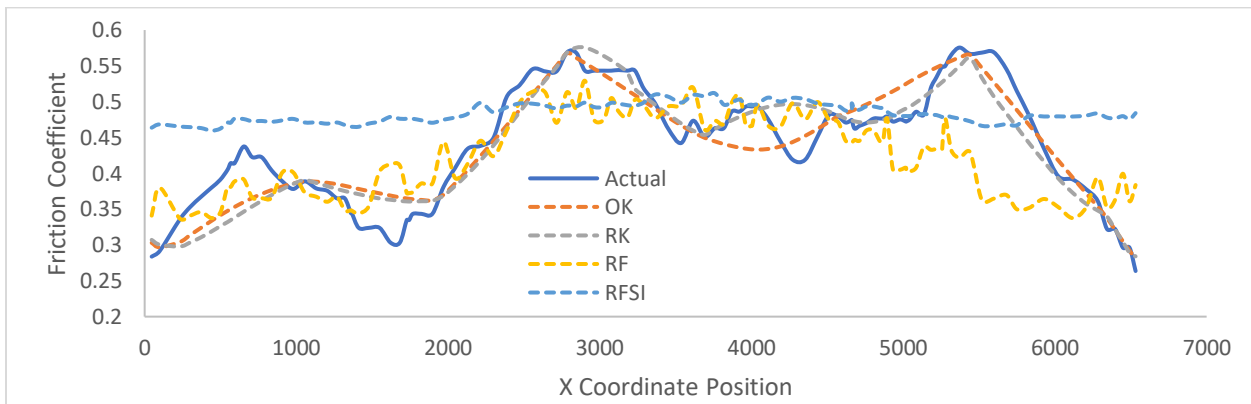
Aside from assessing the accuracy of interpolator performance using RMSE, it is crucial to analyze the form of interpolator predictions to verify that the pattern in the input dataset is being captured, and not simply predicting a constant value at every location. From 100 to 300 m, there was virtually



no difference between the six interpolators, as all could mimic the spatial pattern. Only after 300 m did the performance begin to diverge. The main difference was that RF and RFSI began to lose their ability to capture the local minimums and maximums found in the spatial pattern, which worsened as the separation distance increased (**Figure 5.2**). In other words, the amount of variation in the interpolations decreased as the separation distance increased. This issue was much less prominent in OK and RK. These two interpolators maintained their ability to mimic the spatial pattern even at 900 m. Comparatively, RF and RFSI at 900 m predict what was essentially a straight line, as shown in **Figure 5.3**.

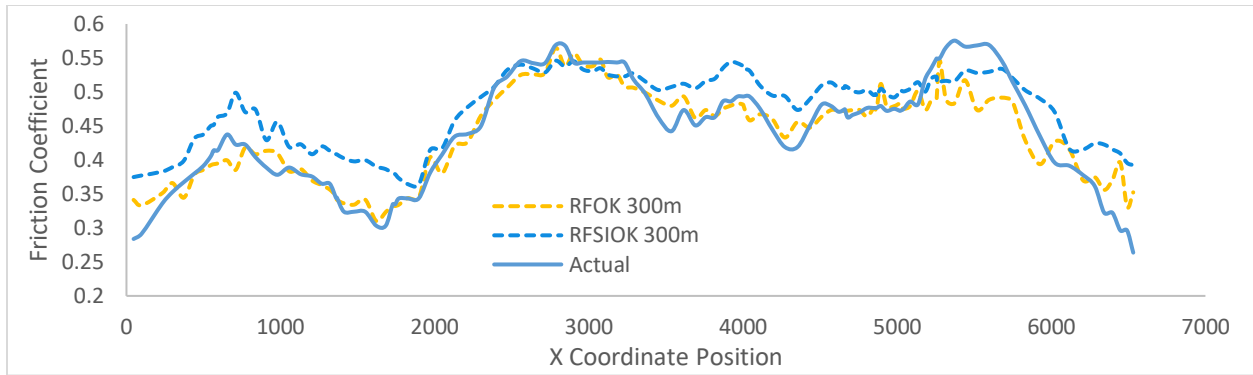


**Figure 5.2 Interpolator Performance at 300 m Separation Distance**



**Figure 5.3 Interpolator Performance at 900 m Separation Distance**

When RF and RFSI were combined with OK, both interpolators saw an improvement in their ability to capture the pattern found in the input dataset. Previously, these two interpolators began to perform poorly at 300 m. After the inclusion of OK, RFOK could mimic the spatial pattern until 900 m and RFSIOK up to 400 m; it can be said that adding OK to ML models can boost performance. **Figure 5.4** below depicts the interpolation improvements made through the inclusion of OK.

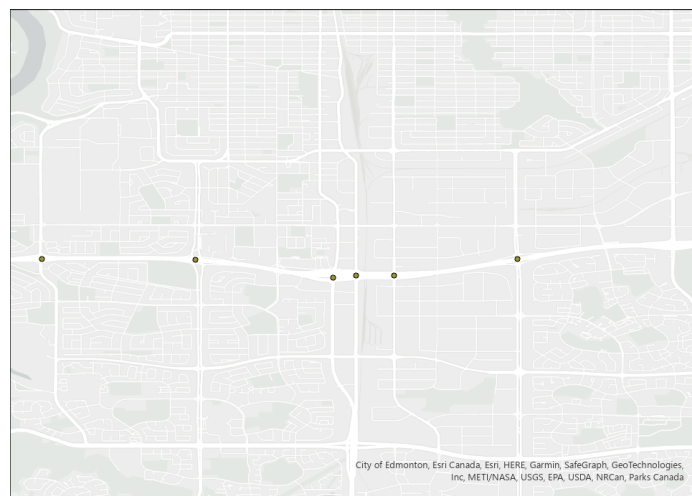


**Figure 5.4 ML Hybrid Model Performance at 300 m Separation Distance**

Ultimately, identical results were obtained from the interpolated values' error comparison and visual inspection. OK was the best of the six interpolators, requiring the fewest input variables while having the lowest error and the ability to mimic the spatial pattern closely.

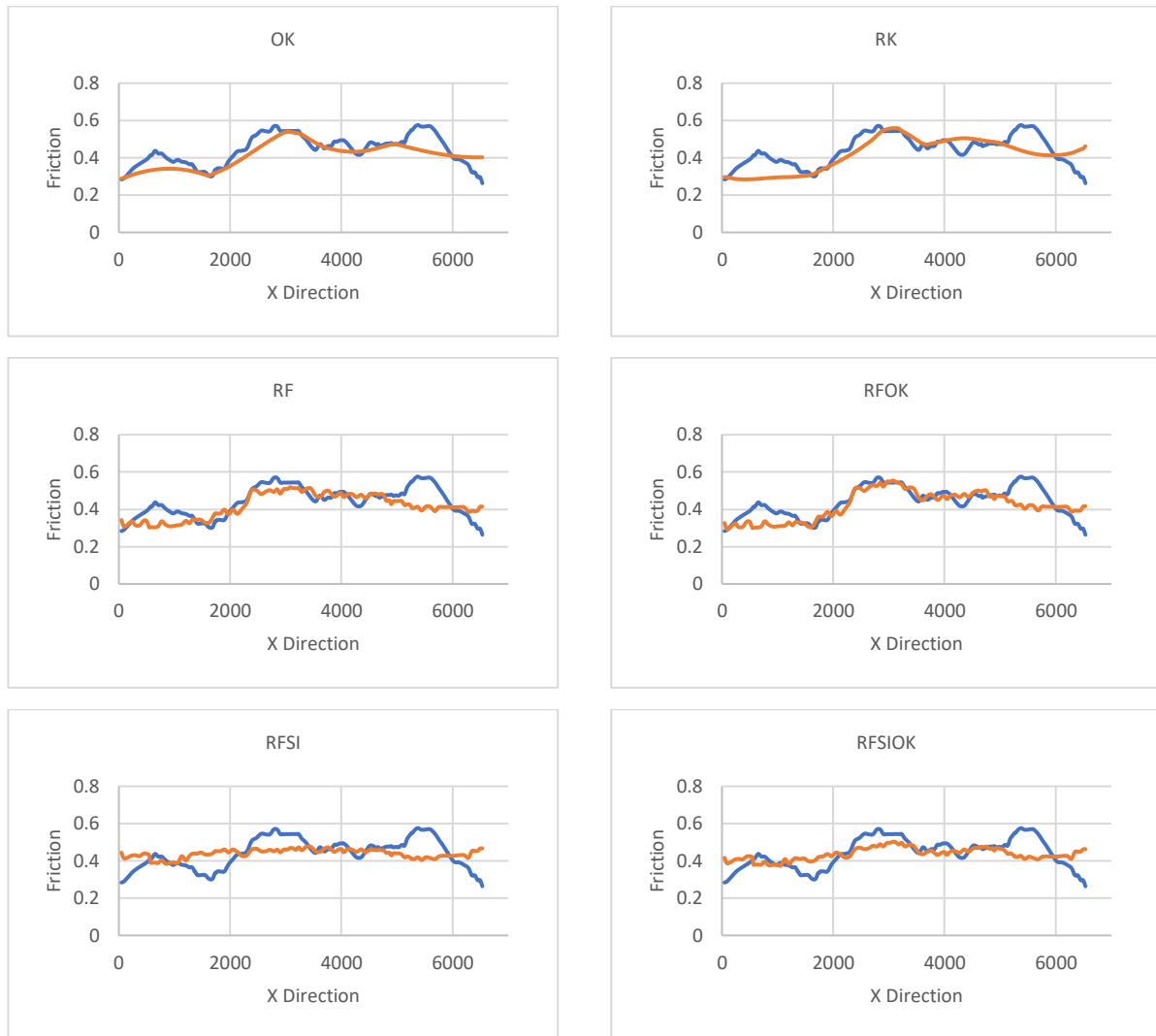
**Real-World Friction Analysis Using Traffic Cameras**

Based on the analysis above, it is clear that all six interpolators possess the ability to interpolate friction values. However, they were only evaluated in artificial scenarios where the observations are equally spaced apart, which is not typically found in the real world. For this reason, an additional dataset was created to simulate a real-world scenario where observations are only available at locations with traffic cameras to simulate friction collected by traffic cameras. The objective is to evaluate the feasibility of using traffic cameras to obtain continuous friction measurements. **Figure 5.5** depicts the dataset generated. It should be noted that the friction data used here was generated from the dash camera images and not from the traffic camera images.



**Figure 5.5 Friction Measurements Obtained at Locations with Traffic Cameras**

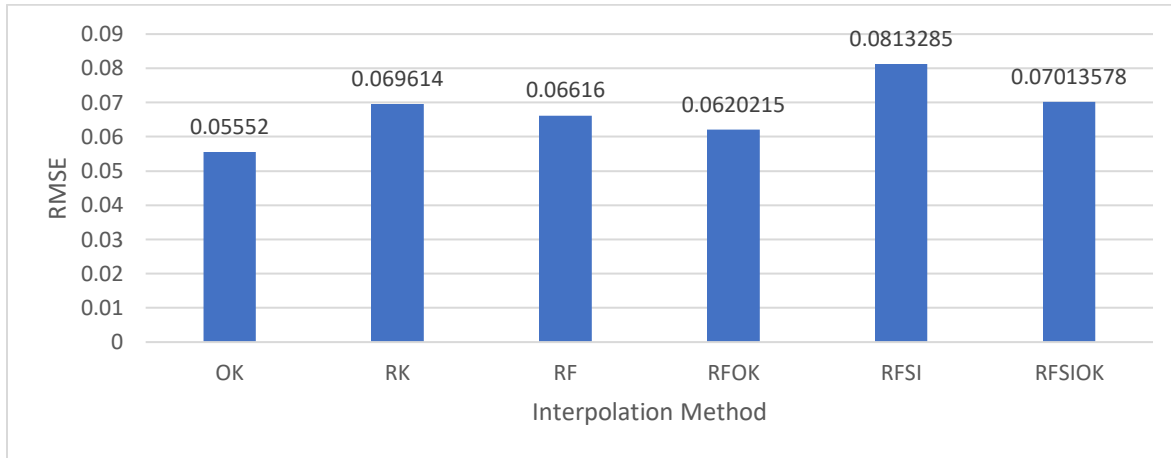
With the above observations as inputs, the following results were obtained from the six interpolators.



**Figure 5.6 Traffic Camera Interpolation Curves**

By looking at the interpolated curvature found in **Figure 5.6**, it appears that OK, RK, RF, and RFOK could capture the global variations with only five input measurements. Nevertheless, that are some differences; one is that OK and RK outputted much smoother estimations than the ML-based models. This observation is expected as Kriging is known to produce smooth estimates (Olea, 1999). Another difference is that the models with Kriging were better at capturing the peaks and valleys, which is consistent with the distance sensitivity analysis performed prior. In contrast, the RFSI-based did not appear to have captured the spatial trend, as its estimations are essentially constant.

From qualitative analysis alone, it is difficult to determine which interpolator performs better among OK, RK, RF, and RFOK, as their ability to mimic the spatial pattern is all very similar. For this reason, RMSE (shown in **Figure 5.7**) was calculated for all six interpolators to investigate interpolator performance more closely.



**Figure 5.7 Traffic Camera Interpolator Error Comparison**

Excluding the RFSI models that could not reproduce the spatial structure, the remaining four models all produced low errors. The average RMSE for the four models (OK, RK, RF, RFOK) is 0.063, among which OK performed best, with an RMSE of 0.05552.

It is also important to note that while pure ML models like RF and RFSI do not perform as well as OK or hybrid models, they have the advantage of requiring fewer data to develop. Kriging models require the development of a semivariogram, which is impossible to construct without a sufficiently large dataset. When it comes to measuring road surface friction, relying solely on traffic camera images is insufficient for constructing a semivariogram. To increase the number of measurements, a vehicle would need to drive through an entire section of the road. Only then can a semivariogram be constructed. Although this process may not be too difficult for a single road section, the data collection effort required for the entire road network would be extensive. In addition, it is unknown how long the semivariogram will remain valid. Image data may need to be collected periodically to update the model. In contrast, pure ML interpolators are independent of the semivariogram. They are built using only the data collected from the traffic camera locations, making them much easier to develop and update. Although RF is inferior in performance, RF may be the better choice if both accuracy and data requirements are considered.

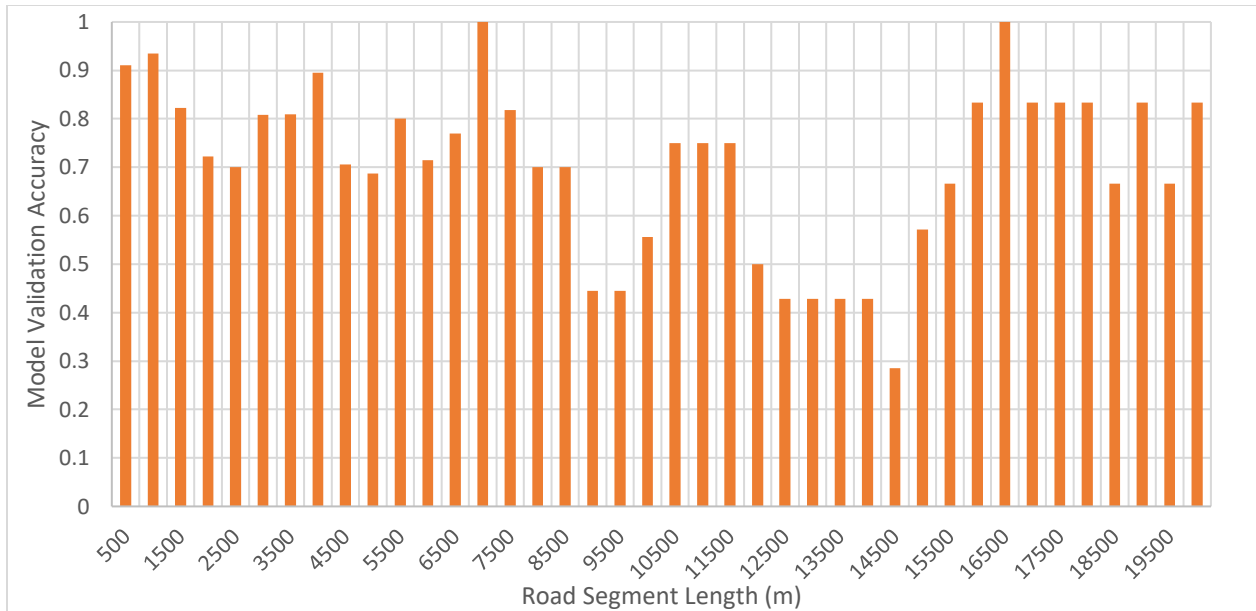
## 5.2 Collision Model Development for Safety Assessment

### Binary Collision Model

With the friction interpolator developed, the next step is to construct a binary collision likelihood model using continuous friction values. The main reason we opted for developing a binary collision model rather than a collision frequency model hinges on several factors. Primarily, binary classifications offer easier interpretation by presenting a clear “safe” or “unsafe” status, as opposed to relying on the potentially ambiguous interpretations derived from predicted collision numbers. Secondly, binary models excel in treating rare events such as collision—a domain where frequency models might struggle due to data sparsity. Lastly, binary models seamlessly integrate into decision-making systems by directly triggering actions like alerts or rerouting in connected vehicles (to be dealt with in subsequent sections), making them exceptionally practical for safety warnings.

To construct the binary model, the continuous friction values were averaged to represent how slippery a road segment is, for which the specific length has to be calibrated due to the class imbalance problem (Javaheri et al., 2014). If the segment length is too short or too long, the model will predict mostly one class because one category dominates most of the input data. Therefore, a calibration process is needed to determine the optimal segment length, which requires continuous friction values so that aggregation can be done at all lengths. In addition, model implementation also requires continuous measurements to ensure that the friction value assigned to each road segment is representative. Hence, because of the importance of spatially dense measurements, we developed a friction interpolator beforehand to demonstrate how continuous friction values can be obtained. Nevertheless, since we already had extremely dense friction values, interpolated values were not used for model development to prevent interpolation errors from carrying over to the collision model.

In this study, all models were developed using the decision tree algorithm due to it being non-parametric and highly interpretable. Concerning aggregation length, segment lengths between 500 m and 20 km (500 m increment) were evaluated to identify the optimal value. **Figure 5.8** shows the obtained results.



**Figure 5.8 Model Validation Accuracy at Various Road Segment Lengths**

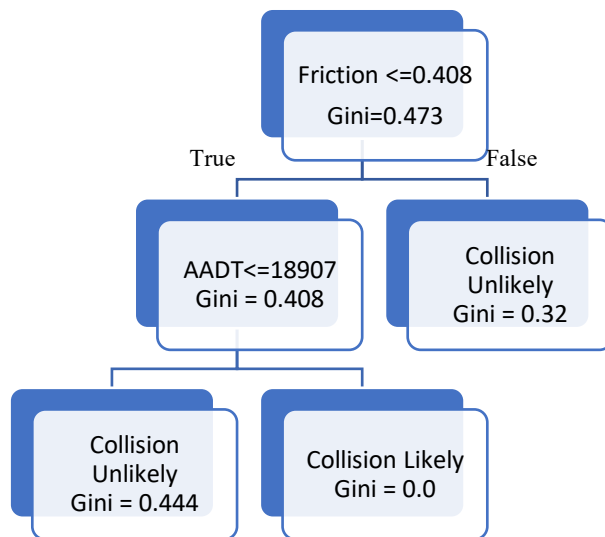
When evaluating each model, its training and validation dataset, model structure, and validation accuracy were examined. This process ensures that the model performance obtained is genuine and not due to class imbalance.

Between 500 m and 6 km, the validation accuracy fluctuated between 70 and 90%. The models developed at these segment lengths only had one node, meaning regardless of the input, the model always outputted “collision unlikely.” Consequently, the change in accuracy was not due to model structure but validation dataset differences. The observed accuracy only reflected the percentage of road segments in the validation dataset that did not have an associated collision event. From 6.5 to 8.5 km, due to the increase in the number of segments with collisions, the models utilized friction values as a key explanatory variable to identify “collision likely” segments. Other features like AADT, elevation, etc., were also used but not to the extent of friction that it was in every model. Between 9 and 11.5 km, a difference in the training dataset caused the accuracy difference observed in this region. The 9 and 9.5 km training dataset contained mostly segments with collisions, resulting in the model only outputting “collision likely.” In contrast, 10-11.5 km had a much more balanced dataset where the model utilized friction and other supporting variables. Since the validation dataset in this region contained a mix of segments with and without collisions, the 9 and 9.5 km performed much worse than the 10-11.5 km models. Starting at 12.5 km, friction was no longer used as a feature due to low feature variation resulting from averaging friction over

long distances. Now that the main feature was too homogeneous to be useful as a predictor, a drop in accuracy of about 30% was observed. From this point on, most of the models could only predict “collision likely,” meaning that validation accuracy became a measure of the proportion of road segments in the validation dataset associated with a collision event.

Based on the extent to which friction was used in the models developed, it is evident that friction is a crucial predictor of collision events. Excluding cases where significant class imbalances and friction values became too homogenized, every model had friction as a decision node, which cannot be said about any other variable selected in this study. In addition, when friction became a poor input variable because of low variation, a major accuracy decrease was observed, signifying the importance of the variable.

Overall, the optimal segment length was between 6.5 and 8.5 km because the friction values in this region were the least homogenized and had a balanced dataset. When we evaluated the internal logic of these models, we found that 7.0, 7.5, and 8.5 km produced highly complex models. Although they showed excellent performance, it was difficult to assess whether the generalizations made were reasonable. In contrast, the 6.5 km models were simple and intuitive (as shown in **Figure 5.9**) but had a slightly lower accuracy of 76.9%. Ultimately, the choice of model depends on user preference, as there is a trade-off between model performance and intuitiveness. Herein, intuitiveness was prioritized.



**Figure 5.9 Decision Tree Model Structure With 6.5 km Segment Length**

The model shown in **Figure 5.9** assumes that collisions are influenced by friction and AADT. Collisions are unlikely when the coefficient of friction is above 0.408. For road segments with friction less than 0.408, a traffic flow rate below 19,000 vehicles per day is considered safer. This model aligns with the Federal Highway Administration (FHWA) mandate of including AADT as a variable (Bauer & Srinivasan, 2013). Furthermore, these findings are in line with existing literature that associates higher AADT and lower friction with more collisions and identifies 0.38 as the unsafe friction threshold (Zhao et al., 2022).

### **Quantifying Safety Benefits in a CV Environment**

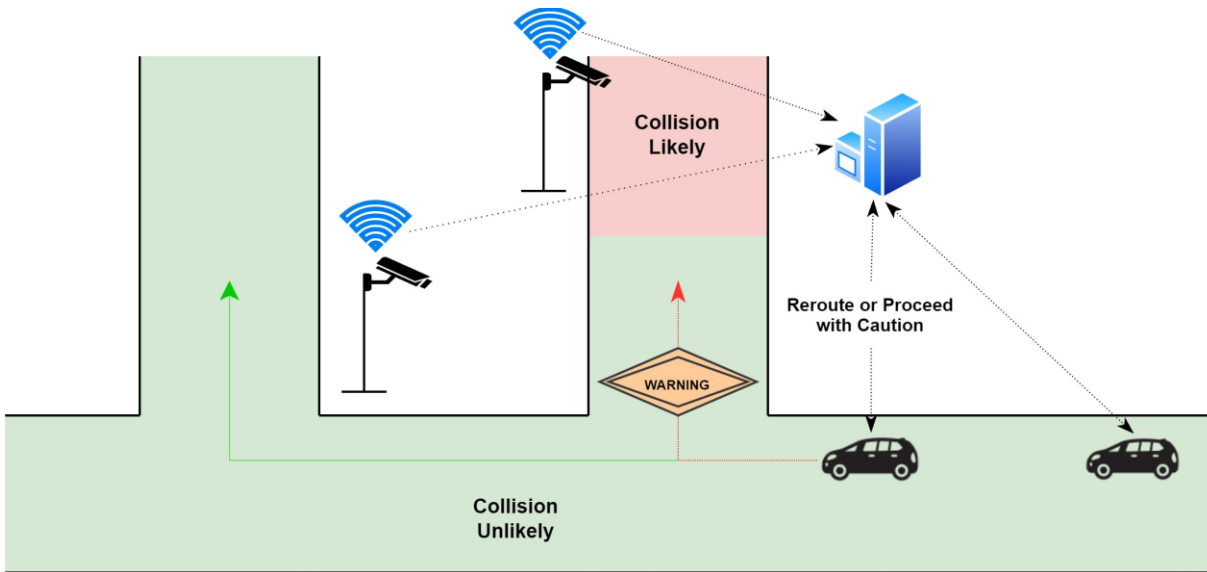
As we venture into a future where Connected Vehicles (CV) and Intelligent Transportation System (ITS) infrastructure are increasingly ubiquitous, the ability to estimate and quantify the safety benefits of such technologies becomes paramount. As demonstrated in the previous sections, spatially rich friction estimations have emerged as critical data points that could potentially enhance road safety. They pave the way for real-time RSC monitoring and data-driven interventions for improved traffic safety.

The proposed framework in this study capitalizes on these technologies and presents a potential use case to improve road safety during winter months. Assuming maturity and widespread availability of CV and ITS infrastructure, this framework offers a method for processing and harnessing the wealth of data captured by these technologies. In the envisioned scenario, snapshots of the RSCs captured by CVs or ITS equipment are transmitted to a nearby server. This server then converts these images into point friction values using the models previously developed and discussed. Subsequently, based on the attached location information, the point friction values are transformed into continuous friction values using an interpolator.

In a scenario where CVs and ITS equipment sufficiently saturate the road network, continuous friction values for the entire road network become accessible. The server can aggregate these friction values and convert them into collision likelihood ratings. This processed information is then relayed back to the CVs, triggering critical safety interventions. For instance, vehicles might be rerouted if an approaching road segment is identified as dangerous. If re-routing is not feasible, drivers receive warnings about potential dangers on the road ahead, thereby promoting cautious



driving. The entire process is repeated every few seconds to ensure constant provision of the most current information. **Figure 5.10** illustrates the above-mentioned process.



**Figure 5.10 Proposed Framework Implementation in CV and ITS Environment**

This sophisticated system lends to the quantification of safety benefits through crash avoidance estimations. For each 6.5 km road segment, the number of Property Damage Only (PDO), injuries, and fatalities were calculated for the four days where friction values were available. These segments were then evaluated through the binary collision likelihood model to ascertain the probability of a collision. It was assumed that all sections marked as “collision likely” had the potential to be avoided with the proposed framework, as the driver would have been prompted to re-route or warned to take necessary precautions. The resulting model demonstrated a labeling accuracy of 80%, identifying 11 road segments as potentially dangerous. These segments were linked with 22 PDOs and one injury.

In this study, two collision reduction scenarios were examined. The first scenario represented an ideal condition where all vehicles were effectively re-routed before reaching a hazardous road section, leading to a 100% Crash Reduction Factor (CRF). The second scenario applied a more conservative CRF of 18%, drawn from a similar study by the state of California on dynamic real-time road condition warning systems (Ye et al., 2012). It is worthwhile highlighting that this 18% CRF was employed for broad estimations of safety benefits when rerouting is unattainable. **Table 5.1** lists the number of collisions reduced and the safety benefits associated with each scenario.

**Table 5.1 Potential Number of Collisions Reduced Over Four Days Within the Study Area**

Collision Severity	Cost (Direct Plus Indirect)	Number of Collisions	Collision Reduction (100% CRF)	Safety Benefits (100% CRF)	Collision Reduction (18% CRF)	Safety Benefits (18% CRF)
PDO	\$14,065	22	22	\$ 309,430	4	\$ 56,260
Injury	\$137,749	1	1	\$ 137,749	0	\$ 0
Fatality	\$2,450,139	0	0	\$ 0	0	\$ 0
			Total	\$ 447,179	Total	\$ 56,260

According to **Table 5.1**, Scenarios 1 and 2 would have saved \$447,179 and \$56,260 over four days, respectively, if the proposed framework had been implemented. It is important to note that these monetary benefits only apply to a small portion of Edmonton’s road network; the actual savings should be greater.

In order to get a sense of the potential savings on a city-wide scale, traffic collision records for the same four days were obtained. Because these records did not have friction values, they could not be inputted into the developed model to determine whether a warning or re-routing would have been triggered. The solution is to use random sampling. Since our model operates at a 76.9% accuracy, we can assume that there is a 76.9% probability of our model detecting a dangerous road section. However, this assumption applies to only collisions that occurred on surface conditions our model was trained on—snowy and icy conditions. After filtering the data and removing duplicate reports, a total of 336 collisions remained. Among these collisions, a random sampling without replacement was implemented to select 76.9% of the dataset (294 collisions); these represent the hypothetical collision events that could have been avoided due to our model. In the case of scenario one, the process stopped here as all 294 collisions were considered avoidable due to re-routing. In comparison, for scenario 2, an additional sampling procedure was performed to select 18% of the 229 collisions to account for the 18% CRF.

**Table 5.2** identifies the potential number of collisions mitigated in each scenario using the proposed framework.

**Table 5.2 Potential Collision Reduction Over Four Days City-wide**

Collision Severity	Cost (Direct Plus Indirect)	Number of Collisions	Collision Reduction (100% CRF)	Safety Benefits (100% CRF)	Collision Reduction (18% CRF)	Safety Benefits (18% CRF)
PDO	\$14,065	243	243	\$ 3,417,795	42	\$ 590,730
Injury	\$137,749	14	14	\$ 1,928,486	4	\$ 550,996
Fatality	\$2,450,139	1	1	\$ 2,450,139	0	\$ 0
			Total	\$ 7,796,420	Total	\$ 1,141,726

The total savings for the four days was estimated to be \$7,796,420 (\$1,949,105 per day) and \$1,141,726 (\$285,431 per day) for scenarios one and two, respectively. These benefits were larger than those calculated in **Table 5.1**, which was expected since this was a city-wide evaluation. While the findings presented herein offer valuable insights into the potential safety benefits of utilizing Connected Vehicles (CV) for real-time hazardous road surface conditions monitoring, it is important to acknowledge that several assumptions have been made in generating these results. These assumptions include the availability of city-wide friction information, the model's ability to maintain its performance when data variation increases, the rerouting process successfully resolving the incident without causing issues elsewhere, and the driver responding to the safety message by slowing down. As a result of these assumptions, it may introduce certain limitations to the generalizability of our conclusions. Nonetheless, the proposed framework serves as a pioneering effort to quantify the safety benefits of this advanced technology. It provides a structured pathway for future research and practical applications, thereby underscoring the transformative potential of CV in enhancing road safety, especially during challenging weather conditions.

### 5.3 Summary

This chapter focuses on the development of the friction interpolator and the binary collision likelihood model.

In order to identify the most accurate interpolators, six interpolators were evaluated using datasets with increasing distances between measurements (from 0.1 to 1 km): Ordinary Kriging (OK), Regression Kriging (OK), Random Forest (RF), Random Forest Spatial Interpolator (RFSI), and hybrid models (RFOK and RFSIOK). The results show that OK had the lowest error and was least affected by increasing separation distance. Additionally, to evaluate each interpolator in a real-world scenario where measurements are not evenly spaced, only friction measurements near the City of Edmonton traffic cameras were used as input data. The results obtained were identical to the distance sensitivity analysis, where OK produced the lowest error and demonstrated the ability to capture the spatial pattern with only five available observations.

After demonstrating that continuous friction values could be accurately generated, we developed binary collision models using segment lengths from 500 m to 20 km. Based on model accuracy and intuitiveness, the most optimal model was found at 6.5 km, which used friction and AADT as parameters and displayed a classification accuracy of 76.9%. The proposed framework can be used to generate advance warnings for road users, which, if implemented in an environment with intelligent transportation systems and connected vehicles, there is an expected minimum safety benefit of approximately \$285,431 per day city-wide.

## Chapter 6 Conclusion and Future Work

This chapter provides an overview of this thesis and highlights the major findings and contributions of the research conducted herein. In addition, limitations of the research and recommendations for future research can also be found in this chapter.

### 6.1 Overview

The provision of real-time winter Road Surface Condition (RSC) information has the potential to improve road safety by allowing road users to take the necessary precautions needed on hazardous roads or avoid them altogether. This potential, however, has yet to be realized fully due to limitations in how RSC is represented, its spatial coverage, and its update frequency. In most jurisdictions, winter RSC is categorized based on the degree of snow coverage, using terms such as bare, partially snow-covered, and fully snow-covered. This kind of system is problematic because the determination process is subjective, and the intermediate classes can represent both safe and unsafe conditions simultaneously. Both problems are solved by using friction. Because friction is a measured variable, it eliminates the need for subjective interpretation, and since friction is a continuous measure between 0 and 1, it allows for the establishment of safety thresholds to separate risky from non-risky roads. Nevertheless, there is currently a lack of cost-effective methods for measuring friction, preventing it from being implemented at a large scale.

In terms of spatial coverage and update frequency, these issues are driven by the cost of collection. It is simply not feasible to collect dense friction measurements on a regular basis without a significant increase in cost. As a result, the RSC data collected is sparse, and the update frequency can be as little as once a day, which means it is likely that the information provided does not reflect actual conditions.

To address these challenging issues, this thesis proposes a novel framework that leverages advances in computer vision and ML. Computer vision makes it possible to effortlessly convert images into predictor features, which are then utilized by ML algorithms to learn the relationship between friction and the extracted features. Compared to traditional collection methods, our proposed framework is more effortless and cost-effective as the collection process is fully automated, all of which are improvements that make the proposed framework more appealing for

large-scale implementations. At the core of our proposed framework are three interconnected components. Component one is an automated image-based friction model to predict friction values; component two is a friction interpolator to generate continuous friction values from limited point measurements; and the final component is a binary collision model to identify high-risk road sections. The combination of these three components forms a novel framework that transforms winter road images into valuable information that allows road users to take precautions on risky road sections and maintenance personnel for targeted treatment. Furthermore, with the proliferation of CVs and ITS equipment, this framework provides the ability to transform the wealth of information collected into valuable real-time RSC data.

## 6.2 Research Findings

### Component One: Image-based Friction Model

- The image-based friction prediction model was developed using 128 pairs of images and friction values. These images were converted into predictor features by RSC categorization, Local Binary Pattern, Gray Level Co-occurrence matrix, and image thresholding. Then, three tree-based algorithms—Decision Tree, Random Forest, and Gradient Boosting—were used to find the relationship between the extracted predictor features and measured friction values. All three displayed high performance with an average RMSE of 0.0796 or 79.3% accuracy based on RMSPE.
- A method called permutation feature importance was used to determine the relative importance of the features based on their impact on the prediction error. The results show that RSC categorization is the most important feature extraction method—a model without RSC as a feature produced seven times more errors than a model with RSC.
- A Convolution Neural Network (CNN) model was used to fully automate the friction model. The highest-performing CNN model found had a validation accuracy of 85.8%. When applied to the friction model’s validation dataset, it labeled all RSC categories with over 85% accuracy. To determine the effect of including a CNN model in the friction prediction process, friction predictions were generated using the CNN-generated labels. The results show that adding a CNN model had minimal impact on model performance; only a 0.05 and 0.6% increase in RMSE and RMSPE were observed, respectively.

### **Component Two: Friction Interpolator**

- Six Interpolators—Ordinary Kriging (OK), Regression Kriging (RK), RF, RF Spatial Interpolator (RFSI), RFOK, and RFSIOK—were evaluated on datasets with increasing separation distance (from 100 m to 1 km with a step size of 100), OK produced the highest accuracy among these interpolators while showing the least sensitivity.
- The same six interpolators were evaluated using a dataset with measurements only available at locations with traffic cameras. This was done to simulate the idea of using traffic camera images to generate continuous friction values. Similar to the sensitivity analysis, OK was found to be the highest-performing interpolator with an RMSE of 0.05552 and the ability to capture the spatial pattern.

### **Component Three: Binary Collision Model**

- A binary collision model was developed using segment length calibration to find the optimal length at which there was a balanced number of road sections with and without collisions. In an effort to find this distance, at each segment length from 500 m to 20 km with an interval of 500 m, a decision tree model was developed. Among these models, the 6.5 km segment model was selected based on model accuracy and intuitiveness. It had a 76.9% accuracy and utilized AADT and friction as input features. By examining the internal logic of the 6.5 km friction model, it was found that lower friction and higher AADT make collisions more likely.
- The proposed framework has the potential to offer real-time collision likelihood ratings for every vehicle on the road network by using continuous friction values obtained from connected vehicles and Intelligent transportation systems equipped with cameras. This can help reroute vehicles to avoid risky road sections or warn road users of upcoming risky conditions. If implemented, the framework would provide significant safety benefits.

### **6.3 Research Contributions**

This thesis has made significant methodological and practical contributions to the field of winter road condition monitoring and safety.

#### **Methodological Contributions**

Methodologically, a novel framework is proposed that uses image processing techniques and machine learning algorithms to convert road surface images into continuous friction measurements. This approach addresses the consistency and subjectivity issues inherent in qualitative RSC descriptors and overcomes the challenges associated with the high cost of friction data collection.

Until this thesis, the plausibility of developing a winter road friction model using road imagery as the sole source of information remained uncertain, as none of the existing efforts in this area appear to have been based on actual friction data. This study marks the first time that a model with robust performance has been developed using collected friction data, demonstrating the potential of image-based friction collection methods. In addition, this work also serves as a reference in the selection of predictive features for future research. Not only have we identified the features used in friction prediction, but we have also ascertained the most critical feature for friction prediction through feature importance. The benefit of this finding is that future researchers will know which features to prioritize when developing models and which features to replace with alternative predictor features extracted using other computer vision techniques.

This thesis is also a pioneer in the exploration of the performance of ML and ML hybrid interpolators on road-related variables. The use of ML as interpolators has only been explored in environmental studies with measurements available in all directions, where ML showed higher accuracy than geostatistical methods. In this thesis, we have rigorously evaluated the performance of ML interpolators in different scenarios to determine their advantages and disadvantages, which is an invaluable reference that future modelers can refer to when choosing the optimal interpolation method for their specific circumstances. Lastly, from the development of the binary collision model, we were able to validate the relationship between friction and winter collisions. Prior to this study, only one author had investigated this relationship. This work further substantiates the notion that lower friction leads to a higher probability of collisions, highlighting the importance of monitoring RSC via friction.



### **Practical Contributions**

Practically, the framework introduced in this thesis is an innovative solution that can greatly enhance winter road condition monitoring and forecasting. By using continuous friction measurements in conjunction with connected vehicle technology and intelligent sensors, we can envision a system that not only predicts hazardous conditions, but also proactively alerts drivers or reroutes vehicles to improve road safety. The implementation of such a system can lead to improvements in traffic management, road safety, and the overall driving experience during the winter season, ultimately leading to a potential reduction in winter-related accidents and their associated costs.

The benefits of the proposed framework extend beyond prospective implementation; they also enhance the value of existing infrastructure and equipment. For example, the current need for hiring contractors to manually inspect the road surfaces to determine RSC can be streamlined using a dash camera to capture road surface footage. Our proposed framework can then process the collected footage to generate dense friction measurements, which can be converted into collision likelihood measures and published on platforms like 511 for advance trip planning. This paradigm shift implies that the time contractors spend inspecting the road surface can now be allocated to collecting more data, thereby increasing the spatial coverage of RSC information. For urban road networks, dash cameras can be installed on buses to collect road condition information. Since buses travel over most of a city's road network at regular intervals, they are a great way to collect spatially extensive and timely information with minimal additional investment. Furthermore, existing infrastructure with imaging capabilities, such as traffic cameras, can be integrated into the RSC collection process. The images collected by these devices can be converted into friction values and collision likelihood readings. With sufficient camera density on a given road section, producing continuous RSC information may be possible, as demonstrated in the traffic camera case study. A compelling advantage of this approach is that because these devices collect data around the clock, the update frequency of RSC information can be adjusted to any interval without a significant increase in expenditure, thereby solving the update frequency issue.

This thesis presents innovative advancements in monitoring winter road conditions that could create a new path for enhancing road safety. By combining image processing techniques, machine

learning algorithms, interpolation methods, and existing infrastructures; this study improves our understanding of road safety management during the winter season.

#### **6.4 Limitations and Future Research**

The research conducted in this thesis shows that the conversion of winter road surface images into continuous friction values is possible and can be done accurately. It also indicates that a collision model that identifies high-risk road sections can be developed based on the resulting continuous friction measurements. However, there are a number of limitations that should be addressed in future studies.

- Investigate how the friction model performance changes when more friction and supplementary maintenance operation data are available. Increasing friction dataset size allows us to evaluate whether the accuracy exhibited by our proof of concept model persists when dataset variation increases, and it may reveal more information regarding why certain generic features extracted from LBP and GLCM are deemed important.
- Replace image thresholding with Semantic Segmentation (SS), as SS allows for more accurate pixel separation between asphalt and road precipitants. It is possible that the inaccuracies introduced during the separation process caused the thresholding features to be underused as input predictors.
- Explore the possibility of using maintenance data as predictors. Maintenance data provides information on which roads have been treated with road additives, which are known to impact friction. Incorporating this kind of information has the potential to make the proposed friction model more accurate.
- Re-evaluate the interpolators over a much longer stretch of road. This will allow more training data to be used for model development. The superior performance of OK may be due to the road segment being relatively short in our case study.
- Develop a traffic camera-focused friction prediction model. The model we created in this project was built using dash camera images; friction tests were not performed in the vicinity of traffic cameras to allow us to develop a model that linked traffic camera images with friction values. Although we could use the friction model developed in this study with traffic camera images, the accuracy would suffer significantly due to vastly different

camera angles and environmental factors. Creating such a model would require the exploration of potentially more advanced machine learning algorithms to compensate for the constantly changing traffic camera angles and the increase in environmental variables like vehicles, traffic signals, and other road infrastructure.

- Confirm the relationship identified between friction and collisions using collision records from an entire winter season. The current model was developed using four days of collision data because these were the only four days for which friction data were available. Increasing the size of the model development dataset will allow a more representative model to be developed, which may provide more insight into the friction-collision relationship.
- Adopt an extreme value theory approach for modeling collision likelihood. This change would allow us to use collision surrogate measures like time-to-collision for model development. Unlike actual collision data, time-to-collision data and its corresponding friction value can be generated using simulation software, providing us with more data to develop a more reliable collision likelihood model.

## References

- 511 Alberta. (2023). <https://511.alberta.ca/>
- Abohassan, A., El-Basyouny, K., & Kwon, T. J. (2021). Exploring the associations between winter maintenance operations, weather variables, surface condition, and road safety: A path analysis approach. *Accident Analysis and Prevention*, 163. <https://doi.org/10.1016/j.aap.2021.106448>
- Abohassan, A., El-Basyouny, K., & Kwon, T. J. (2022). Effects of Inclement Weather Events on Road Surface Conditions and Traffic Safety: An Event-Based Empirical Analysis Framework. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(10), 51–62. <https://doi.org/10.1177/03611981221088588>
- Al-Qadi, I. L., Loulizi, A., Flintsch, G. W., Roosevelt, D. S., Decker, R., & Wambold, J. C. (2002). *Feasibility of Using Friction Indicators to Improve Winter Maintenance Operations and Mobility*.
- Armi, L., & Fekri-Ershad, S. (2019). TEXTURE IMAGE ANALYSIS AND TEXTURE CLASSIFICATION METHODS-A REVIEW. In *International Online Journal of Image Processing and Pattern Recognition* (Vol. 2, Issue 1).
- Bauer, K., & Srinivasan, R. (2013). *Safety Performance Function Development Guide: Developing Jurisdiction Specific SPFs*.
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Carrillo, J., Crowley, M., Pan, G., & Fu, L. (2019). *Design of Efficient Deep Learning models for Determining Road Surface Condition from Roadside Camera Images and Weather Data*.
- Chen, L. (2009). Curse of Dimensionality. In *Encyclopedia of Database Systems* (pp. 545–546). Springer US. [https://doi.org/10.1007/978-0-387-39940-9\\_133](https://doi.org/10.1007/978-0-387-39940-9_133)
- Das, S., Tsapakis, I., & Khodadadi, A. (2021). Safety performance functions for low-volume rural minor collector two-lane roadways. *IATSS Research*, 45(3), 347–356. <https://doi.org/10.1016/j.iatssr.2021.02.004>
- Davies, E. R. (2005). *Machine Vision : Theory, Algorithms, Practicalities* (3rd ed.). Morgan Kaufmann.

- Du, Y., Liu, C., Song, Y., Li, Y., & Shen, Y. (2020). Rapid Estimation of Road Friction for Anti-Skid Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*, 21(6), 2461–2470. <https://doi.org/10.1109/TITS.2019.2918567>
- Erdogan, G., Alexander, L., & Rajamani, R. (2011). Estimation of Tire-Road Friction Coefficient Using a Novel Wireless Piezoelectric Tire Sensor. *IEEE Sensors Journal*, 11(2), 267–279. <https://doi.org/10.1109/JSEN.2010.2053198>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5). <https://doi.org/10.1214/aos/1013203451>
- Gaur, M., Faldu, K., & Sheth, A. (2021). Semantics of the Black-Box: Can Knowledge Graphs Help Make Deep Learning Systems More Interpretable and Explainable? *IEEE Internet Computing*, 25(1), 51–59. <https://doi.org/10.1109/MIC.2020.3031769>
- Giles, C. G., Sabey, B. E., & Cardew, K. H. F. (1965). Development and Performance of the Portable Skid Resistance Tester. *Rubber Chemistry and Technology*, 38(4), 840–862. <https://doi.org/10.5254/1.3535703>
- Gu, L. (2019). *Developing Models for Estimating Winter Road Weather and Surface Conditions—An Empirical Investigation*. University of Alberta.
- Hall, J. W., Smith, K. L., Titus-Glover, L., Wambold, J. C., Yager, T. J., & Rado, Z. (2009). *Guide for Pavement Friction*. <https://doi.org/10.17226/23038>
- Hall-Beyer, M. (2017). GLCM Texture: A Tutorial. In *PRISM: University of Calgary's Digital Repository*. <https://doi.org/http://dx.doi.org/10.11575/PRISM/33280>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer New York. <https://doi.org/10.1007/978-0-387-84858-7>
- Hohn, M. E. (1999). The semivariogram. In *Geostatistics and petroleum geology* (pp. 15–80). Springer Netherlands. [https://doi.org/10.1007/978-94-011-4425-4\\_2](https://doi.org/10.1007/978-94-011-4425-4_2)
- Iowa DOT 511. (2023). <https://www.511ia.org/@-94.9603,42.19251,7?show=iowaAppIncident,iowaAppRoadwork,weatherWarningsAreaEvents,otherStateInfo>
- Javaheri, S. H., Sepehri, M. M., & Teimourpour, B. (2014). Response Modeling in Direct Marketing. In *Data Mining Applications with R* (pp. 153–180). Elsevier. <https://doi.org/10.1016/B978-0-12-411511-8.00006-2>

- Kensert, A., Harrison, P. J., & Spjuth, O. (2019). Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. *SLAS Discovery*, 24(4), 466–475. <https://doi.org/10.1177/2472555218818756>
- Khaleghian, S., Emami, A., & Taheri, S. (2017). A technical survey on tire-road friction estimation. *Friction*, 5(2), 123–146. <https://doi.org/10.1007/s40544-017-0151-0>
- Khaleghian, S., Ghasemalizadeh, O., & Taheri, S. (2016). Estimation of the Tire Contact Patch Length and Normal Load Using Intelligent Tires and Its Application in Small Ground Robot to Estimate the Tire-Road Friction. *Tire Science and Technology*, 44(4), 248–261. <https://doi.org/10.2346/tire.16.440402>
- Kuttesch, J. S. (2004). *Quantifying the Relationship Between Skid Resistance and Wet Weather Accidents for Virginia Data*. Virginia Polytechnic Institute and State University.
- Lee, H. S., Hall, J., & Stanley, M. (2019). *Develop Statistical Models Quantifying the Relationship Between Pavement Surface Friction Characteristics and Traffic Accident Rates*.
- Leirvik, T., & Yuan, M. (2021). A Machine Learning Technique for Spatial Interpolation of Solar Radiation Observations. *Earth and Space Science*, 8(4). <https://doi.org/10.1029/2020EA001527>
- Li, J., Heap, A. D., Potter, A., & Daniell, J. J. (2011). Application of machine learning methods to spatial interpolation of environmental variables. *Environmental Modelling & Software*, 26(12), 1647–1659. <https://doi.org/10.1016/j.envsoft.2011.07.004>
- Matilainen, M. J., & Tuononen, A. J. (2011). Tire friction potential estimation from measured tie rod forces. *2011 IEEE Intelligent Vehicles Symposium (IV)*, 320–325. <https://doi.org/10.1109/IVS.2011.5940528>
- McCarthy, R., Flintsch, G. W., Katicha, S. W., McGhee, K. K., & Medina-Flintsch, A. (2016). New Approach for Managing Pavement Friction and Reducing Road Crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2591(1), 23–32. <https://doi.org/10.3141/2591-04>
- Molnar, C. (2022). *Interpretable Machine Learning* (2nd ed.).
- Niskanen, A., & Tuononen, A. (2015). Three Three-Axis IEPE Accelerometers on the Inner Liner of a Tire for Finding the Tire-Road Friction Potential Indicators. *Sensors*, 15(8), 19251–19263. <https://doi.org/10.3390/s150819251>

- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987.  
<https://doi.org/10.1109/TPAMI.2002.1017623>
- Olea, R. A. (1999). *Geostatistics for Engineers and Earth Scientists*. Springer US.  
<https://doi.org/10.1007/978-1-4615-5001-3>
- Olea, R. A. (2006). A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment*, 20(5), 307–318.  
<https://doi.org/10.1007/s00477-005-0026-1>
- Pichler, M., Boreux, V., Klein, A., Schleuning, M., & Hartig, F. (2020). Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, 11(2), 281–293. <https://doi.org/10.1111/2041-210X.13329>
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.  
<https://doi.org/10.1007/BF00116251>
- Roychowdhury, S., Zhao, M., Wallin, A., Ohlsson, N., & Jonasson, M. (2018). Machine Learning Models for Road Surface and Friction Estimation using Front-Camera Images. *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8.  
<https://doi.org/10.1109/IJCNN.2018.8489188>
- Schulze, K., Gerbaldi, A., & Chavet, J. (1977). Skidding Accidents, Friction Numbers, and the Legal Aspects Involved Report of the PIARC Technical Committee on Slipperiness and Evenness. *Transportation Research Record*.
- Sekulić, A., Kilibarda, M., Heuvelink, G. B. M., Nikolić, M., & Bajat, B. (2020). Random Forest Spatial Interpolation. *Remote Sensing*, 12(10), 1687. <https://doi.org/10.3390/rs12101687>
- Smoothing Time Series*. (2023). STAT 510 The Pennsylvania State University.  
<https://online.stat.psu.edu/stat510/lesson/5/5.2>
- TCHANGOU TOUDJEU, I., & TAPAMO, J.-R. (2019). Circular Derivative Local Binary Pattern Feature Description for Facial Expression Recognition. *Advances in Electrical and Computer Engineering*, 19(1), 51–56. <https://doi.org/10.4316/AECE.2019.01007>
- Wackernagel, H. (1995). Ordinary Kriging. In *Multivariate Geostatistics* (pp. 74–81). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-03098-1\\_11](https://doi.org/10.1007/978-3-662-03098-1_11)

- Wallman, C.-G., & Åström, H. (2001). *Friction measurement methods and the correlation between road friction and traffic safety*.
- Wu, M., & Kwon, T. J. (2022, January 12). Continuous Mapping of Winter Road Surface Conditions via Deep Learning and Geostatistics. *Transportation Research Board 101st Annual Meeting*.
- Xie, Q., & Kwon, T. J. (2022). Development of a Highly Transferable Urban Winter Road Surface Classification Model: A Deep Learning Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(10), 445–459.  
<https://doi.org/10.1177/03611981221090235>
- Xie, Q., & Kwon, T. J. (2023). Developing Machine Learning-based Approach for Predicting Road Surface Frictions using Dashcam Images – A City of Edmonton, Canada, Case Study. *Canadian Journal of Civil Engineering*. <https://doi.org/10.1139/cjce-2023-0015>
- Yang, G., Wang, K. C. P., & Li, J. Q. (2021). Multiresolution analysis of three-dimensional (3D) surface texture for asphalt pavement friction estimation. *International Journal of Pavement Engineering*, 22(14), 1882–1891. <https://doi.org/10.1080/10298436.2020.1726350>
- Ye, Z., Veneziano, D., & Turnbull, I. (2012). Safety Effects of Icy-Curve Warning Systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2318(1), 83–89. <https://doi.org/10.3141/2318-10>
- Zhao, G., Jiang, Y., Li, S., & Tighe, S. (2022). Exploring implicit relationships between pavement surface friction and vehicle crash severity using interpretable extreme gradient boosting method. *Canadian Journal of Civil Engineering*, 49(7), 1206–1219.  
<https://doi.org/10.1139/cjce-2021-0337>
- Zhu, Q., & Lin, H. S. (2010). Comparing Ordinary Kriging and Regression Kriging for Soil Properties in Contrasting Landscapes. *Pedosphere*, 20(5), 594–606.  
[https://doi.org/10.1016/S1002-0160\(10\)60049-5](https://doi.org/10.1016/S1002-0160(10)60049-5)



# Appendix

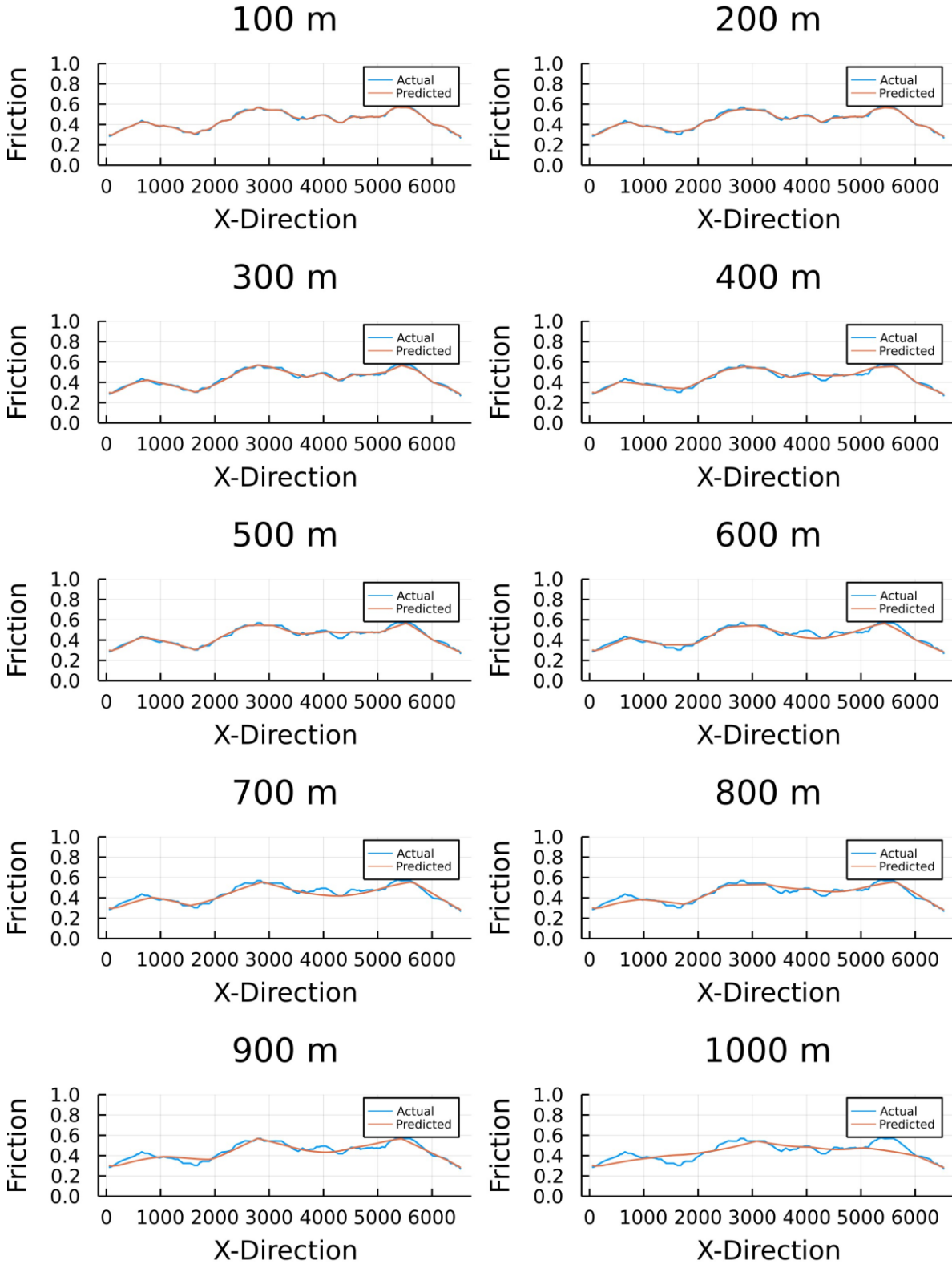
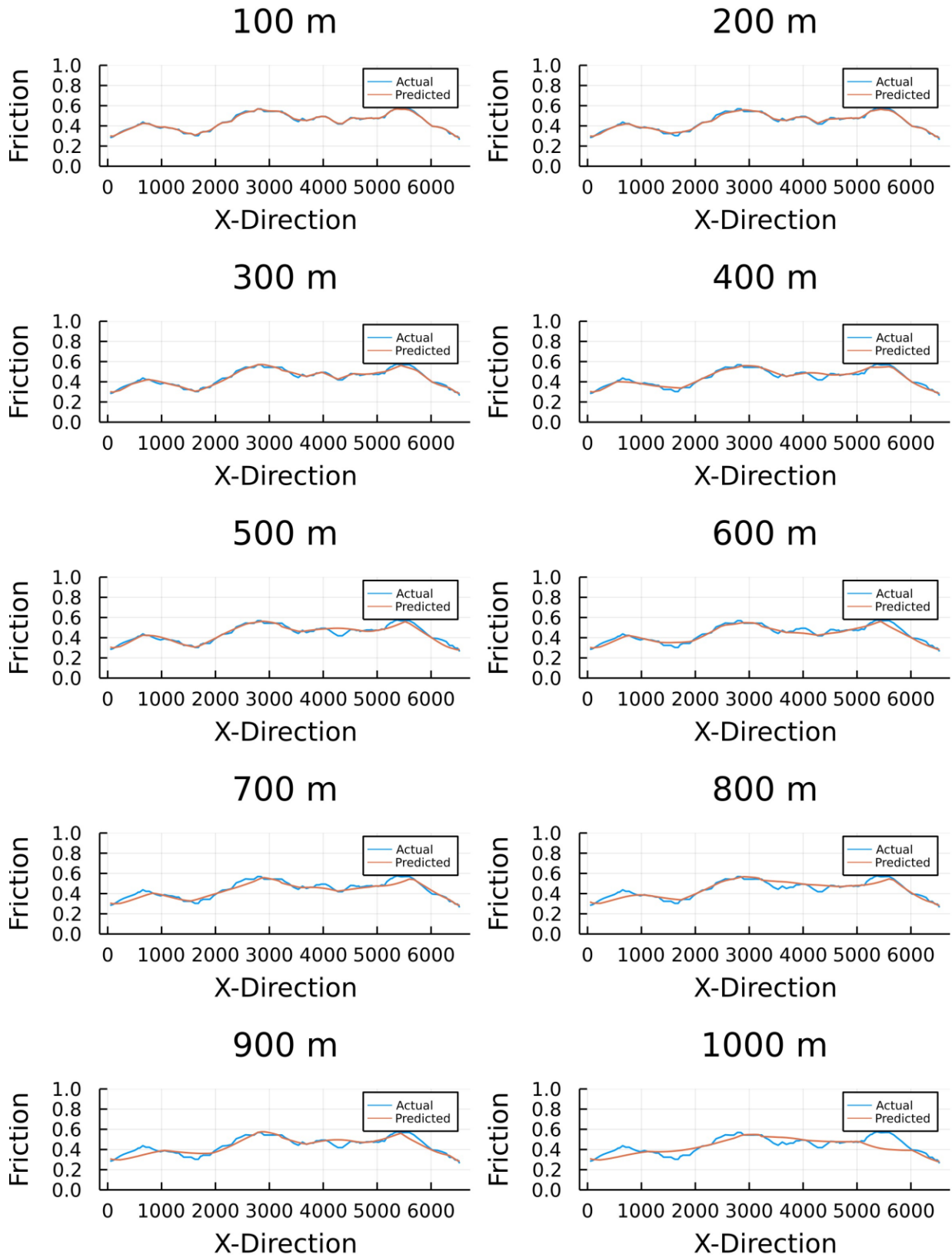
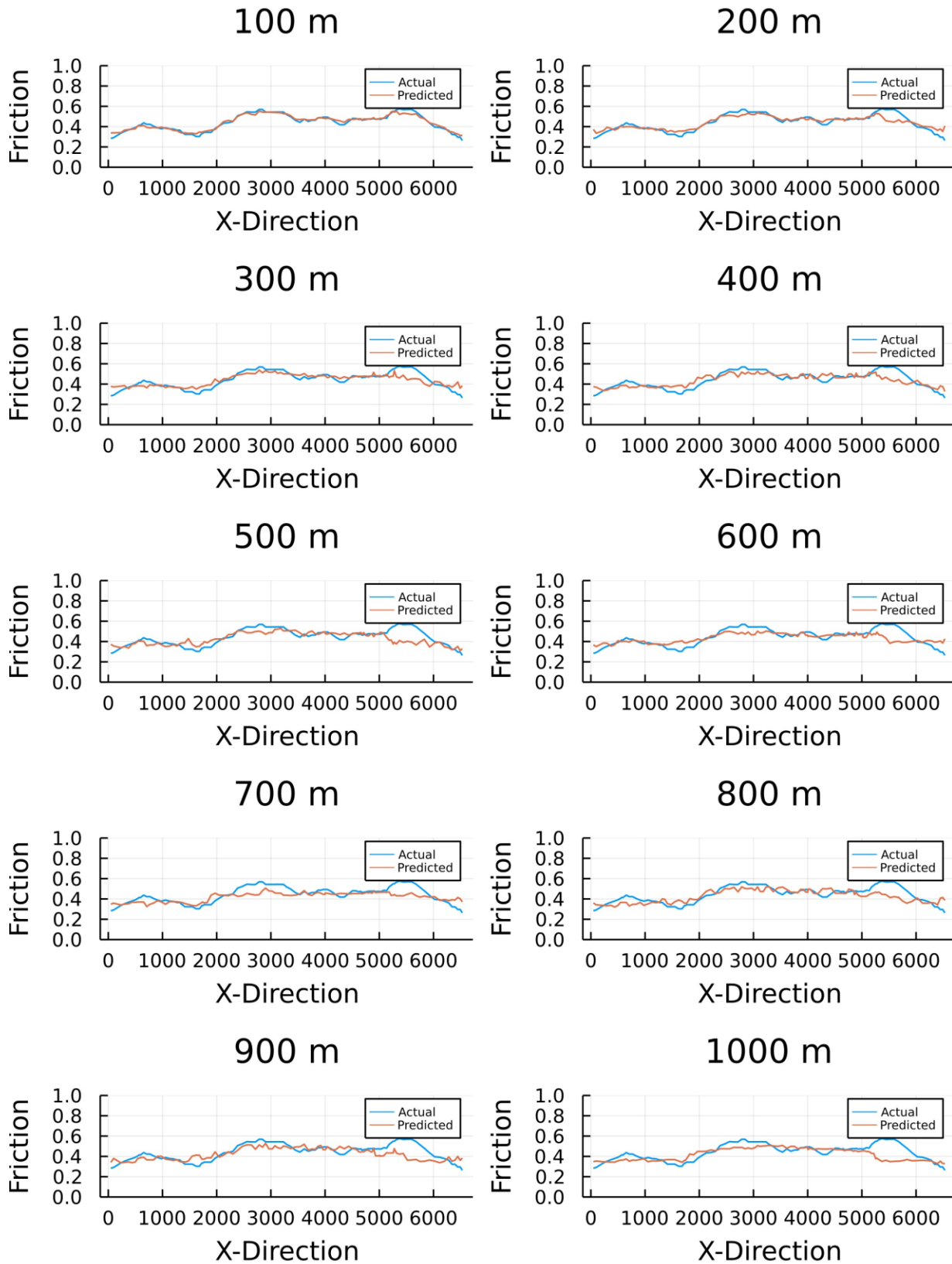


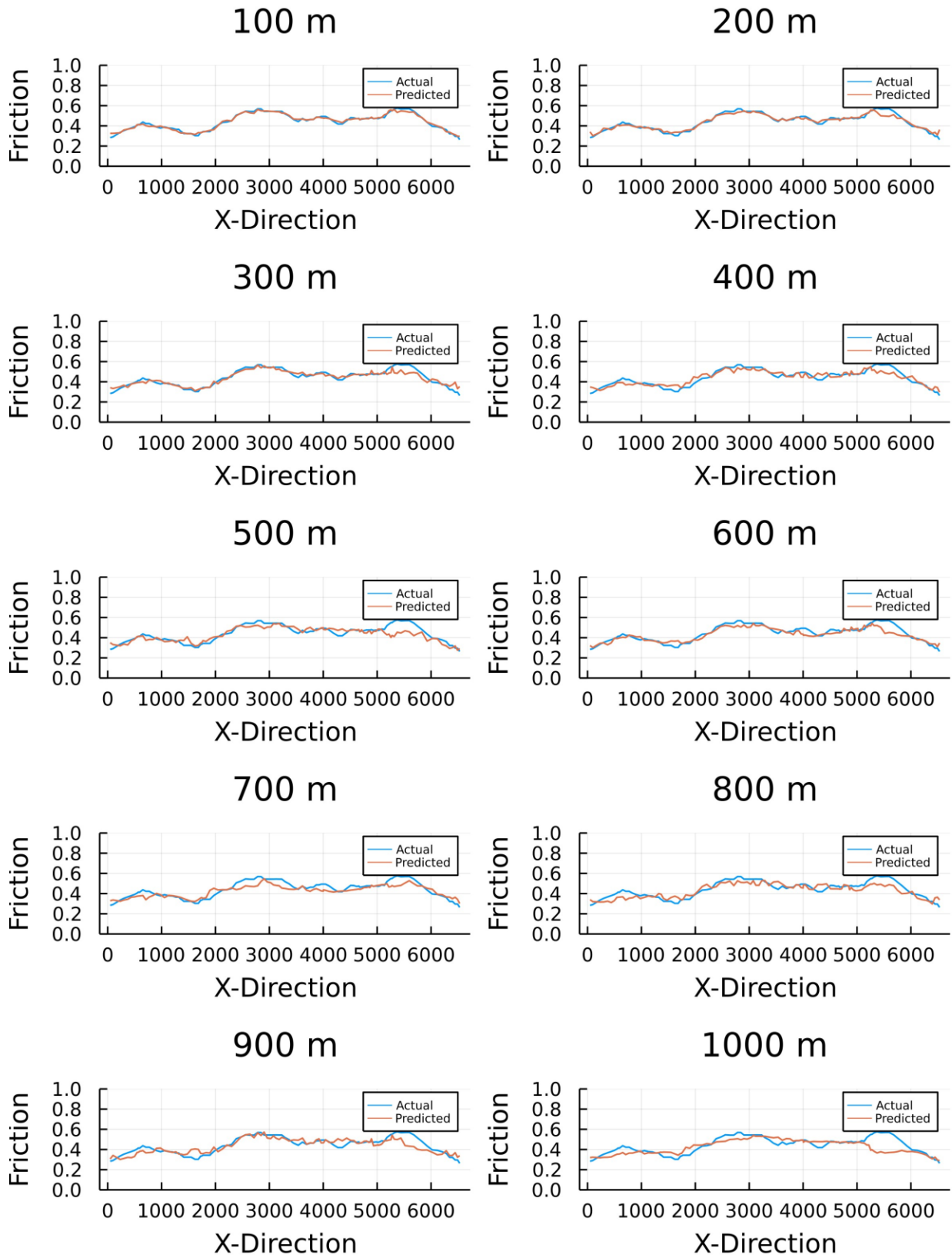
Figure A1. Ordinary Kriging (OK) Interpolation Curve



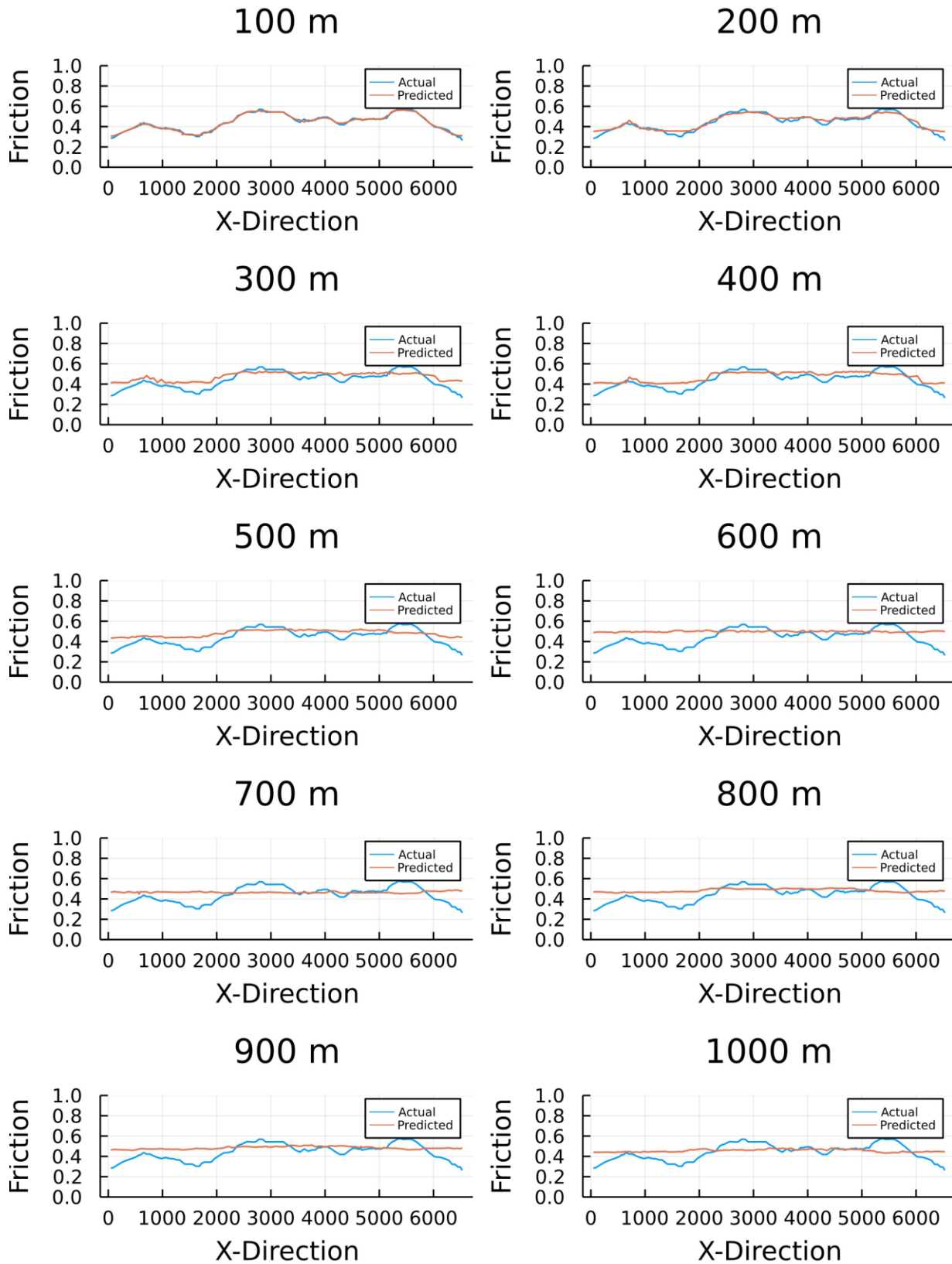
**Figure A2. Regression Kriging (RK) Interpolation Curve**



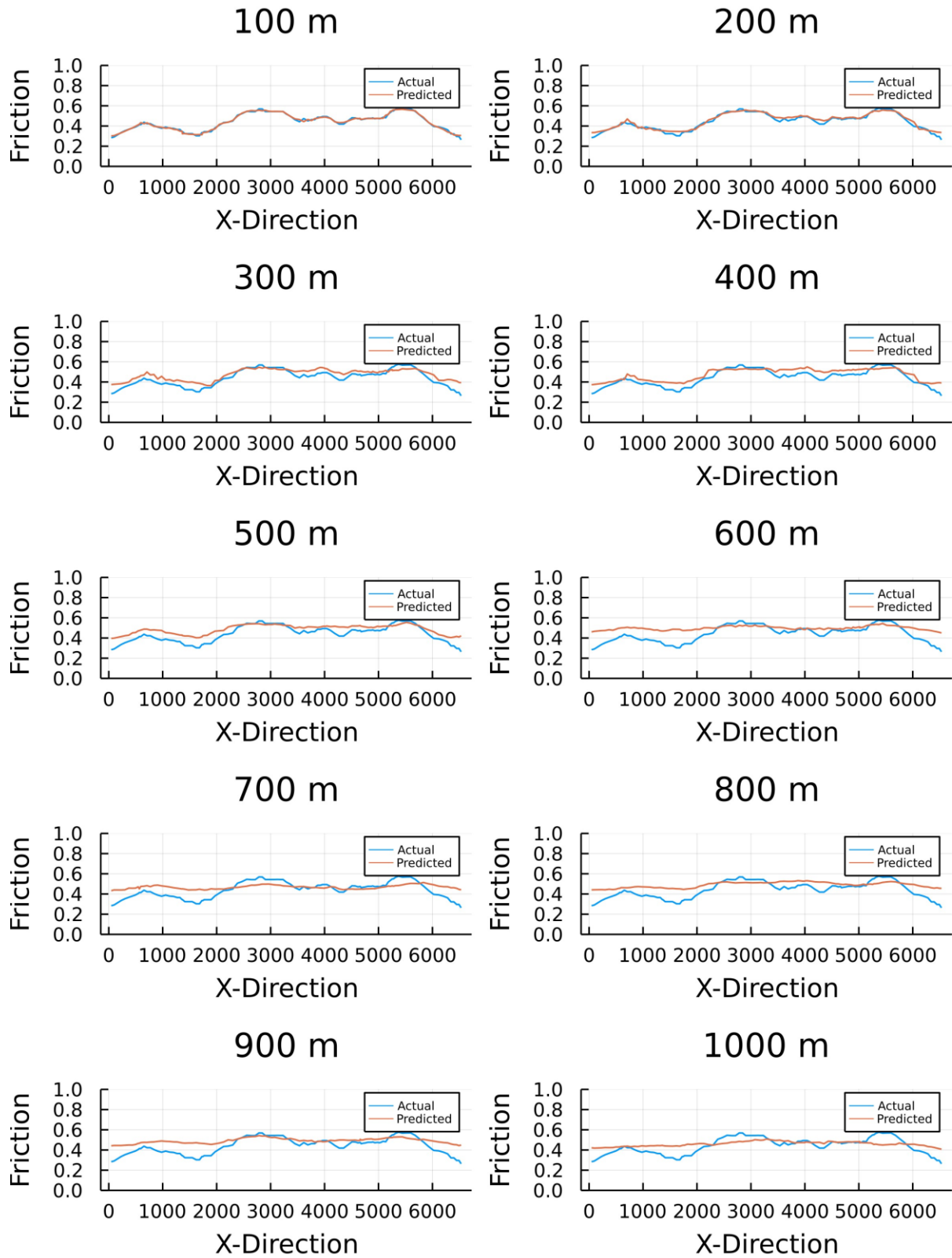
**Figure A3. Random Forest (RF) Interpolation Curve**



**Figure A4. RFOK Interpolation Curve**



**Figure A5. Random Forest Spatial Interpolator (RFSI) Interpolation Curve**



**Figure A6. RFSIOK Interpolation Curve**



**Table A1. Sensitivity Analysis Error Comparison**

<b>Observation Gap</b>	<b>OK</b>	<b>RK</b>	<b>RF</b>	<b>RFOK</b>	<b>RFSI</b>	<b>RFSIOK</b>
100	0.006068	0.006664	0.018416	0.013724	0.010684	0.010142
200	0.009511	0.010221	0.036411	0.022055	0.024523	0.019866
300	0.012689	0.013969	0.044642	0.03023	0.059214	0.04812
400	0.017556	0.01889	0.047665	0.034488	0.055107	0.048822
500	0.015638	0.021129	0.063017	0.047634	0.0703	0.059196
600	0.025845	0.023545	0.056933	0.035731	0.096406	0.082898
700	0.027872	0.028861	0.059824	0.043463	0.087096	0.073138
800	0.025905	0.030293	0.058074	0.044286	0.083844	0.075116
900	0.032585	0.032088	0.065228	0.043344	0.084657	0.075525
1000	0.05119	0.057832	0.070466	0.063669	0.07782	0.067425