A Smartphone-based Connected Vehicle Solution for Winter Road Surface Condition Monitoring

by

Michael André Linton

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Masters of Applied Science
in
Civil Engineering

Waterloo, Ontario, Canada, 2015

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The monitoring of winter road surface conditions (RSCs) is essential to transportation agencies and the traveling public, since the former needs to be aware of the location and severity of existing RSCs in order to effectively maintain safe roadways with minimal environmental impact, while the latter uses RSC information to make informed travel decisions. However, current RSC monitoring practice still relies on methods that are time-consuming, labour-intensive and lacking in objectivity, therefore limiting their ability to provide sufficient spatial and temporal coverage across a road network. This research was motivated by the need for accurate, timely and reliable RSC monitoring for winter maintenance personnel and the travelling public. To achieve this objective, the field performance of a smartphone-based RSC monitoring system was evaluated on a section of Highway 6 in Ontario, Canada during the winter of 2014. A comparison between this system and current monitoring methods indicated that the former was capable of providing reliable results particularly at the maintenance route level; however, classification accuracy was found to vary according to RSC type.

To improve the results produced by the smartphone-based system, this thesis proposes a connected-vehicle (CV) based RSC monitoring system that utilizes Road Weather Information System (RWIS) data in addition to the smartphone-based system's data. Three techniques in artificial neural networks (ANNs), random trees (RTs), and random forests (RFs) were tested as the underlying models of the CV system, and the results indicated that all three models successfully increased the classification accuracy of the smartphone-based system. RFs were found to provide the most accurate RSC classifications for the standard (three-class) classification scheme while RTs were found to be most accurate when using a more detailed (five-class) classification scheme. Model transferability was also tested using data captured from a different test site during the winter of 2015; and it was found that although the proposed CV system significantly increased the reliability of RSC classifications, the underlying models were non-transferable and would therefore require local calibration before being used at different sites across a road network.

Acknowledgements

Firstly, without a doubt, this thesis would not have been possible without the unwavering support of my supervisor Dr. Liping Fu. I would like to express my sincere gratitude for his encouragement, motivation and guidance throughout my Master's studies. His mentorship was truly invaluable.

Secondly, I would like to acknowledge the financial and operational assistance of the many public and private sector partners that I've had the pleasure of working with during my graduate studies: Ontario Ministry of Transportation (MTO), the Aurora Program, Transportation Association of Canada (TAC), CIMA+, Viaesys, Steed and Evans Ltd., Integrated Maintenance and Operations Services (IMOS), and the Town of Oakville.

Thirdly, I would like to thank my iTSS lab mates and colleagues for their many inspiring conversations and assistance. To Tae J. Kwon, Lalita Thakali, Matthew Muresan, Raqib Omer, Kamal Hossain, Zhengyang (John) Lu, Faranak Hosseini and Sajad Shiravi, thank you for your suggestions, expertise and stimulating discussions. We've found many solutions whether in the office, coffee shop or hallways. I would like to extend a special thank you to Ramona Mirtorabi for her constant words of encouragement, advice and inspiration.

Last but by no means least, I would like to thank my family and loved ones for their unconditional support, especially my parents and brother for their continuous words of guidance and encouragement - I truly appreciate it. Thank you to my Uncle Denis and Aunt Shannon for giving me the home away from home. Thank you Tyson for your constant and timely motivation and advice. Thank you Jabari, Ramon, Dario and Damian for the fun conversations. Thank you Aleshia for your support and motivation.

Dedication

To my parents: Ken Hamilton Linton & Ona Loretta Linton

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AMC	Area Maintenance Contractor
ANN	Artificial Neural Networks
BP	Bare Pavement
BPRT	Bare Pavement Regain Time
CART	Classification and Regression Trees
CCTV	Closed Circuit Television
CFME	Continuous Friction Measurement Equipment
CV	Connected Vehicle
DNN	Deep Neural Network
DOT	Department of Transportation
FS	Fully Snow Covered
ID3	Iterative Dichotomiser
LOS	Level of Service
MDCU	Mobile Data Collection Unit
MDSS	Maintenance Decision Support System
MLP	Multilayer Perceptron
MTO	Ontario Ministry of Transportation
PS	Partly Snow Covered
RCWIS	Road Condition and Weather Information Sheet
RF	Random Forest
RSC	Road Surface Condition
RT	Random Tree
RWIS	Road Weather Information System
SVM	Support Vector Machines
TAC	Transportation Association of Canada
TRIP	Traveller's Road Information Portal
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
WADT	Winter Average Daily Traffic
Weka	Waikato Environment for Knowledge Analysis

WOR	Winter Operation Records
WPR	Winter Patrol Records
WRM	Winter Road Maintenance

Chapter 1
Introduction

1.1 Background

For regions experiencing significant snow events, reliable and timely monitoring of road weather and surface conditions is essential to the success of any winter road maintenance program. Highway agencies and maintenance personnel need to be constantly aware of the location and severity of road conditions in order to maintain roadways at safe and acceptable levels of service (LOS). Additionally, motorists could use this information to make better travel decisions about routes, modes and even when to commence trips. Unreliable information on winter road conditions can therefore lead to inefficient maintenance decision-making, increasing costs and environmental impacts; similarly, uninformed driver decision-making, can increase congestion, travel time and accident risk.

At the onset of every winter storm, maintenance personnel must make decisions about fleet dispatching and chemical use to restore snow-covered roads to bare pavement as quickly as possible. Reliable information detailing the location and severity of road surface conditions (RSC) during and after a snowstorm therefore allows maintenance personnel to prioritize activities in order to achieve this goal efficiently. Maintenance agencies that are able to treat snow-covered roads more strategically are also able to improve road safety, reduce operating costs and better manage public expectation. Furthermore, highway agencies are often required to provide RSC information to the travelling public, who need the most recent and accurate data available to make travel decisions.

The most commonly used methods for monitoring winter roads are Road Weather Information Systems (RWIS) stations and manual patrolling. An RWIS station is a group of environmental sensors that collects real-time localized road and weather data such as precipitation type and intensity as well as air and pavement temperature. Typically fixed to the roadside, this technology provides information that is accessible remotely. Some stations are even equipped with cameras that capture images of the pavement surface, allowing for visual identification of the existing conditions (Buchanan & Gwartz, 2005). This technology, though valuable to the winter maintenance industry, is costly and lacks spatial coverage as a result of fixed installation (Kwon & Fu, 2013; Kwon et al., 2014; Jin et al., 2015). Manual patrolling involves personnel driving along a route and manually recording visually observed RSCs. This method addresses the previous issue of limited spatial coverage but lacks objectivity, repeatability and timeliness. As a result of its extensive spatial coverage, manual patrolling is an integral source of RSC information for highway agencies.

The laborious and subjective process of manual patrolling and the spatial sparseness of RWIS information have prompted the development of technologies aimed at addressing these issues while providing reliable and cost effective RSC monitoring. Technologies such as friction trailers and spectroscopic sensors, which have gained popularity as viable RSC monitoring tools, add information to the visual description offered by manual patrolling. For instance, friction trailers continuously measure friction or grip along the road surface and provide a quantifiable RSC measure; however, they cover only a single wheel track and are therefore limited in lateral spatial coverage across the roadway. Moreover, there remains a concern about inter-device calibration as well as standardization of friction measurements as an indicator of performance and safety measurement (Al-Qadi, et al., 2002; Erdogan et al., 2008). Spectroscopic sensors, using the reflection of light, estimate friction and provide information about pavement surface status; however, these devices can be costly and may offer performance advantages too limited to be considered a cost-effective RSC monitoring alternative.

In most North American jurisdictions, LOS is assessed by using a bare pavement standard, where a particular highway route is required to be restored to bare pavement status (or part thereof) within a specified time following a snow event. This standard means that the safety of a winter road and the quality of winter maintenance have to be assessed primarily through visual means, thus reinforcing the need to manually patrol the highway network. Vehicle-mounted video cameras that capture images or videos of the roadway have therefore been used more frequently to remotely monitor pavement status, and to provide video-based evidence if legal issues arise. Although this technology adds some measure of reliability to patrol reporting, manual interpretation is still required to translate the results of such recordings, resulting in a process that is as time consuming and subjective as the patrolling itself. As a result, several researchers have developed systems to automatically categorize the RSCs of a highway based on data acquired from vehicle-mounted camera systems (Conrad & Foedisch, 2003; Foedisch & Takeuchi, 2004; Hong et al., 2009). The main results were more reliable and less subjective automated RSC classification systems that provided visual-based evidence of road conditions for later retrieval and confirmation if necessary (Omer, 2011; Omer & Fu, 2010).

Traditional systems were not without drawbacks since RSCs were primarily classified offline via batch processing. Ideally, for practical decision-making purposes RSC information provided by the vehicle-mounted cameras should be real-time or near real-time. Moreover, maximizing spatial and temporal coverage requires that equipment be as inexpensive as possible for multiple vehicular installations in order to optimize highway network coverage. One of the proposed solutions for addressing these issues, a smartphone-based RSC monitoring system called AVL-Genius, provides real-time automated road condition information to the end user. Using smartphone cameras to capture images of the road surface at specified spatial intervals, this system automatically classifies each image according to level of snow coverage. These classifications are then visualized on a Google Maps interface so that the travelling public and maintenance operators can easily identify road conditions along highway routes. Originally designed to solve RSC monitoring problem by offering real-time RSC information on a ubiquitous and inexpensive smartphone platform, the system is currently used primarily as an experimental research tool (Linton & Fu, 2015; Fu & Linton, 2014). Prior to consideration for practical use by maintenance personnel, the system needs to be assessed for its accuracy, performance and reliability as a RSC monitoring tool. Moreover, it is hypothesized that initial system performance could be improved through the inclusion of localized weather information captured from RWIS stations.

Researchers have recently combined processing of road surface images with local weather data in order to produce more reliable RSC information; however, this procedure was performed using the camera images and weather data obtained from stationary RWIS stations (Jonsson P, 2011b). To acquire such RSC information with high spatial coverage across a highway network, this concept must be extended to a mobile system. With the objective of improved performance, this research proposes a smartphone-based connected vehicle RSC monitoring system in which localized road and weather data obtained from RWIS stations is combined with processed images of the roadway captured from a smartphone system. Connected vehicle technology allows vehicles to communicate with each other and surrounding transportation infrastructure in order to achieve a desired goal. One potential application of connected vehicles is road and weather monitoring; however, most proposed systems are conceptual and have not yet been tested for preliminary evaluation. This thesis research therefore explores the potential of a connected vehicle RSC monitoring solution.

1.2 Research Objectives

The objectives of this thesis are defined as follows:

1) Review existing literature on winter road surface condition (RSC) monitoring technologies and models used in image based RSC discrimination;
2) Evaluate a smartphone-based RSC monitoring system developed in a previous study;
3) Identify sources of errors in the smartphone-based system;
4) Propose a connected vehicle RSC monitoring system to improve the smartphone-based system RSC monitoring results and provide more detailed RSC classifications for use by the maintenance community

The scope of this thesis is limited to the evaluation of an image-based RSC smartphone monitoring system and its improvement by using proposed connected vehicle structure.

1.3 Thesis Organization

This thesis is divided into five main chapters and an additional section of supporting appendices. Each chapter is comprised of several subsections that systematically address research objectives. Figure 1.1 illustrates the framework of this thesis while the details of each chapter are summarized below.

- Chapter 1 provides a background on the research problem and outlines research objectives and the scope of this thesis
- Chapter 2 provides a detailed literature review of existing RSC monitoring methods and technologies and identifies the issues unaddressed in previous research.
- Chapter 3 describes the procedure and results of the evaluation of a smartphone-based automatic RSC monitoring system.
- Chapter 4 explains the hypothesis and testing of a connected vehicle RSC monitoring system.
- Chapter 5 summarizes the major conclusions, implications, and limitations of this research, suggesting recommendations for future study.

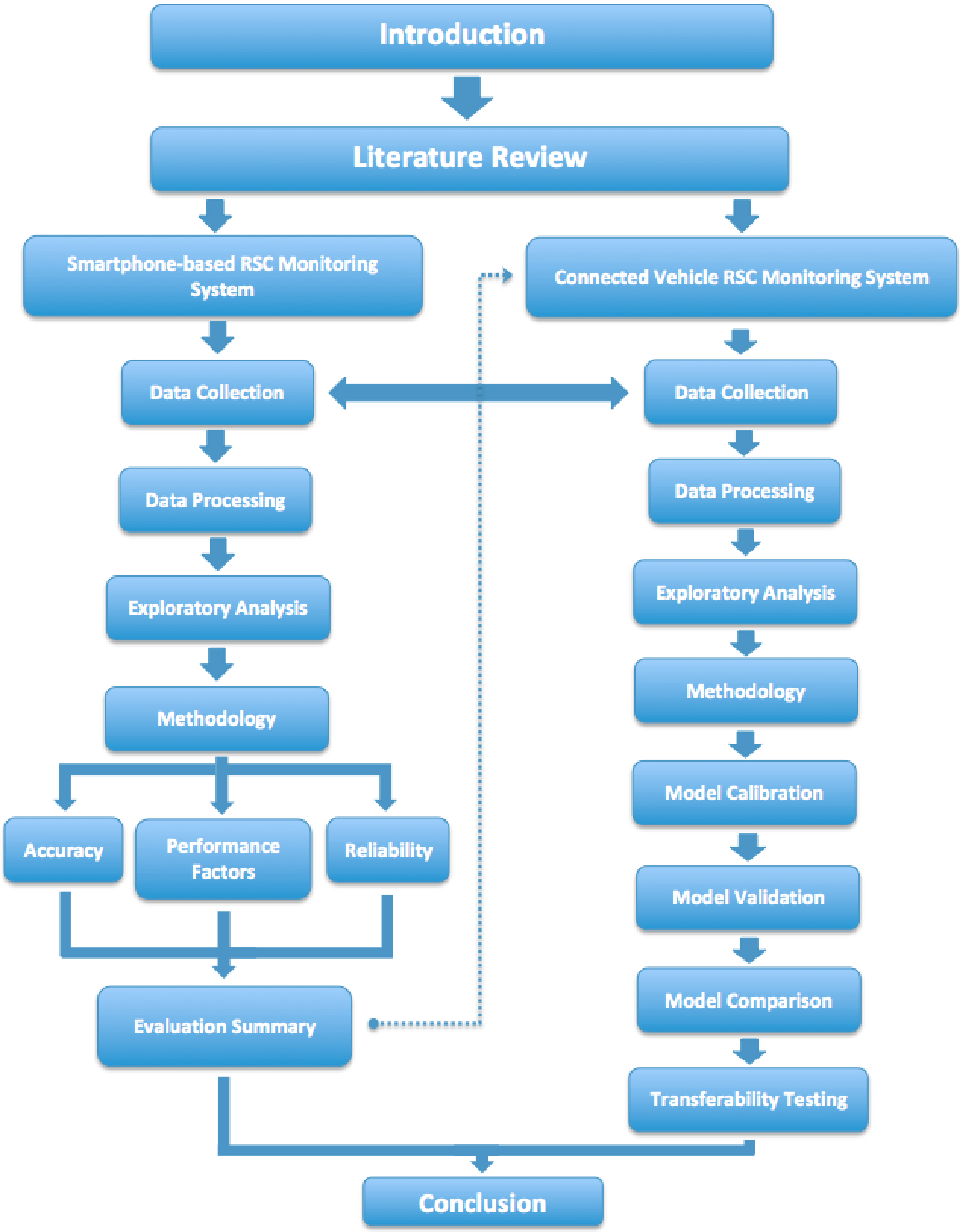


Figure 1.1 - Flowchart of Thesis Framework

Chapter 2
Literature Review

One of the primary objectives of this thesis is to evaluate a smartphone-based road surface condition monitoring system. In reviewing existing road surface condition (RSC) monitoring technologies and their usability as common winter maintenance tools, this chapter discusses the most frequently used systems, including those that include image recognition for identifying winter road surfaces.

2.1 Road Surface Condition Monitoring

Winter road maintenance (WRM) personnel are required to restore snow and ice covered roads to a safe state as defined by highway agency guidelines, within a specified time following winter weather events. The safe state is determined by the performance measure legislated by the jurisdiction, while the specified time is determined by highway class. WRM personnel therefore must know when to begin operations after the safe state has been compromised and when to cease operations once the safe state has been restored.

The most popular measure used today is visual observation, resulting in Bare Pavement (BP) policies being enforced by highway agencies across the world, while Friction-based policies have been utilized in some Nordic countries as an alternative or supplement to BP policies. The difference is that the former requires highways to be restored to bare pavement, whereas the latter requires highways to be restored to an acceptable friction level. However, both standards must be met within a certain time following a winter event. Regardless of the performance measure, highways need to be monitored to ensure that motorists are safe and that LOS standards are being met. The following sections detail the methods and technologies used in the RSC monitoring process, both in practice and in research.

2.2 Non-Visual Road Surface Condition Monitoring Technologies

2.2.1 Road Weather Information System (RWIS) Stations

An RWIS station consists of a group of environmental sensors that collects real-time localized weather and pavement condition data, such as air and pavement temperatures, type and intensity of precipitation, dew point, as well as surface contaminants, amount of deicing chemical on the roadway. The Ontario Ministry of Transport (MTO) currently has over 140 RWIS stations installed across the provincial highway network, thus providing an important decision support tool for winter road maintenance (Buchanan & Gwartz, 2005; Kwon & Fu, 2013).

Figure 2.1 - Typical RWIS Station with Environmental Sensors (North Dakota Department of Transportation, n.d.)

An RWIS station with basic functionality carries an installation cost of more than $50,000. The overall cost increases when additional in-pavement sensors are added and maintenance is included (Buchanan & Gwartz, 2005). However, even if advanced sensors are added to an existing RWIS station, measurements will have limited spatial coverage due to the technology being fixed to the roadside. On the other hand, being fixed to the roadside gives RWIS stations high temporal resolution, providing observers with a reliable assessment of the changing conditions at a particular site. Considering the costs of RWIS stations, it is not feasible to consistently install RWIS stations with high spatial density along the highway network. This lack of high spatial density creates an incomplete picture of roadway conditions along the network, since it is possible to obtain only spot-wise measurements scattered along different highway sections. With limited resources, selecting RWIS installation sites has become an optimization problem, since highway agencies are forced to balance technology costs with societal benefits. Researchers are currently investigating methods to address this issue, since RWIS stations play a vital role in the success of winter maintenance programs (Kwon & Fu, 2013; Kwon et al., 2014; Jin et al., 2015).

Areas susceptible to snowfall frequently experience microclimatic effects, where weather conditions in a localized area are very different to those of the surrounding areas. As a result of this effect as well as the variability in weather patterns, a single maintenance route may experience a multitude of conditions such as drifting snow, snowfall, sunlight and freezing rain. Similarly, several types of maintenance operations such as plowing, sanding and salting may also be performed on the same route. Since weather as well as varying traffic conditions all affect the RSCs to varying degrees, multiple RSCs often occur along a particular route. This possible significant variation in RSCs makes it questionable to use RWIS observations at a single spot to represent the conditions along an entire route. This is one of the main drawbacks of using RWIS data to report RSC, especially to the travelling public.

2.2.2 Thermal Mapping

Thermal mapping (TM) is a process of determining the spatial distribution of pavement temperature over a highway or highway network using temperature sensors such as infrared (IR) thermometers. Thermal surveys are usually carried out by a fleet of vehicles equipped with IR devices measuring the road surface temperature on winter nights under various weather conditions. Measurements are taken during the winter nights in order to avoid temperature errors due to the rising sun (The Institution of Civil Engineers, 2000; Marchetti et al., 2014). Thermal fingerprints (or maps) can subsequently be generated for different types of weather and climate conditions, as illustrated in Figure 2.2 (Erpicum et al., 2005). The resulting thermal maps allow visual identification of areas prone to freezing and other “cold-spots” in which specialized maintenance activities are necessary to keep the road safe. Thermal maps are also used to forecast pavement temperatures for maintenance decision-making. The effectiveness of anti-icing is highly dependent on pavement surface temperature, and thermal mapping is used to optimize anti-icing routes and to select locations for road weather outstations (Handa et al., 2007; Marchetti et al., 2014). Identification of these locations allows maintenance operators to adjust material selection and application rates accordingly, or notify road users via appropriate media.

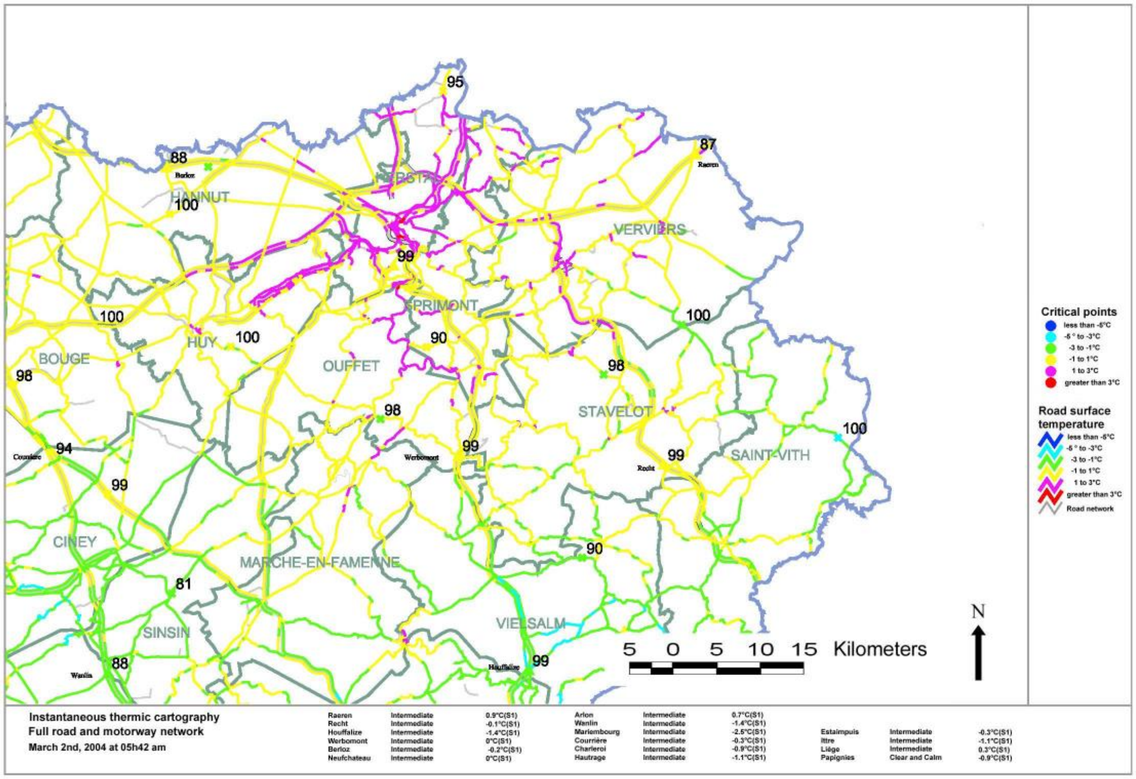


Figure 2.2 - Thermal Mapping Example (Erpicum et al, 2005)

While TM can capture spatial variation in pavement temperature across a highway network, a feature lacking in point measuring technologies such as RWIS stations, it is limited by low temporal resolution and incomplete representation of various winter events and conditions. For example, this technology does not provide any information on the state of snow and ice cover. Since the ideal time to take pavement temperature surveys is prior to dawn as well as during various weather conditions, completing a survey for an entire highway network can be time-consuming. Other issues can be prevalent during the measurement-taking process such as the location of hotspots at road junctions caused by stationary traffic and the lane changing of the measurement-taking vehicle due to presence of slow-moving vehicle on multilane highways (Marchetti et al., 2014). In previous studies, pavement temperature across and along multilane highways has even been found to vary by more than 2°C, a variation that can be critical for material selection and maintenance operation decisions (Chapman & Thornes, 2005). Therefore, while the technology can provide high spatial resolution in pavement temperature, it cannot provide all of the information about winter RSC that is required by maintenance operators and the travelling public.

2.2.3 Spectroscopic Sensors

Spectroscopic sensors represent the latest technology available for monitoring road surface conditions during winter events. Unlike the embedded pavement temperature sensors often used as part of an RWIS station, spectroscopic sensors work in a non-intrusive way by emitting light towards the road surface in one or several different wavelengths, usually in the near infrared (NIR) spectrum, and then receiving and analyzing the reflected light to estimate the status and amount of the contaminants on the surface spot being detected (Pilli-Sihvola et al., 2006; Riehm, 2012; Jonsson et al., 2015b; Casselgren, 2007). These sensors can estimate the surface contaminants because water, ice and snow have different spectral responses in the NIR region (Casselgren et al., 2007). Figure 2.3 shows the spectral responses obtained in a laboratory setting for varying RSC types in different wavelengths using a standard spectrometer. The distinguishable differences in surface contaminant absorption allows for reasonably accurate RSC detection according to the wavelength used.

Some spectroscopic devices can also provide additional information such as grip level, freezing point temperature, water film depth, or percentage of ice in water. However, similar to the embedded pavement sensors, stationary spectroscopic sensors are also restricted to the small spots being monitored, a limitation which is even more an issue for monitoring conditions of high spatial variation such as snow cover. Nevertheless, commercial spectroscopic sensors have gained popularity due to their non-intrusiveness and ability to measure multiple features of the pavement surface.

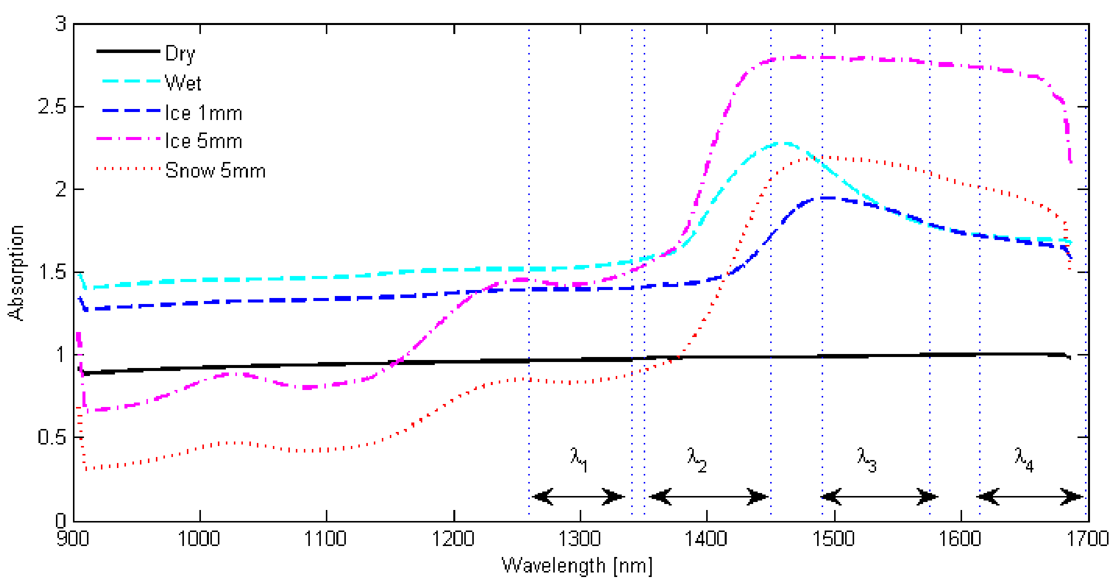


Figure 2.3 – Infrared Light Absorption by Different Surfaces (Jonsson et al., 2015a)

Popular devices include the Vaisala DSC111, Lufft NIRS31 and the Teconer RCM411 (Vaisala, n.d.; Lufft, n.d.; Teconer, n.d.). Research has assessed the performance of such spectroscopic sensors as compared to in-situ instruments, and Feng & Fu (2008) found that the spectroscopic sensors generally provide similar environemntal mesurements when compared to in-situ instruments. However, friction measurements can vary when comparing results to dedicated friction measuring equipment.

Figure 2.4 – DSC111 Non-Intrusive Laser Based Condition Sensor (Vaisala, n.d.)

Mobile variants of the sensors have been developed to address the issue of low spatial coverage that is associated with a fixed sensor, providing maintenance personnel with real-time information about the existing conditions along maintenance routes (Haavasoja et al., 2012). Mobile spectroscopic sensors can also to visually identify conditions by communicating results to map interfaces, allowing maintenance personnel to clearly identify unsafe locations. Figure 2.5 shows such an interface, which is one of the more valuable components of an overall mobile spectroscopic system. However, the relatively high cost of these mobile sensors may inevitably lead to low spatial and temporal coverage if the equipment is used only when dedicated patrol vehicles traverse highway sections.

Figure 2.5 – RCM411 Spectroscopic Data on Map (Teconer, n.d.)

Several transportation agencies, including MTO, have evaluated the field performance of these sensors as an add-on to the existing RWIS stations; while a few have also tested the sensor for mobile monitoring (Feng & Fu 2008; Joshi, 2002; Ye et al., 2012). Nevertheless, because of their high costs and limited performance advantages, significant feature improvements are needed before they can become a cost-effective RSC monitoring alternative.

2.2.4 Continuous Friction Measuring Equipment (CFME)

Continuous friction measurement equipment (CFME) measures the coefficient of friction or the friction number, of a pavement surface using specially designed tires attached to a device mounted on a travelling vehicle. CFME can collect spot-wise friction data along a maintenance route during winter events, thus having the potential to support maintenance decision-making and performance management. This type of high spatial resolution data, when made available in real-time or near real-time, allows maintenance operators and road users to make informed decisions in a timely manner (Perchanok, 1998; Al-Qadi, et al., 2002; Feng, 2013). Moreover, friction measurements represent a quantitative RSC measure instead of the traditional descriptive, qualitative measures. For example, friction measurements allow for identification of maintenance "hotspots" (slippery surfaces) in a road network indicating areas where greater attention may be needed, similar to those shown in Figure 2.5. Friction data with sufficient spatial and temporal coverage can also be used for performance measurement. Maintenance personnel can also use measurements to make analytical decisions that optimize operations while highway agencies can make objective assessments about RSCs. In practice, several Nordic countries have already used friction as a performance measurement tool for improved WRM decision-making (Cloutier & Donaldson, 2007). Agencies have also been experimenting with this technology for many years; however, it has mostly been used as a research tool to evaluate alternative snow and ice control methods and technologies (Fu et al., 2008; Feng & Fu, 2009 ; Feng et al. 2010; Takahashi et al., 2013; Al-Qadi, et al., 2002; Salimi et al., 2014; Cloutier & Donaldson, 2007).

Despite the advantages associated with the use of CFME to monitor RSCs, several issues exist with this technology. For example, uniquely mapping friction levels to road surface snow cover and type remains a challenge since friction data alone does not fully describe RSCs that may be observed in the real world, as is often required by both maintenance operators and travellers. This drawback is partially due to the fact that CFME covers only a small area, typically the wheel path of the cross section of a roadway, resulting in limited lateral representation. This misrepresentation of RSCs can become more apparent in multilane highways, where each lane may have varying RSCs. Researchers have developed different scales and standards to map RSC types to an equivalent friction value; however, the inconsistency of these scales among organizations can lead to discrepancies in RSC descriptions. Tables 2.1 and 2.2 demonstrate the overlapping characteristic of different friction standards.

Table 2.1 – Correspondence between Friction Values and Road Conditions (Finnish Road Administration, 2008)

Friction Value	**0.00 - 0.14**	**0.15 - 0.19**	**0.20 - 0.24**	**0.25 - 0.29**	**0.30 - 0.44**	**0.45 - 1.00**
Description of the road surface	wet ice, **very slippery**	icy **slippery**	smooth compacted snow **fair winter road conditions**	antiskid compacted snow and ice **good winter road conditions**	clear and wet **antiskid road conditions**	clear and dry **antiskid road conditions**

Table 2.2 – Friction Ranges of RSC Types (Öberg & Gregersen, 1991)

RSC Type	**Friction Value**
bare dry	0.8-1.0
bare wet	0.7-0.8
packed snow	0.20-0.30
loose snow/slush	0.20-0.50
black ice	0.15-0.30
loose snow on black ice	0.15-0.25
wet black ice	0.05-0.10

Feng (2013) identified several issues with such RSC classification schemes. For example, particular friction values may represent different RSCs depending on the study or country in which they are used. These friction values may also overlap, depending on jurisdiction, making it difficult for maintenance personnel to identify the nature of surface contaminants. In addition, the fact that some schemes define RSC types with a wide range of friction values indicates large tolerances for RSC discrimination that can compromise the reliability of that particular friction-based classification scheme. Moreover, different friction devices also have varying degrees of accuracy and reliability. Halliday RT3 (Halliday Technologies Inc, n.d.) is one of the popular devices used for testing in highway agencies across North America, but other devices used in the market include the Traction Watcher One (TWO) and the Griptester (GripTester, n.d.). Although it is conceptually possible to inter-calibrate these devices in order to unify classification schemes and CFME, actually doing so is not feasible.

Figure 2.6 - An Example of CFME (Halliday Technologies Inc., n.d.)

CFMEs are also costly and require a significant amount of work for installation and calibration, making their cost-effectiveness for application as a network-wide monitoring tool questionable. This issue when coupled with the problems in availability and reliability of friction data can compromise its value to highway agencies.

2.3 Visual Road Surface Monitoring Technologies

2.3.1 Patrol Reporting

Patrolling the road network and reporting its prevailing road weather and surface conditions represent the state-of-the-practice method for collecting RSC data used by most maintenance agencies. Patrollers travel along designated routes and record their conditions on a patrol report, describing the bare pavement status, the extent and types of surface contaminants, and active maintenance operations being deployed.

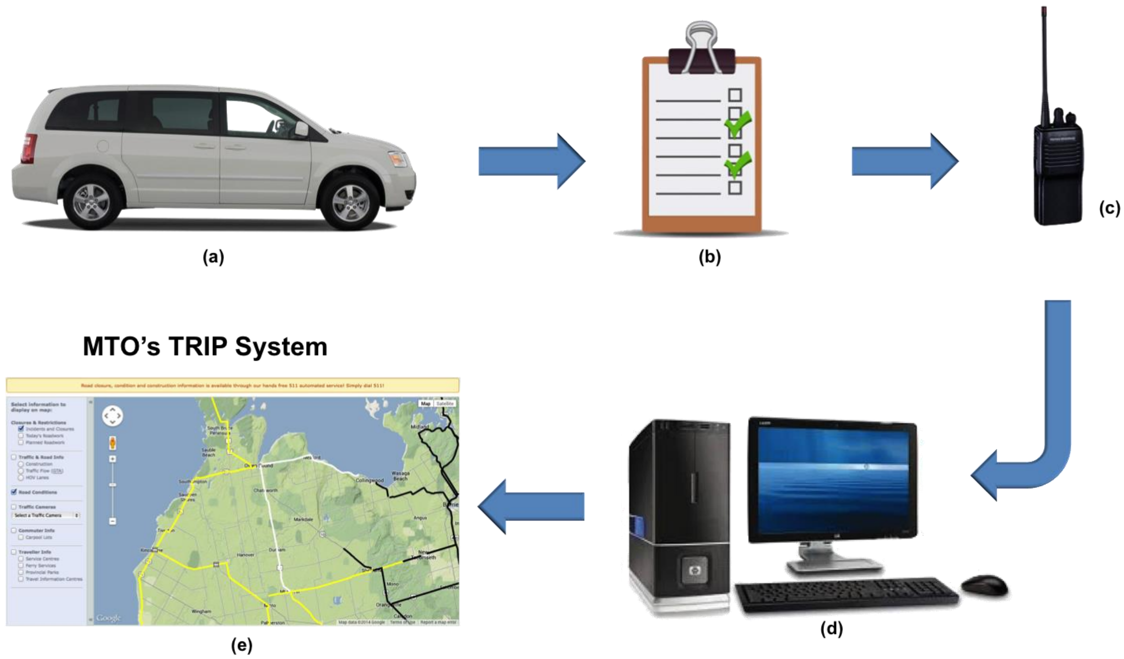
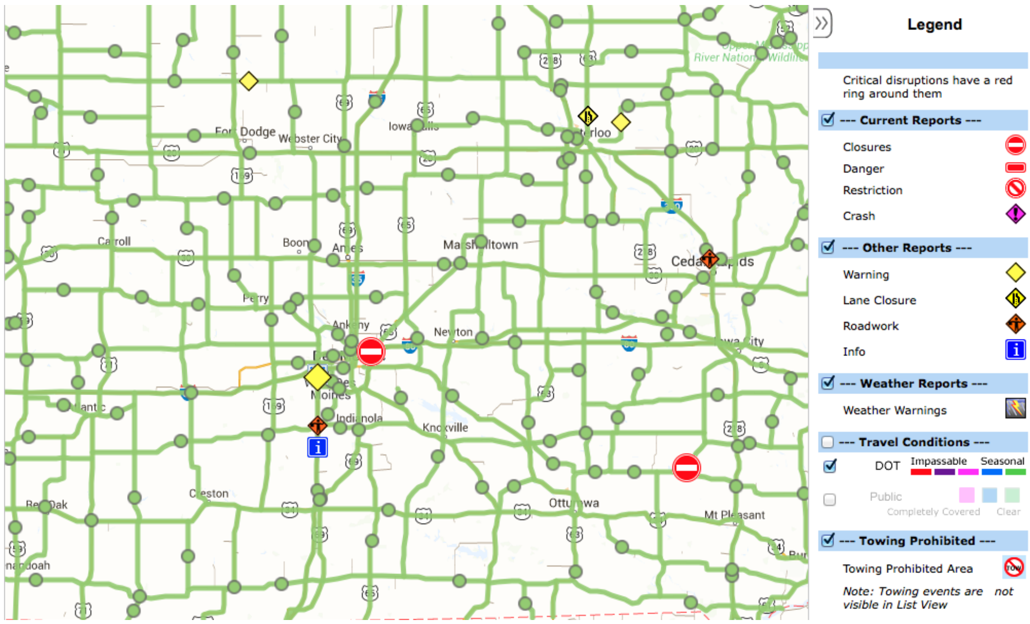


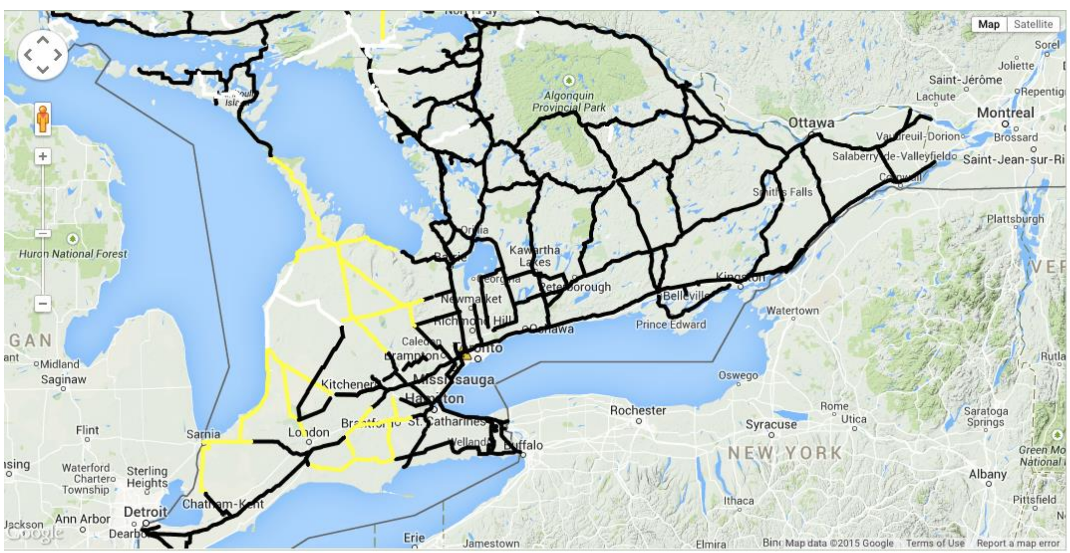
Figure 2.7 – The Current Process of RSC Reporting in Ontario

Figure 2.7 shows the steps involved in a typical reporting procedure, from surveying the maintenance route to recording and reporting, to publication on the Ministry's Travellers Road Information Portal (TRIP) for the travelling public. MTO currently adopts a self-monitoring approach in which Area Maintenance Contractors (AMCs) are responsible for patrolling their maintenance routes and reporting the conditions during a winter season, with the number of daily observations dependent on the weather conditions and RSCs experienced. The Ministry also sends out its own personnel to check the road conditions on a random basis as a form of contract oversight to ensure the accuracy of patrol reports submitted by AMCs. Additional patrols are also conducted at least 5 times daily, the results of which become available to the public at designated times through a visual online interface (TRIP). This process is similarly adopted by Departments of Transportation (DOTs) across North America, although some opt not to communicate results to the public (Iowa Department of Transportation, n.d.; Ontario Ministry of Transport, n.d.). However, those that do still engage in a manual observation and notification procedure. As a manual and labour-intensive process, patrol reporting has the drawbacks of low efficiency, high subjectivity and low granularity. Figure 2.8

shows examples of interfaces used to communicate RSC information to the public. Details of the patrol monitoring process and the types of data collected are described in the following Chapter.



(a)



(b)

Figure 2.8 – Web-based Traveller Information Services: (a) Iowa Department of Transportation Traveler Information (Iowa Department of Transportation, n.d.) (b) MTO's Traveller's Road Information Portal (Ontario Ministry of Transport, n.d.)

2.3.2 Web Based Surveillance Video

Road surface conditions can also be monitored remotely using a video based system that transfers video or images of the road surface in real time to maintenance personnel and road users via the Internet. For decades, Closed-Circuit Televisions (CCTVs) and web cameras have been used to remotely monitor highway conditions as a part of maintenance decision-making process and more recently public reporting. These media are frequently components of a more integrated system, such as found in RWIS stations, and provide accompanying data to highway agencies. Not all RWIS stations are equipped with functional video cameras, however, so it remains a logistical challenge to install fixed video cameras along the roadside for the sole purpose of RSC monitoring due to the infrastructure required to power the device and transmit data in real time.

CCTV cameras provide a snapshot of the RSC for the sections in view and can archive the images for retrieval in the future. Since images are restricted to the viewable section of roadway, there is a limited spatial coverage for RSC estimation along a highway route. Figure 2.9 shows an example of real-time web-based surveillance that allows remote RSC observation. Physical exposure to precipitation also means that captured images may be obscured by low visibility and dried precipitation on the camera lens. One of the biggest issues with using CCTV cameras is that the images require manual observation and classification. Judgment of observed images depends on the experience of the observer, making the process subjective and time-consuming. Manually analyzing images from a network of cameras may then require considerable human resources if the results of these images are to be used for WRM decision-making and public reporting (Ye et al., 2012; Yamamoto et al., 2005).

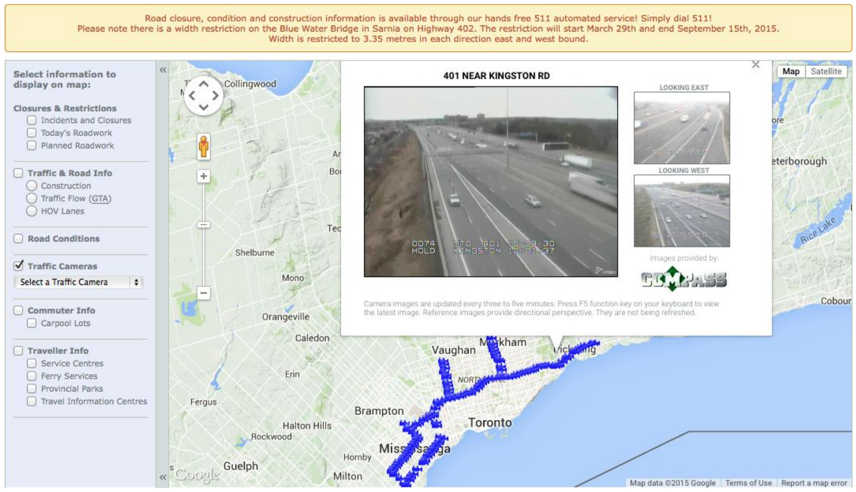


Figure 2.9 - MTO's Traveller's Road Information Portal – Traffic Cameras

2.3.3 Automatic Road Surface Condition Image Recognition

Since the 1990s when researchers began to understand the challenge of manually monitoring CCTV images, investigations into automatically classifying RSCs from images have been ongoing. The main concept of RSC identification through image recognition is centered on the development of algorithms that extract features from each image in order to automatically distinguish between RSC types. Figure 2.10 illustrates a recent automatic road condition imaging system based on a stationary camera. Over time, different image features have been found important to the image recognition process, which is also dependent on the RSC classification scheme used and the type of surface contaminant (snow, ice, water). Kuehnle & Burghout (1998) used neural networks to analyze image features and classify images of winter roads with moderate accuracies up to 50%. Since then, neural networks have been one of the most popular methods for winter RSC image classification as well as a comparison tool to measure the effectiveness of alternate image analysis methods. More advanced neural network models were later found to classify road images into specific categories such as bare, wet, icy, snowy and bare wheel tracks with accuracies of over 90%. These models, though initially developed for images captured by stationary camera systems, later evolved to facilitate automatic RSC classification from images captured by mobile in-vehicle camera systems (Conrad & Foedisch, 2003; Foedisch & Takeuchi, 2004; Hong et al., 2009; Jonsson P, 2011a; Omer, 2011; Omer & Fu, 2010).

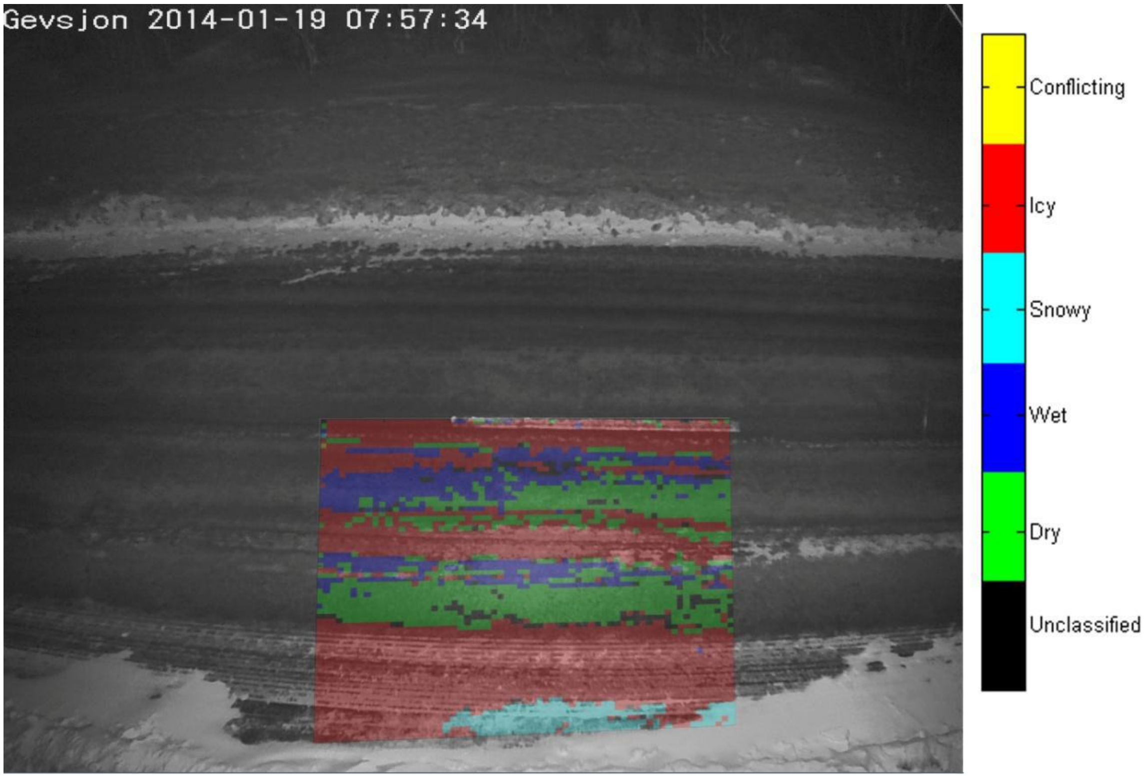


Figure 2.10 – Road Condition Imaging System (Jonsson et al., 2015b)

The relative maturity of artificial neural networks (ANN) development and overall performance for RSC classification via image recognition have encouraged the investigation of further model development. One popular alternative to traditional ANN models for RSC image classification is Support Vector Machines (SVMs). Conrad & Foedisch (2003) used SVMs were able to increase classification accuracy compared to traditional ANN models, but calculation times also increased. Specifically, SVMs took several seconds to classify an image while ANN completed the task in less than one second. While accuracy is the primary performance factor for RSC classification, the development of a real-time automatic system for this purpose would also require that computation time be considered in any decision on its practical use, especially when hundreds of images are waiting to be classified. Omer & Fu (2010) also successfully trained SVMs to classify images obtained from an inexpensive mobile video camera attached to the inside of a patrol vehicle, into three categories (bare, tracks and fully snow covered) with accuracies of 80-90%. The use of inexpensive video cameras for automatic RSC monitoring therefore allows for increased spatial and temporal resolution by facilitating multiple device use along a highway network, considering a budget equivalent to that for CCTV video cameras.

Even though some recognition models have correctly classified road images over with 90% accuracy, some instances may be situational and a bias may exist within the training and testing sets Omer & Fu (2010). Researchers are constantly trying to improve classification methods and model performance as well as testing such models on completely new data sets. Most attempts to improve RSC detection with road images have involved re-evaluating image classification algorithms and models (Jonsson, 2011a; Jonsson, 2011b). This method, while optimizing the information deduced from the image, does not account for situations in which intuitive weather data could provide useful classification information. For instance, if snowfall does not occur for 48 hours prior to image capture and the temperature is above freezing, it is unlikely for the road surface to be completely snow covered. Jonsson (2011b) created a model that simultaneously included image features and RWIS data to discriminate between conditions such as dry, wet, icy and the presence of wheel tracks. This study found that the combined method of RSC identification was more reliable than image classification alone. Along with features extracted from the RWIS camera images, the model included several important variables used to describe RSCs such as:

- Air temperature
- Air humidity
- Air dew point
- Precipitation count, number of particles
- Surface temperature
- Wind speed
- Wind speed as last 10 minute average
- Wind speed as last 30 minutes average
- Wind speed as maximum value of last 30 minutes
- Wind direction
- Wind direction as last 10 minute average
- Day/Night Indicator

In finding that most meteorological variables, along with the camera image features contributed to defining RSCs, this work confirmed the intuitive assumption that weather data could play an instrumental role in RSC discrimination when combined with image processing. Although the RSC discrimination in this work produced models with very high accuracies, the automatic RSC identification would be valid only for a fixed location where the RWIS station was installed, resulting in limited spatial coverage along a highway network. This drawback means that automatic RSC classifications would be limited according to the density of a highway agency's RWIS network. The success of such a model at a fixed location indicates that RWIS data and image classifications can identify RSCs using images captured by a mobile device throughout a highway network, thus increasing reliability and spatial resolution of RSC classifications.

2.3.4 Smartphone-based Automated Road Surface Condition Monitoring System

The smartphone-based automatic RSC monitoring system tested in this research, AVL-Genius, includes a front-end device for collecting RSC data and a cloud based server for data processing and reporting. The data collection data device consists of an Android smartphone with a dedicated App and an optional hardware integration that interfaces with other sensors such as an infrared pavement thermometer, salt rate controller and GPS, as shown in Figure 2.11. Once started, the smartphone takes pictures of the roadway at configurable intervals. The images can be uploaded to the cloud server either in real time via wireless cellular data connection or off-line at any Wi-Fi spots. The uploaded images are processed and classified in terms of snow coverage using an automated image recognition algorithm similar to those described in the previous section. The RSC classification results are then displayed in a standard colour scheme on a Google Maps interface. The device operates with little human intervention, and the customizable frequency of taking images offers flexibility in spatial resolution for the kinds of information needed by WRM operators and the travelling public.

Figure 2.11 - AVL-Genius Device - Smartphone and Control Box

The device takes as little as a few seconds to automatically classify captured images of a road surface, with a classification as either bare, partly snow covered, or fully snow covered. This system can also provide additional characteristics of the road surface such as percentage of snow cover, and quality control measures to indicate if a roadway is detected in an image, Each image is GPS-tagged and time-stamped, facilitating both aggregated and disaggregated views of RSC for a particular route. For example, it can provide detailed visualization of snow and ice covered hotspots along a route accompanied by images as illustrated in Figure 2.12. It can generate route level statistics for classification of the overall condition of a route (e.g., bare pavement regain status). If there is sufficient temporal coverage of a maintenance route, it is also feasible to derive critical performance information such as bare pavement regain time (BPRT).

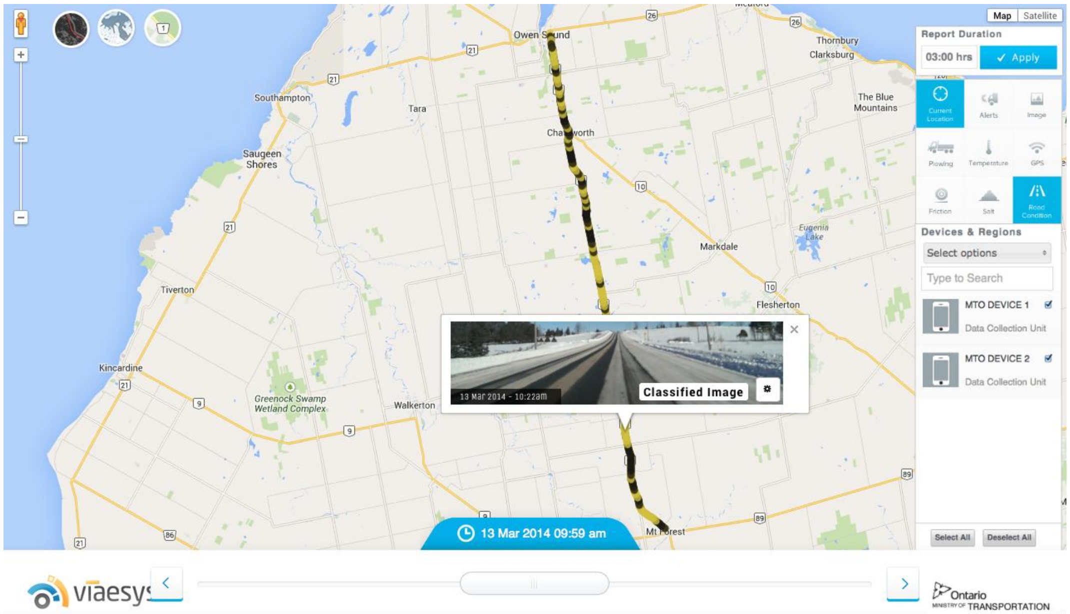


Figure 2.12 - AVL-Genius Web Interface Showing Classified RSC and Images

2.4 Summary

In this chapter, the major road surface condition (RSC) monitoring methods and technologies have been discussed, including the device tested in this research. Developments in standards and technologies were also identified as well as their shortcomings.

Previous studies have highlighted the various ways that practitioners and researchers can effectively monitor RSCs. One of the most widely accepted metrics is visual identification of RSCs. As a result, manual patrolling is still one of the most popular methods of RSC monitoring even though it is process that is not only time-consuming and labour-intensive but also somewhat subjective. Thus, saving time and increasing accuracy have led to the development of image recognition models that classify camera images of pavement surfaces to identify RSCs. The final performance of these models varies according to prevailing conditions, and researchers have consistently tried to improve this automatic classification process. Unfortunately, the focus on RSC classification improvement has been limited to improvement of image classification alone. One previous study successfully integrated RWIS data with camera image processing for RSC identification, but this was limited to fixed roadside RWIS cameras, resulting in a system that is still lacking spatially coverage across a road network. The next chapter introduces the test site and details the field evaluation of the smartphone-based RSC monitoring system tested in this thesis.

Chapter 3

Evaluation of a Smartphone-based Winter Road Surface Condition Monitoring System

Highway agencies and contractors face the consistent challenge of obtaining timely, reliable, accurate RSC information using inexpensive means, and of providing the public with this information so that they can make informed travel decisions. For any RSC monitoring system to be of practical use, one must assess its accuracy and its ability to provide similar or additional information to maintenance personnel and the public. In this chapter, a real-time smartphone-based RSC monitoring system called AVL-Genius is evaluated to determine its reliability as a monitoring tool. Of primary importance is system accuracy (How correct are the system's results?); and of equal significance is its comparability with current monitoring methods (How do the system's results compare with current RSC monitoring methods). In order to evaluate the performance of the system with respect to these issues, a field test was conducted during the winter season of 2013-2014 from Feb 24, 2014 to April 2, 2014. This chapter details the test site, data collection method, processing procedures, and results of the system evaluation.

3.1 Data Collection

3.1.1 Test Site

Field Tests were carried out in the winter season of 2013-2014 on a section of a two-lane, two-way Class 2 highway – Highway 6 near Owen Sound, Ontario as shown in Figure 3.1. The test section is approximately 70km long with a winter average daily traffic (WADT) volume of 4900 (Ontario Ministry of Transport, 2010). The site has uniform geometrical features, few horizontal curves and a combination of open and sheltered agricultural fields and woodlots. The area experiences an annual average of 59 days of snowfalls with at least 0.2cm (Environment Canada, 2014). An Area Maintenance Contractor (AMC) maintains the route with typical WRM activities, including plowing, sanding and salting.

The overall test area is covered by six (6) RWIS stations (thumbnails in Figure 3.1), two of which are located on the test route (SW-25 and SW-13, identified by the green thumbnails in Figure 3.1). These RWIS stations provide additional data on road and weather conditions around the test site.

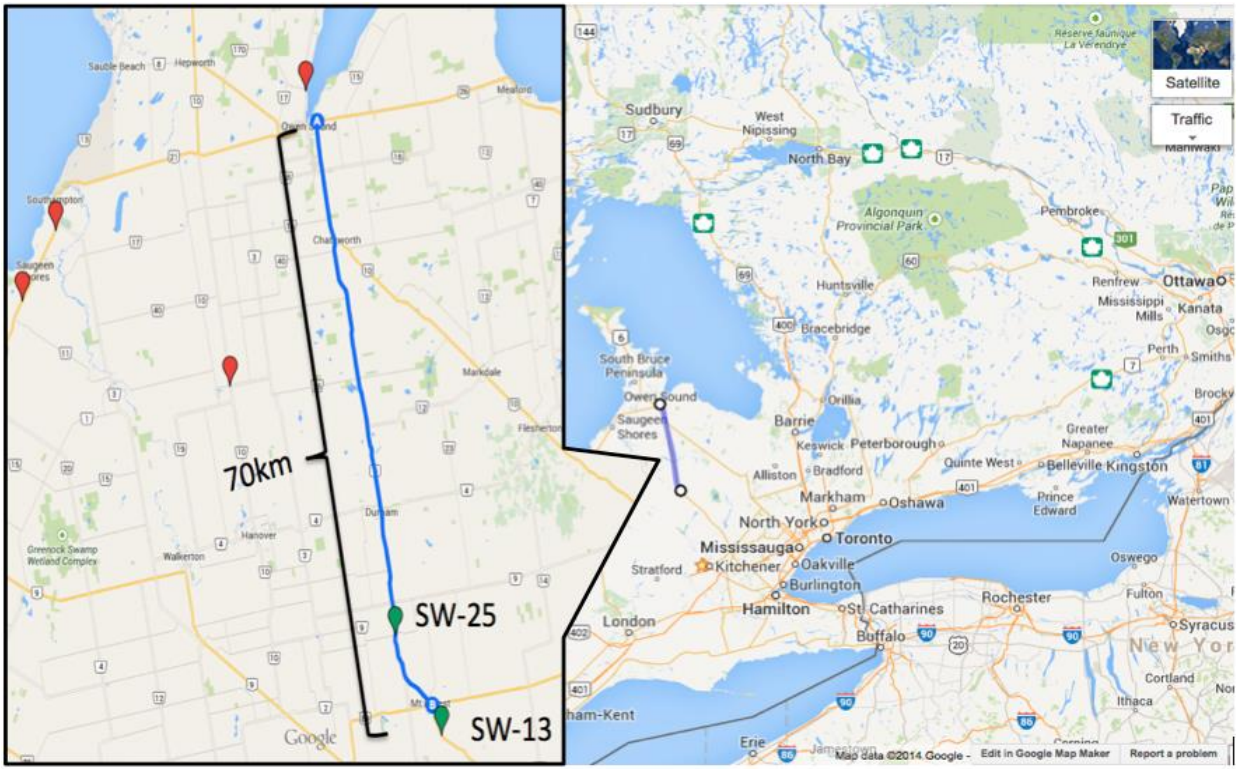


Figure 3.1 – Test Route and RWIS Stations

3.2 Data Sources

3.2.1 Smartphone-based System

For the purpose of this research, the smartphone-based system was installed on four patrol vehicles, two each from MTO and the contractors, and one dedicated mobile data collection unit (MDCU).

Figure 3.2 – Dedicated Mobile Data Collection Unit

Conducting their usual observational trips during and after snowstorms along the test route, patrollers manually recorded RSC observations on paper forms. Once turned on at the start of each trip, the AVL-Genius unit operates automatically, recording images at a spatial frequency of 450m and 350m for patrol vehicles and the MDCU, respectively. GPS-tagged and time-stamped images of the roadway were then automatically uploaded to a cloud server and classified by image processing software. As previously discussed, AVL-Genius classifies images into one of three distinct types: bare, partly snow covered, fully snow covered. Following the Transportation Association of Canada's (TAC) guidelines, this classification scheme is currently adopted by MTO and many other transportation agencies in Canada (Transportation Association of Canada, 2011).

3.2.2 Patrol Records

For each maintenance route, field operators prepare three types of winter road maintenance reports:

- Winter Patrol Records (WPR): Prepared by MTO and AMC staff, these reports include information such as weather conditions (precipitation and highway conditions (RSC)) and maintenance operations observed.
- Winter Operations Records (WOR): Prepared by AMC staff, this report includes information on maintenance operations performed as well as material type and amount.
- Road Condition and Weather Information Sheets (RCWIS): These reports, prepared by MTO patrollers five times a day as indicated in the literature review, contain information on precipitation, atmospheric and road conditions are included, as well as maintenance operations observed.

The details of the data collection process for the patrol reports are described below. The patroller drives along the maintenance route and records the observed RSCs on a patrol form. Fixed categories of RSCs are included on the patrol form as checkboxes to be filled out by the patroller. There is no provision to indicate the frequency of observation of a particular road condition. In addition to the RSCs, information on the type of maintenance operations observed and the resulting weather conditions is recorded on each patrol record. The possible conditions according to the forms are described in Table 3.1.

Table 3.1 – Data Provided by Patrol Records

Winter Patrol Records (WPR)			
Temperature (°C)	Weather Conditions	Highway Conditions	Operations
Air	Clear	Bare and Dry	Patrolling
Pavement	Partly Cloudy	Bare and Wet	Plowing
	Overcast	Track Bare	Sanding
	Rain	Centre Bare	Salting
	Snow	Snow Covered	Snow Blowing
	Freezing Rain	Snow Packed	
	Fog	Drifted Sections	
	Visibility	Icy Sections	
	Wind	Frost	
	Wind Direction	Slushy	
Road Condition and Weather Information Sheet (RCWIS)			
Precipitation Conditions	Atmospheric Conditions	Road Conditions	Maintenance Operations
No Precipitation	Air Temperature	Bare and Dry	Patrolling
Rain	Wind Direction	Bare and Wet	Plowing
Freezing Rain	Wind Speed	Partly Snow Covered	Sanding
Snow	Visibility	Snow Covered	Salting
Ice Pellets	Cloud Condition	Partly Snow Packed	Anti-icing
Mix (Rain/ Snow)	Fog	Snow Packed	Clean Up
	Drifting	Partly Ice Covered	
		Ice Covered	

Both the maintenance contractor and the MTO patroller follow these steps (parts (a) and (b) shown in Figure 2.7). These types of patrols are usually carried out for a few hours to allow MTO personnel to monitor contractor performance, while AMCs update winter patrol records as often as deemed necessary according to the prevailing conditions. During some intense storms, it is not uncommon for a single contractor to continue for over 8 hours of observation on a maintenance route.

After a patroller traverses a route and records observations on the patrol form, the results are radioed to the central location responsible for the maintenance area. This information is then manually entered into the system where it later becomes available to the public on MTO's TRIP website (now part of Ontario 511). The reported RSCs adhere to the Transportation Association of Canada's guidelines; graphically displayed on the TRIP website, they are color coded to represent the intensity of the reported conditions as follows (Transportation Association of Canada, 2011):

- Bare (Black)
- Partly snow covered (Yellow)
- Fully snow covered (White)

Road and weather information sheets provide atmospheric, weather and precipitation conditions in addition to RSCs according to the categories outlined in Table 3.1. These conditions are reported to the public five times daily through the TRIP system at 03:00, 09:00, 13:00, 15:00 and 21:00. The resulting visualization of RSCs on the TRIP system is dependent on the order in which conditions are entered into the system, an example of which is shown in Table 3.2. For instance, a patrol form may indicate partly snow packed and snow packed conditions. If "partly snow packed" is entered first into the system, indicating primary condition, followed by "snow packed", the resulting RSC would be "partly snow packed with snow packed sections", displayed in yellow on the TRIP website. Alternately, if "snow packed" is entered first, indicating its primary condition, the resulting RSC would be "snow packed with partly snow packed sections", translating to a white color on the TRIP website**.**

Table 3.2 – Example of Entry Order Effect on TRIP System Output

Scenario 1		
Order of Entry	**Road Surface Condition**	**TRIP Output**
1	Partly Snow Packed	
2	Snow Packed	Partly Snow Covered
Scenario 2		
Order of Entry	**Road Surface Condition**	**TRIP Output**
1	Snow Packed	
2	Partly Snow Packed	Snow Covered

According to the forms, there is no indication of what the primary RSCs are; the only distinguisher is the visualization offered by TRIP, as identified via radio by the patroller along a route. This means that patrol forms often show a myriad of RSCs observed along the route without any indication of dominant conditions. In the absence of an order specified by the patroller, it is common practice to lead with the more intense condition, i.e., Scenario 2 in Table 3.2. While safe from an accountability perspective, this can lead to an exaggeration of current RSCs, which when reported to road users can affect their trip decision-making. Additionally, updates to the website are made only five times per day, so displayed RSCs can be hours old at the time of accessing the TRIP website. For instance, TRIP data used to plan a morning commute at 8:00am are the conditions observed prior to 3:00am, since this is time when the most recent update is obtained. Several hours could make a big difference in terms of road surface conditions, which can vary between a fully snow covered highway to a bare highway section if weather changes and/or maintenance has been performed.

All manual patrol reports were processed and entered into a database via Microsoft Access forms. Field interviews were also conducted with maintenance personnel to further understand not only the process in which the winter road surface conditions were observed and made available to the public but also any accompanying issues involved in the process. Samples of each type of patrol report are included in Appendix A.

3.2.3 Traveler Road Information Portal (TRIP) Data

TRIP data was made available through MTO's web-based interface, which is updated five times daily: 03:00, 09:00, 13:00, 15:00 and 21:00. Screenshots were taken after the RSCs were updated throughout the day, as illustrated in Figure 2.8b.

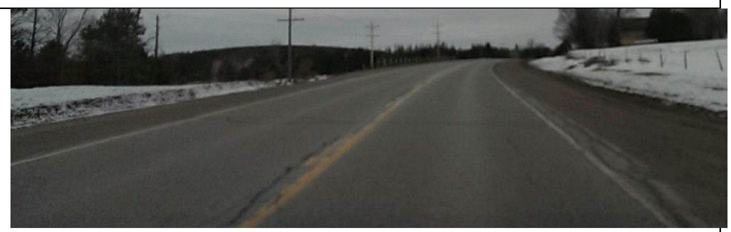
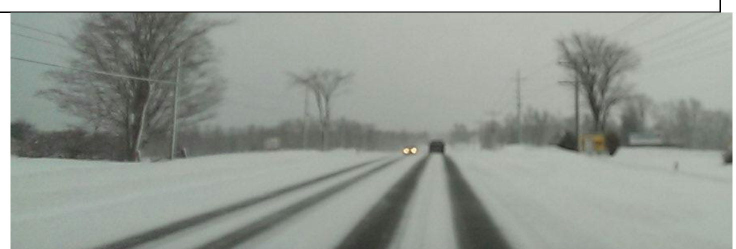
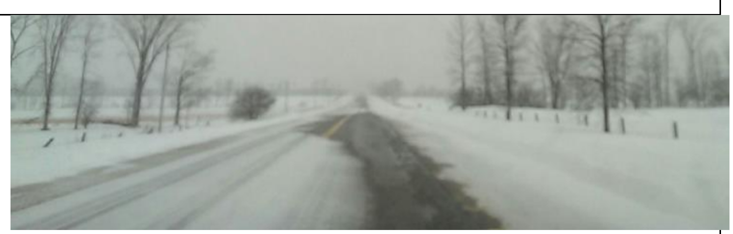
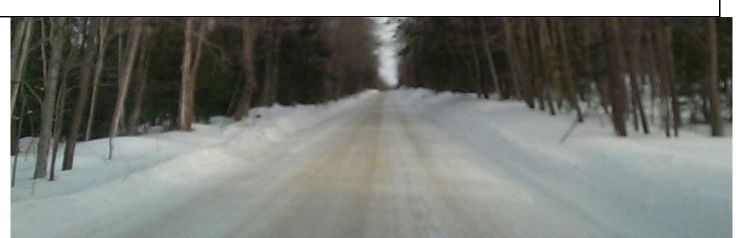
3.3 Data Processing and Preparation

Data was processed according to the requirements involved in answering the research questions. Details on the method are explained below.

3.3.1 Spot-wise RSC Classifications

In order to evaluate the accuracy of AVL-Genius's RSC classification algorithm, one must first obtain a form of "ground-truth" to which AVL-Genius classification results can be. AVL-Genius captured a sequence of images along the test route with each capturing the RSCs of the short section covered by the image view. The individual images could be considered as point or spot observations along the test route. For the purpose of comparison, each image was manually classified in terms of lateral snow coverage, which was then compared to the computer classification result. The manual classification task was completed by a group of students trained with the same level of understanding of the classification rules to minimize the possible inconsistency and subjectivity. If the automatic RSC classification is the same as the corresponding manual RSC classification, the status of that image was said to be a match. If the opposite occurred, the image status was said to be mismatch. Presented with images taken by AVL-Genius's smartphone camera, the user is asked to choose one of the categories described in Table 3.3 in terms of snow/ice coverage on the pavement surface. This process was carried out for 23 days, capturing over 15,000 images that cover a variety of weather conditions.

Table 3.3 - Definition of Different Types of Lateral Snow Coverage

Lateral Snow Coverage	Description	Sample Image
Bare	At least 3 meters of the pavement cross-section in all lanes is clear of snow or ice.	
<25	Track between two wheel paths are clear of snow or ice.	
25 to 50	Both wheel paths are clear of snow or ice.	
50 to 75	Only one wheel path is clear of snow or ice.	
Full	No wheel path is clear of snow or ice.	

3.3.2 Route-level Classifications

To obtain summary statistics of the RSCs observed along an entire patrol route, point level classifications obtained from the system can be aggregated to the route. Assuming each point observation (i) represents segment i of length l_i, the classification results of each trip run are combined to generate summary statistics for the whole route using Equation 3.1.

$$\boldsymbol{P_k = \frac{\sum_i l_i \times \delta_i^k}{L}} \qquad \textbf{……… ………. (3.1)}$$

where P_k = percentage of the route having RSC class k;

l_i = length of the segment i;

L = total length of the route, $L = \sum_i l_i$

δ_i^k = 1 if the segment i has the RSC class of k; 0 otherwise.

If a single RSC class is to be designated for the entire route, TAC’s winter RSC classification guidelines can be followed to determine the class based on the frequency of occurrence of a RSC. For instance, if less than 10% of the route is affected by snow or ice, the RSC is considered to be bare. The resulting data was summarized according to patrol time and event.

Aggregation of spot-wise measurements to route-level classifications is essential in comparing system results to current monitoring methods because manual patrols are conducted at the route-level. Route-level classification accuracy was assessed by aggregating RSCs from manual and automatic system classifications according to TAC definitions and comparing the two methods. Route-level RSC classifications and the resulting summary statistics from the system were directly compared to the descriptive RSCs from corresponding patrol records for corresponding data runs. Visual side-by-side comparisons were also made between route-level RSCs generated by TRIP and those visualized by the smartphone-based system through a Google Maps interface.

3.4 Results

3.4.1 Summary of Test Data

Over 15,000 images were classified during 23 days of test runs, covering approximately 21 events, according to the precipitation data recorded by the RWIS stations. In reality, several more events were covered during data collection, since phenomena such as drifting snow are considered winter events that RWIS measurements may not capture. Moreover measuring precipitation such as snowfall can prove to be a challenging process for environmental stations, as snowfall is sometimes not measured even when an event has occurred. Figure 3.3 and Table 3.4 show the summary statistics of the field tests and the associated event characteristics, where average air temperature was -9.4°C and average daily precipitation was 0.6cm..

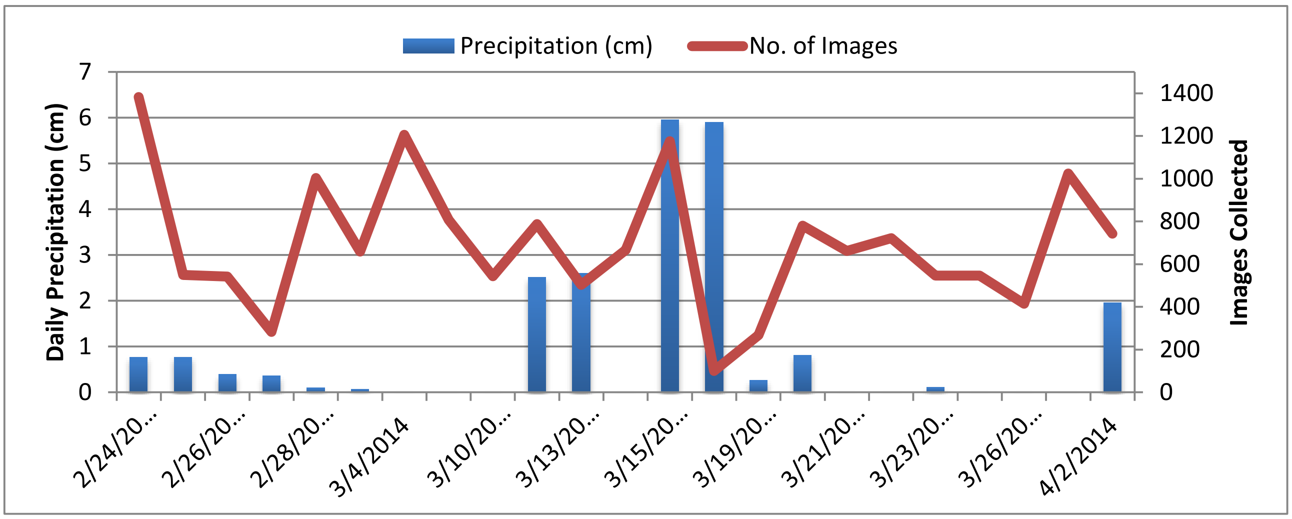


Figure 3.3 – Summary of Events and Images Collected

Table 3.4 – Summary of Event Attributes and AVL-Genius Images Captured

	Min	Max	Mean	Std. Dev.
Events	21			
Total Precipitation (cm)	0	5.95	0.6	1.3
Pavement Temperature (°C)	-23.3	22.5	-5.6	7.9
Air Temperature (°C)	-30	5	-9.4	7.2
Wind Speed (km/h)	0	45	16.7	9
No. of AVL-Genius Images	100	1206	666	312

3.4.2 Classification Accuracy

3.4.2.1 Spot-wise Condition Monitoring Accuracy

One of the key features of the AVL-Genius system is its ability to collect RSC data at specific locations or spots along the route being monitored. The first question of relevance for this investigation is therefore concerned with the accuracy of the system, i.e., how accurately can the system classify the RSC of each image? This performance is important as it reflects the system's ability to identify the time of occurrence and the location of poor road surface conditions, using the GPS and timestamp information associated with each classified image.

In answering this question, we must first obtain the "ground truth" of the RSC in each image. As discussed previously, this was done by manually classifying all images. Table 3.5 shows the confusion matrix of the classification results by AVL-Genius using the manual classification results as the "ground-truth". A total of 15,913 images collected by the data collection units over 20 events were manually classified and used for evaluating the spot-wise condition monitoring performance of the system.

Of all the images collected, a total of 10689 images (67%) were manually classified as bare. The AVL-Genius system correctly classified 82% of these bare condition images. Approximately 15% of these images were misclassified as Partly Snow Covered, which is somewhat expected considering that for some of the images there is only a small difference (in snow coverage) between bare and partly snow covered, especially in events of low precipitation. Three percent of bare conditions were misclassified as fully snow covered, which could be caused by the effect of glaring and residual salt as detailed in the later sections of this chapter.

A total of 4522 images (28%) were manually classified as Partly Snow Covered. Approximately 55% of these images were correctly classified by AVL-Genius, while 41% of them were classified as Bare and the remaining 4% as Fully Snow Covered. Interestingly, 30% of these images were associated with the lower end of the snow coverage scale (< 25%), which could account for a high number of partly snow covered conditions being misclassified as bare. The other main reason for misclassification is that dark coloured slush is not being accurately detected by the current image recognition algorithm.

The classification accuracy for Fully Snow Covered conditions was much lower. Of the 702 Fully Snow Covered images, approximately 38% were classified correctly, with 47% of them classified as

Partly Snow Covered and the remaining 15% as Bare. One of the main reasons for this problem was the high proportion of conditions showing wheel paths covered by slushy snow, appearing to be track-bare and thus classified as partly snow covered. A closer examination of the images automatically classified by AVL-Genius as fully snow covered shows 62% being manually classified as either 50%~75% snow covered or fully snow covered. This result indicates a possible overestimation of snow coverage by AVL-Genius, where if only one wheel track is clear of snow and ice, the resulting automatic classification is fully snow covered. A detailed discussion on the associated issues is provided the following section.

Table 3.5 - Confusion Matrix of AVL-Genius Classification Results

By number				
Manual Classification (Ground Truth)	AVL-Genius Classification			Total
	BP	PS	FS	
BP	8729	1575	385	10689
PS	1840	2484	198	4522
FS	106	330	266	702
Total	10675	4389	849	15913
By percentage				
Manual Classification (Ground Truth)	AVL-Genius Classification			Total
	BP	PS	FS	
BP	81.70%	14.70%	3.60%	100%
PS	40.70%	54.90%	4.40%	100%
FS	15.10%	47.00%	37.90%	100%

Legend: Bare – BP; Partly Snow Covered – PS; Fully Snow Covered – FS

3.4.2.2 Route Level Condition Monitoring Accuracy

The previous section evaluates the performance of the AVL-Genius system in classifying the RSCs based on the point-wise observations or individual images taken at locations along the test route. AVL-Genius can also provide summary statistics at a route level in terms of proportion of individual RSC types detected along a route. These route level statistics could be used to assess the performance of the system in providing aggregate information on the overall conditions of a patrol route in accordance with the current practice and needs of MTO. This section compares AVL-Genius results against manual classification, patrol observations and MTO's TRIP system.

3.4.2.3 AVL-Genius vs. Manual Classifications

Table 3.6 shows the summary statistics of the proportion of RSCs occurring over the route for two sample events. For each run, the proportions of RSCs are listed according to manual and automatic classification, with the single aggregated RSC class conforming to TAC definitions. It is shown that, while the proportions of individual RSC classes vary for each run, the single route-level RSC class matches perfectly with the manual class, with the exception of one run. This observation emphasizes that even though there is performance variation in the classification of individual images, system classification performance is satisfactory at the route level, which is representative of the current state of practice. These results remained consistent over the course of all remaining events.

Table 3.6 – Comparison of AVL-Genius and Manual Classifications for Route-level Conditions

Feb 28th 2014				
Run	Route-Level RSCs		Single TAC RSC	
	Manual	AVL-Genius	Manual	System
1	87% Partly Snow Covered 11% Bare 3% Fully Snow Covered	62% Partly Snow Covered 38% Bare	Partly Snow Covered	Partly Snow Covered
2	93% Partly Snow Covered 4% Bare 4% Fully Snow Covered	79% Partly Snow Covered 21% Bare	Partly Snow Covered	Partly Snow Covered
3	72% Bare 25% Partly Snow Covered 3% Fully Snow Covered	61% Partly Snow Covered 33% Bare 6% Fully Snow Covered	Partly Snow Covered	Partly Snow Covered
4	94% Bare 6% Partly Snow Covered	86% Bare 14% Partly Snow Covered	Bare	Partly Snow Covered
5	83% Bare 17% Partly Snow Covered	53% Bare 47% Partly Snow Covered	Partly Snow Covered	Partly Snow Covered
Mar 15th 2014				
Run	Route-Level RSC		Single TAC RSC	
	Manual	AVL-Genius	Manual	System
1	100% Partly Snow Covered	86% Partly Snow Covered 4% Bare 10% Fully Snow Covered	Partly Snow Covered	Partly Snow Covered
2	75% Partly Snow Covered 21% Bare 4% Fully Snow Covered	71% Partly Snow Covered 25% Bare 4% Fully Snow Covered	Partly Snow Covered	Partly Snow Covered
3	72% Partly Snow Covered 30% Fully Snow Covered	50% Fully Snow Covered 44% Partly Snow Covered 6% Bare	Partly Snow Covered	Partly Snow Covered

3.4.3 Comparative Analysis

3.4.3.1 AVL-Genius vs. Patrol Reports

As discussed previously, MTO relies on patrollers to monitor and report road surface conditions during winter events. The patrol reports give a qualitative description of the road weather and surface conditions over specific routes and an idea of the extent to which these conditions occur. To enable a comparison to the qualitative patrol reports, the point-wise condition classification data from AVL-Genius are aggregated to generate route-level condition statistics such as percentage of route with individual types of snow coverage.

Table 3.7 shows time-stamped conditions as reported by patrollers and corresponding smartphone-based system RSC classifications for a sample event occurring on March 12, 2014. Throughout the event, a variety of partly snow covered conditions were observed by patrollers and the system showed good correspondence and representation of these observed conditions along the route. The major difference was that the system provided quantitative results showing the frequency of RSCs, instead of the descriptive RSCs provided by the patrol records. For instance, on March 12, 2014, three RSC conditions (including track bare, partly snow covered and fully snow covered) were observed by the patroller over the first patrolling trip at 11:49am~12:00am. These conditions were well captured by AVL-Genius (4% Bare, 85% Partly Covered, and 11% Fully Snow Covered). Similar comparative analyses were performed on all events covered by field tests and the findings were similar to the sample indicated in Table 3.7 From this comparative analysis one can reasonably conclude that the system can be an effective alternative to the current method of patrol reporting. More importantly, the system offers the advantage of being objective by providing RSC statistics that allow more systematic performance measurement and more accurate condition forecasting.

Table 3.7 – Patrol Reports vs. AVL-Genius: Sample Event 2

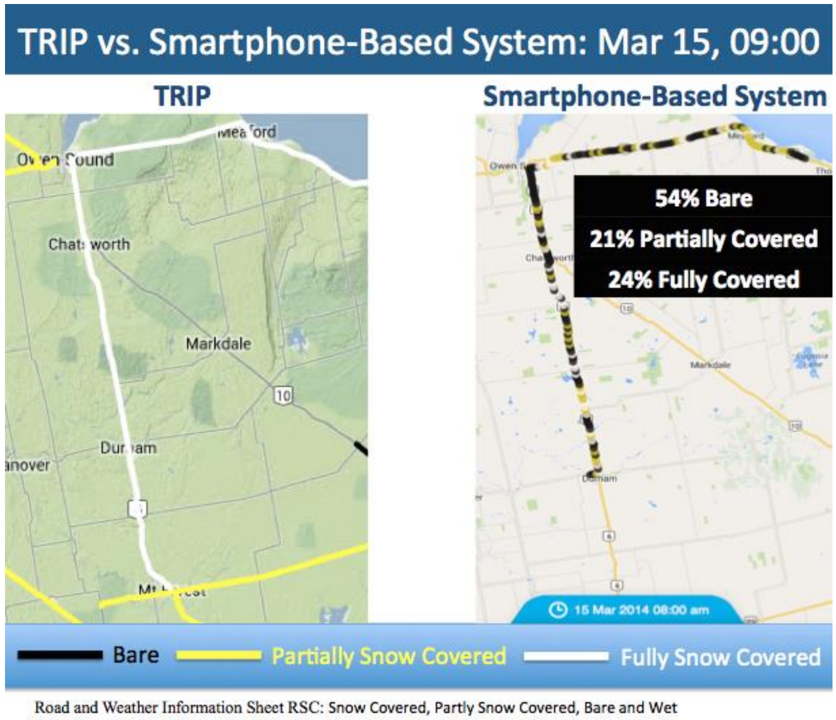
Mar. 12th, 2014		
Time	**Patrol Reports**	**AVL-Genius**
11:49 - 12:00	Track Bare Partly Snow Covered Snow Covered	85% Partly Snow Covered 11% Fully Snow Covered 4% Bare
14:10 - 14:20	Track Bare Partly Snow Covered Snow Covered Bare and Wet Partly Ice Covered	71% Partly Snow Covered 25% Bare 4% Fully Snow Covered
14:35 - 15:00	Track Bare Snow Covered Snow Packed Drifted Sections Slushy	50% Fully Snow Covered 44% Partly Snow Covered 6% Bare
15:00 - 15:30	Snow Covered Snow Packed Partly Snow Covered Partly Snow Packed	66% Partly Snow Covered 34% Fully Snow Covered
15:37 - 15:48	Partly Snow Covered Partly Ice Covered Snow Covered/ Packed Bare and Wet	72% Partly Snow Covered 21% Fully Snow Covered 7% Bare

3.4.3.2 AVL-Genius vs. TRIP System vs. Patrol Reports

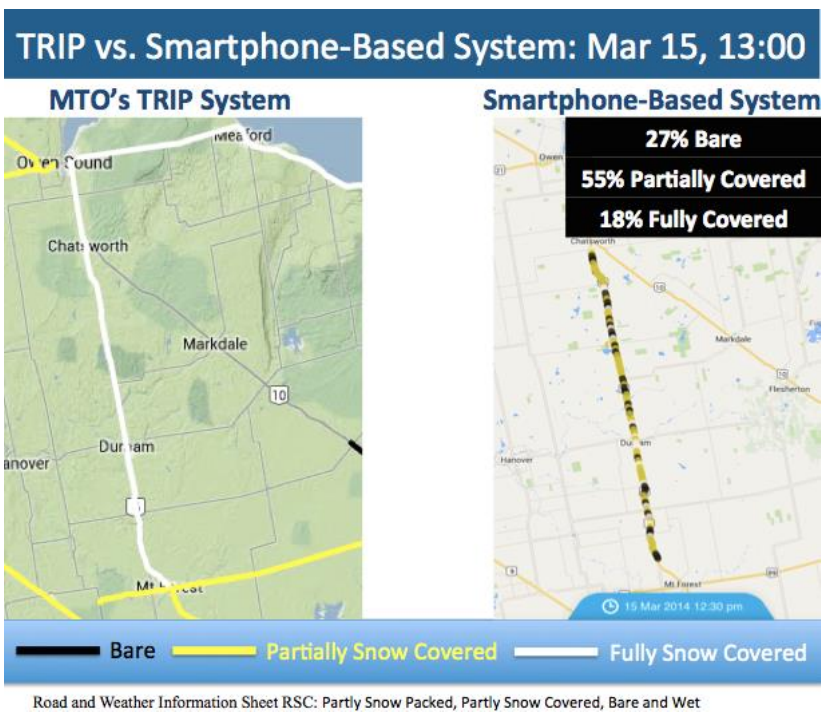
As discussed in the earlier section, MTO's TRIP is a Traveller Information Service providing road information for provincial highways in Ontario via an interactive Internet map application. While the road surface conditions that appear on TRIP's map interface also come from the field patrol reporting channel discussed previously, there is a lag between the time that TRIP updates the conditions and the time that the conditions are actually observed and reported. This section compares the conditions displayed in MTO's TRIP with those reported by AVL-Genius. Since the TRIP condition database was not available during this research, snapshots of TRIP's interactive map display were taken during individual snowstorms over the test period. For comparison, the same was done to obtain a display of the conditions generated by AVL-Genius. Note again that both systems provide visualization of the RSCs according to the national conventions established by TAC.

Figure 3.4a illustrates side-by-side comparisons of the RSCs of the test route between the two systems for an event occurring Mar 15, 2014 at 9:00am and 1:00pm, respectively. At 9:00am, TRIP showed fully snow covered conditions while the AVL-Genius showed predominantly bare, with approximately equal proportions of partly and fully snow covered conditions for the 70km route. However, according to TAC route classifications guidelines, the route would be classified as fully snow covered, thus corresponding exactly with TRIP.

Figure 3.4b shows TRIP displaying fully snow covered conditions while the smartphone system shows more than half of the route having partly snow covered conditions with a low proportion of fully snow covered RSCs. According to TAC guidelines, the route would be classified as partly snow covered, directly conflicting with TRIP. Moreover, patrol records indicate several types of partly snow covered conditions, thus corresponding with the system results, but no mention of fully snow covered RSCs is made. This suggests that TRIP's classification in this instance could be due to a time lag between patrol reporting and TRIP updates, since the TRIP updates originate from the RCWIS reports. However, during this study there was no way of determining the time lag between the scheduled TRIP update and the actual TRIP update. The smartphone system also shows a high degree of granularity as compared to TRIP, which is advantageous because it identifies the locations of poor RSCs. If these hotspots persist during and after winter events, WRM operators can pay a greater attention to the details in an effort to maintain safe roadways and motorists can identify deteriorating RSCs and drive along the route in a way that reduces collision risk.



(a)



(b)

Figure 3.4 - Comparison between Results from TRIP and Smartphone System

Table 3.8 summarizes ten similar side-by-side comparisons between the smartphone-based system and TRIP (screenshots are included in Appendix C). While there is remarkable consistency between the system and patrol reporting, there is also a clear discrepancy between these two data sources and the RSCs indicated by the TRIP system. As previously mentioned, these discrepancies are most likely due to a lag between the patrol reporting and TRIP system updating, both of which are a manual process, further underscoring the need for automating the RSC monitoring, data collection and reporting methods.

Table 3.8 - Summary of Smartphone-based System and TRIP Comparison

Date	Time	Road & Weather Info Sheet	AVL-Genius		MTO's TRIP RSC
			RSC Classification	TAC Classification	
28-Feb-14	13:00	Partly Snow Covered Partly Snow Packed Bare and Wet Bare and Dry	49% Bare 51% Partly Snow Covered	Bare	Partly Snow Covered
04-Mar-14	9:00	Partly Snow Covered Bare and Wet Bare and Dry	91% Bare 6% Partly Snow Covered 3% Fully Snow Covered	Bare	Bare
04-Mar-14	15:00	Partly Snow Covered Bare and Wet Bare and Dry	85% Bare 11% Partly Snow Covered 5% Fully Snow Covered	Bare	Partly Snow Covered
10-Mar-14	9:00	Partly Snow Covered Bare and Wet Bare and Dry	95% Bare 5% Partly Snow Covered	Bare	Bare
10-Mar-14	15:00	Bare and Wet Bare and Dry	97% Bare 2% Partly Snow Covered 1% Fully Snow Covered	Bare	Bare
14-Mar-14	9:00	Partly Snow Covered Bare and Wet Bare and Dry	69% Bare 23% Partly Snow Covered 8% Fully Snow Covered	Bare	Partly Snow Covered
15-Mar-14	9:00	Snow Covered Partly Snow Covered Bare and Wet	54% Bare 24% Fully Snow Covered 21% Partly Snow Covered	Fully Snow Covered	Partly Snow Covered
15-Mar-14	13:00	Partly Snow Packed Partly Snow Covered Bare and Wet	55% Partly Snow Covered 27% Bare 18% Fully Snow Covered	Fully Snow Covered	Fully Snow Covered
21-Mar-14	9:00	Snow Packed Partly Snow Covered Partly Snow Packed Bare and Wet Bare and Dry	78% Bare 22% Partly Snow Covered	Bare	Partly Snow Covered
25-Mar-14	9:00	Partly Snow Covered Partly Snow Packed Bare and Wet Bare and Dry	68% Bare 30% Partly Snow Covered 2% Fully Snow Covered	Bare	Partly Snow Covered

3.5 System Reliability

This section details the factors that affect the smartphone system's ability to perform its intended functions (i.e., determining the RSC of a maintenance route). Two main factors affect the reliability of an image based condition monitoring system such as AVL-Genius: availability of useful images and image classification accuracy. Availability is defined as the proportion of images taken by the smartphone camera on a given data collection run that are classifiable (i.e., useful for classification). Image classification accuracy is defined as the percentage of the classifiable images correctly classified. On average, 86% of the images were found to be classifiable with variations in trips and devices ranging between 60% and 100%. The following section describes the factors contributing to poor images and classification results, including visibility and image quality.

3.5.1 Visibility

In any image driven system, the quality of images captured by the device plays an important role in its eventual results. Since AVL-Genius uses the smartphone camera to capture images of the road surface, visibility is critical to the system's ability to classify RSCs. A variety of conditions such as dirty windshields, heavy precipitation and fog can compromise the visibility of the road surface in an image. These conditions did not frequently occur during the field experiment but still need to be addressed for potential users. For instance, on Mar. 12, 2014, one device showed over 40% of images having poor visibility, resulting in modest system performance with automatic RSC classification accuracy of 59%. Vehicle operators could manually adjust for situations posing a challenge to image quality such as such as poor camera view and dirty windshields, but situations with heavy precipitation and snow squalls will require additional improvement in image processing to make images useful. An automated procedure is being developed to detect and consequently exclude these images from the automatic RSC classification process. Figure 3.5a shows an image with poor visibility due to a dirty windshield.

3.5.2 Road Surface Contamination

The smartphone system essentially estimates the quantity of material that covers the road surface ahead (i.e., level of road surface contamination). Pavement surface contaminants can range in color and consistency from brown slush to white powdery snow. As mentioned during the accuracy evaluation, surface contaminants can affect classification accuracy, particularly with low contrast (dark pavement covered with dark contaminant). For example, automatic RSC classifications can be inaccurate when the road is covered with dried residual salt, which is often similar in appearance to snow cover on a road surface since both are whitish in appearance. Moreover, when there is high color contrast (very dark pavement covered with white dried salt), additional scrutiny and care are required to distinguish between snow cover and dried salt, even for manual classifications. The system currently tends to classify bare images covered with dried residual salt as partly or fully snow covered, but the system developer has partially addressed this issue by improving the RSC classification algorithm. Figure 3.5b shows an image with excess residual salt on the road surface.

3.5.3 Ambient Lighting

One issue that is yet to be addressed during this research is the classification of images captured at night. The system's current image recognition algorithm is not calibrated to classify night images; however, this feature is intended to be included in future iterations of the smartphone system. It is expected that as long as the view of the road surface is not obscured due to high reflectivity (for instance, from oncoming headlights) the system will perform as expected. Presently, the system therefore depends on images captured during times of sufficient ambient light (i.e., in the daytime). For the system to be considered fully practical, this issue needs to be addressed since maintenance personnel must frequently deal with snowstorms that occur at night. Figure 3.5c shows an image captured during the night-time field tests, which was excluded from analysis.

(a) (b) (c)

Figure 3.5 - Sample Images of Key Performance Factors: (a) Image with poor visibility (b) Dried residual salt on road surface (c) Image captured in night-time

3.6 Conclusions and Recommendations

This chapter describes a field study conducted to evaluate the performance of an automated smartphone-based automatic road surface condition (RSC) monitoring system called AVL-Genius. Using an image recognition algorithm to automatically detect the level of snow coverage from an image of the roadway, the system returns the classification to the end user graphically on a Google maps interface. AVL-Genius was deployed on four patrol vehicles during the Winter 2013-14 season and its performance was evaluated with respect to its accuracy and reliability.

When compared to manual spot-level (individual image) classifications, the system accurately classified 72% of 15,913 images. Images manually classified as bare, partly snow covered and fully snow covered were classified with accuracies of 82%, 55% and 38% respectively. Factors contributing to misclassifications were identified, with low/poor visibility being a primary concern. While the system performance was found to be satisfactory overall, partly snow covered conditions could benefit from improved classification accuracy and fully covered conditions require significant improvement. Moreover, the system is not currently calibrated to automatically classify images captured at night.

At the route level, the system presents a reasonable alternative to the current methods of winter patrol monitoring. AVL-Genius was found to be timelier, while offering more granularity than the current map-based Traveller's Information Portal (TRIP). AVL-Genius also currently classifies RSCs using a three-class system: bare, partly snow covered and fully snow covered. While these three RSC types may be sufficient for travellers, maintenance personnel often require additional details about road conditions to make better operational decisions. More detailed RSC classifications (such as extent of partial snow coverage) would therefore prove beneficial to the winter maintenance community. Finally, identifying the situations contributing to system inaccuracies is a prerequisite for exploring and developing improved condition monitoring and classification solutions.

Chapter 4
A Connected Vehicle Solution for Improving Winter Road Surface Condition Monitoring

This chapter describes a proposed methodology for improving the road surface condition monitoring system evaluated in Chapter 3. The major objective involves using data from a connected vehicle paradigm in which the vehicles can connect to and use the data captured by the nearest RWIS station in order to produce more accurate RSC information. Details on the system construct, model calibration and validation using field test data are discussed.

4.1 Problem Definition

As indicated in Chapter 3, the smartphone-based system was found to perform with an overall classification accuracy of 72%. For individual RSC types (bare, partly snow covered and fully snow covered) the classification accuracy was 82%, 55% and 38%, respectively. The test results on the smartphone system have indicated fair performance in classifying partly snow covered conditions, but poor performance in classifying fully snow covered conditions. From the perspective of maintenance personnel and the travelling public, classification accuracy of these RSC types needs to be improved for practical use of this monitoring tool. Moreover, maintenance personnel often require more detailed reports of RSCs occurring along a maintenance route for their decision-making process. The research effort described in this chapter addresses these issues with the following two primary objectives:

1) To improve the RSC classification accuracies observed in Chapter 3 using the existing RSC classification scheme.
2) To provide more detailed classifications for practical use by maintenance personnel, highway agencies and the public using a more detailed classification scheme.

4.2 Background

One proposed method of addressing the various performance issues with the existing RSC monitoring system is to apply a connected vehicle technology. A connected vehicle environment consists of vehicles, infrastructure, information services and travelers sharing information in order to operate more safely and efficiently with reduced environmental impact for all participants (U.S. Department of Transportation, 2014). Connected vehicle (CV) technology applies Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2V) communication to allow vehicles to communicate to other vehicles and the infrastructure in order to achieve transportation goals such as managing congestion and maintaining traffic safety. V2V applications include lane change warning, forward collision warning and left turn assist while V2I applications include curve speed warning, red light violation warning and transit pedestrian warning. One of the major potential applications of CV technology is road weather and condition monitoring. Real-time vehicle data such as harsh braking and initiation of windshield wipers can indicate road and weather conditions, thus providing a valuable source of information for motorists and highway agencies. Researchers have attempted to evaluate road conditions via onboard vehicle sensors for potential inclusion in a connected vehicle framework to be used for winter Maintenance Decision Support Systems (MDSS) (Sukuvaara & Nurmi, 2012; Chapman & Drobot, 2012; Hill, 2013). Technologies including mobile instrumentation and advanced communications have been integrated into maintenance fleets across U.S. jurisdictions such as Minnesota, Nevada and Michigan in order to provide decision-makers with information on windshield wiper status, traction control and road-surface condition data (U.S. Department of Transportation, n.d.). However, at present, this information is largely experimental and used as individual indicators of road and weather conditions for maintenance decision-makers. Therefore, at the writing of this thesis, no known system automatically combines mobile imaging and local weather data comprehensively to provide near real-time winter road condition data across highway networks to the maintenance community and the travelling public. This is the motivation behind solving the RSC monitoring problem using a connected vehicle solution.

In a CV system, data is shared between the vehicle and the existing infrastructure to achieve a specific goal. In this research, it is assumed that the CV technology will establish a communication link between the mobile units (e.g., vehicles with AVL-Genius or other image capturing systems) and nearby RWIS stations (representation of infrastructure). As a result, RWIS data could be used as supplementary information for improving the RSC detection results from a purely image based system. It should be noted that although the images captured in this research are obtained from a smartphone system, this data source could easily be replaced by a vehicle's onboard front end facing camera, which is already a standard component of many passenger cars, but used primarily for safety reasons such as vehicle, pedestrian and lane detection. It is therefore feasible to use onboard vehicle cameras to capture road surface images for automatic RSC monitoring, in a similar manner to the AVL-Genius smartphone-based system discussed in Chapter 3.

4.3 Methodology

The proposed connected vehicle based solution combines RWIS road and weather data with automatic image classification results to produce a more reliable and accurate RSC classifications. Figure 4.1 shows the methodology behind a working system, in which the vehicle captures GPS-tagged, time stamped images and selects the nearest RWIS station and appropriate data based on GPS and time stamps. These data sets are then input to a RSC classification model to obtain a new classification for each image. In practical terms, each RWIS station has a "zone-of-influence" (i.e., road and weather conditions observed at a given RWIS station are assumed to represent the existing conditions within a certain proximity to that RWIS station). Therefore, the size of each RWIS station's zone-of-influence depends on its proximity to another RWIS station. Figure 4.2 illustrates this concept where vehicle data captured within Zone 1 is automatically paired with RWIS Station 1, while vehicle data captured within Zone 2 will only be paired with RWIS Station 2.

The scope of research in this chapter is limited to the simulation of such a connected vehicle system via models with inputs that mirror data captured through a fully functioning connected vehicle RSC monitoring system as previously described. Communication technologies, protocol and network concerns are therefore outside the scope of this research.

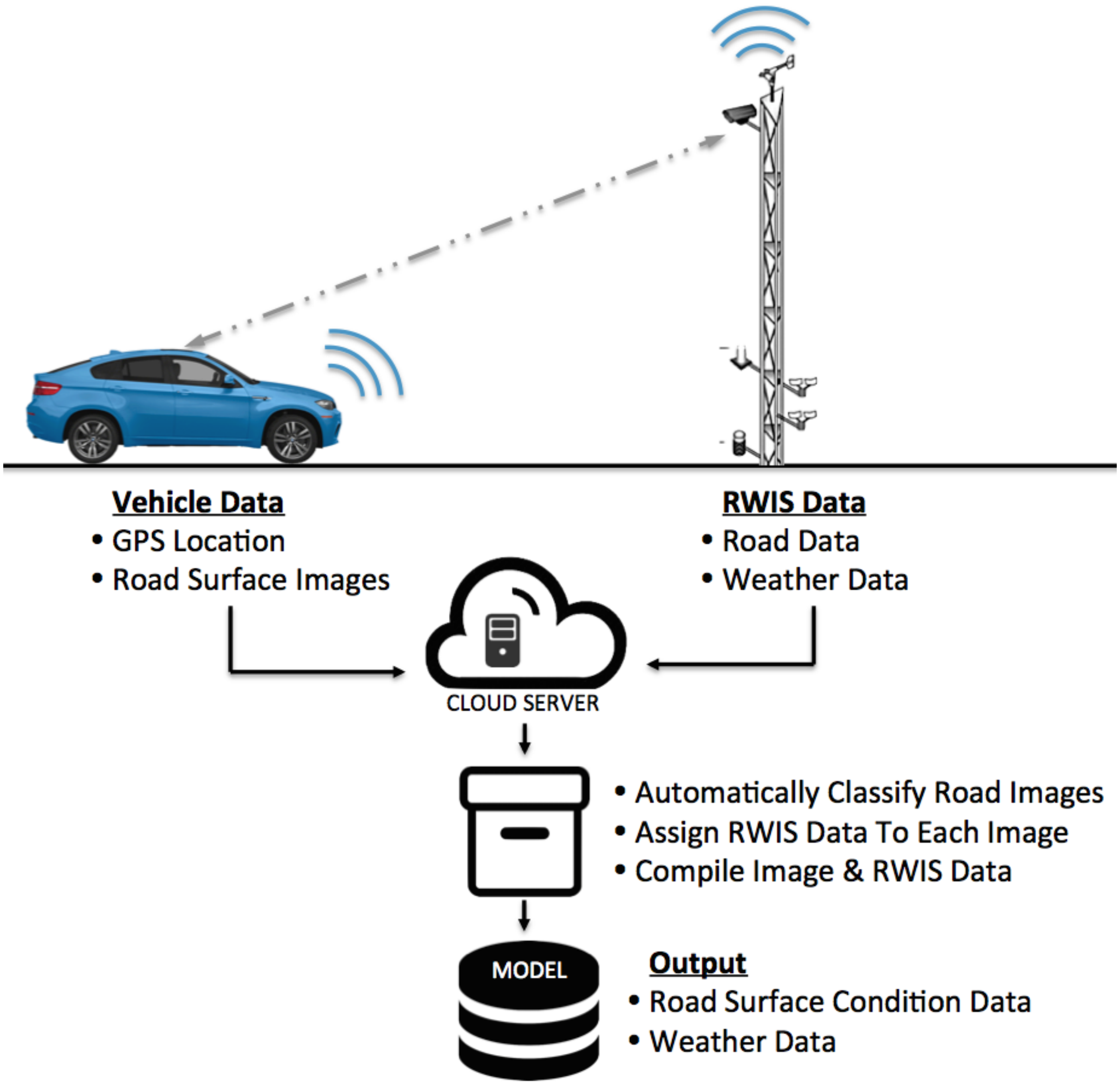


Figure 4.1 – Connected Vehicle Road Surface Condition Monitoring Framework

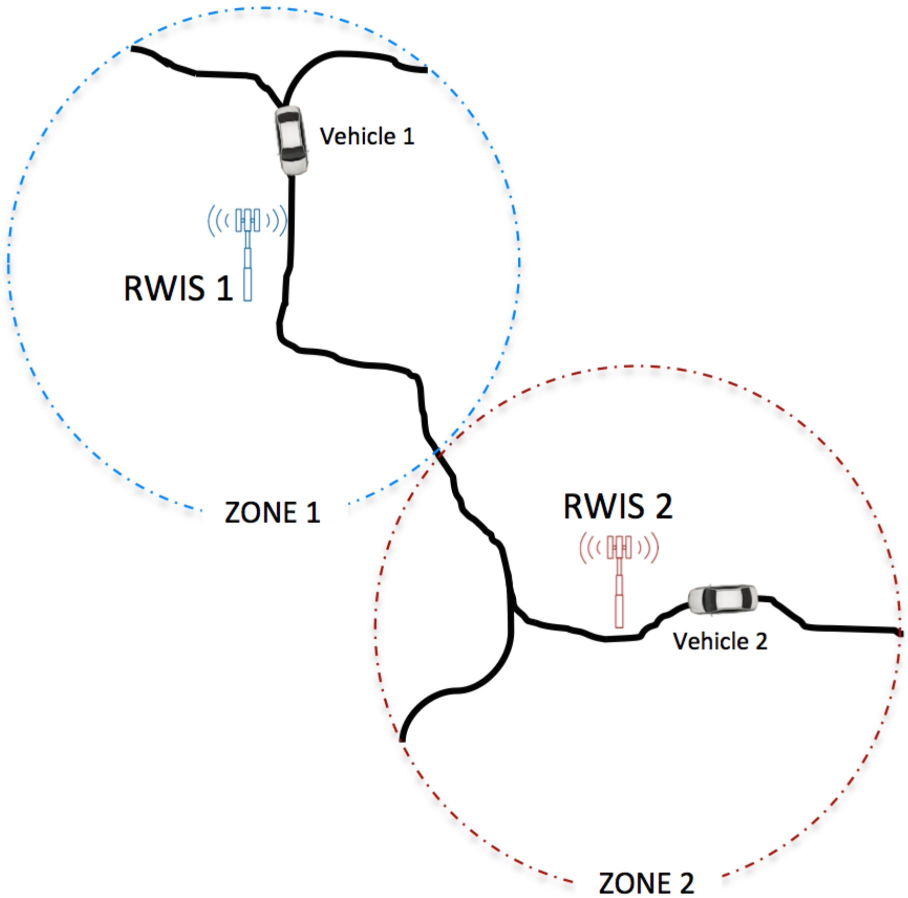


Figure 4.2 – Example of RWIS Zone-of-Influence

4.3.1 Modelling Framework and Background

The key component of the proposed solution is a RSC classification model that makes use of both AVL-Genius image classification results and RWIS data. As in Chapter 3, each image is manually classified into one of five classes according to the descriptions in Table 3.3. For a given image, if the class predicted by the model matches the manual class, the image status is considered a match. If the opposite occurs, the image status is considered a non-match. Accuracy is therefore defined as the proportion of matching images within the data set.

Model outputs are categorical classes of RSCs while input variables include the automatic AVL-Genius image classification and associated RWIS data. As indicated in the literature review, image-based RSC classification was traditionally done using Artificial Neural Networks (ANN), which provided accurate levels of model performance. Now considered one of the default mechanisms for RSC identification using road surface images, NN models are also used as a performance standard against which alternative models are compared. Predictive models used to solve these types of problems can be categorized as either statistical models or machine learning algorithms, with each having differing properties.

Statistical models are characterized by their transparent structure that allows users to interpret the effects of variables and understand their interaction with the model output. The advantage of these models is that they facilitate an understanding of their structure and make improvements easier because variable effects can be clearly interpreted. Popular types include linear regression, logistic regression and classification trees (Freedman, 2009; Bishop, 2006; Breiman et al., 1984).

Machine learning algorithms are often more complex models that capture the non-linearity between variables in complex data sets. As a result, their structures are difficult to interpret, and variable effects and interactions are unknown. Their performance may be better than traditional statistical models, but the user may have limited understanding of how the variables affect model performance and why. Popular machine learning algorithms include support vector machines (SVMs) and artificial neural networks (ANN) (Cortes & Vapnik, 1995; Ripley, 1996).

Both statistical models and machine learning algorithms have been used historically to solve classification problems. With the development of machine learning, black-box classifiers have gained popularity for solving pattern recognition problems. The following sections describe machine learning algorithms frequently used in solving classification problems, including decision trees and multilayer NN.

4.3.1.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) are black box classifiers used in machine learning to recognize patterns and model complex variable relationships. ANN models create outputs by learning patterns usually non-linear) from a training data set and by making predictions based on new inputs. As a result of their ability to handle large number of inputs as well as complex non-linear relationships, ANN have gained popularity for classification problems, especially in the fields of image and pattern recognition. The method behind many image-based automatic RSC classification systems is neural network modeling, and researchers have continued to advance this work through variations in the network structure and image feature input consideration.

There are different types of ANN models, but this study utilizes the popular multilayer perceptron (MLP) model due to its success in solving many classification problems (Kuehnle & Burghout, 1998; Conrad & Foedisch, 2003; Foedisch & Takeuchi, 2004; Hong et al., 2009). MLP models are supervised since they require an output from a training set in order to learn desired outputs. The model framework involves three main layers (input, hidden and output) with each consisting of interconnected nodes (neurons) as shown in Figure 4.3. The nodes in the input layer represent the potential feature variables while the output layer includes nodes for the prediction targets. When the training data is fed into the ANN, the input data is processed through a series of neurons and an output is generated. The output is then compared to the desired output and the difference is used to adjust the weights of the neurons in the hidden layer(s) to create a newer, more accurate output. This process, called backpropagation, is repeated until the errors between the model output and the desired output are minimized.

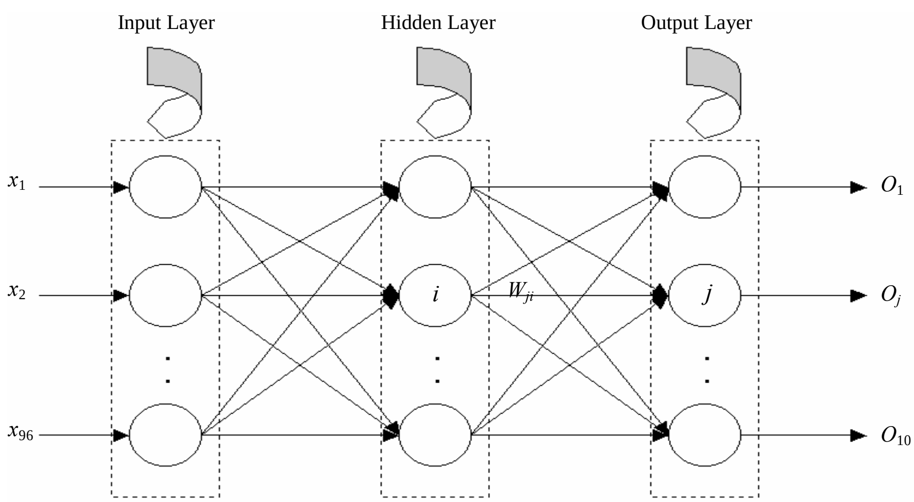


Figure 4.3 – Typical Neural Network Model Structure (Roy et al., 2005)

Although neural networks can produce accurate outputs with complex non-linear variable relationships in large data sets, their model structure does not allow the relationships between inputs and outputs to be understood. This limitation can prove challenging especially when solution optimization depends on understanding the relationship between variables.

Researchers have furthered developed neural networks to improve on initial algorithm postulation. The most recent evolutionary step of neural networks is the concept of deep learning. A deep neural network (DNN) is a feed-forward neural network with multiple layers of hidden units. DNNs learn hierarchical layers of representation from input in order to recognize patterns (Bengio, 2009; Hinton, 2007; Nguyen et al., 2015; Hinton et al., 2006; Hinton, et al., 2012; Deng et al., 2013). Capable of classifying objects in images with near-human-level performance, DNNs have also achieved notable success in areas such as acoustic modelling and speech recognition. Although they performed at human-competitive levels for image recognition tasks, DNNs can also be easily fooled into classifying unrecognizable images with near-certainty as members of a recognizable class. As a result, there are concerns about generalization for some tasks since there is a potential for costly exploitations of DNN-based solutions (Nguyen et al., 2015).

4.3.1.2 Classification Trees

A decision tree is a machine learning classification technique that predicts the value of a dependent variable using a set of rules in the form of a tree or flowchart structure as illustrated in to Figure 4.4. Decision trees used to solve classification problems are called classification trees, which have been used in various applications such as remote sensing, speech recognition, medical diagnosis and image recognition.

As a form of supervised learning, classification trees are developed by continuously partitioning a training data set into increasingly homogenous groups, through selection of the variables or attributes that result in the highest classification accuracy. The classification rules used in the attribute-partitioned process are illustrated using a tree-based structure in Figure 4.4, where every component of the decision-making process is included in the tree.

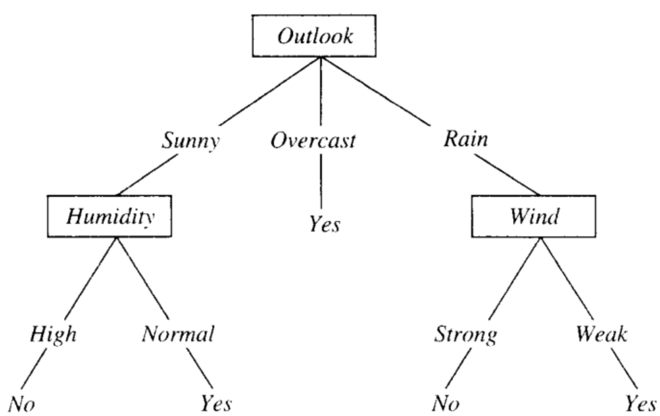


Figure 4.4 – Decision Tree (Mitchell, 1997)

On the one hand, classification trees have the advantage of implicitly screening variables since the nodes closest to the top (root) of the tree are essentially the most important. This automatic prioritization of variables occurs internally through the tree development process. Since linearity is not assumed, non-linearity between parameters does not affect model performance. Outliers also have little effect on decision tree classifiers since splitting at the branch level depends on the proportion of samples within the split ranges instead of absolute values. Their rule-based structure also makes them easy to interpret and explain, while requiring relatively little user effort in most cases to produce an output (Rokach & Maimon, 2007; Mitchell, 1997; Danilo, n.d.).

On the other hand, classification trees can be complex and may over-fit data as a result of being sensitive to the training set, sometimes requiring the change of an entire sub-tree if a minor change is made to a split close to the root. Diagrams can also become unreadable with large, complex trees, dramatically increasing user effort. Most algorithms also require that the target attribute have only discrete values, possibly limiting the types of problems they can optimally solve (Rokach & Maimon, 2007). Nonetheless, classification trees are commonly used to solve operational problems where the desired output is a class or dependent categorical variable.

Researchers have continuously modified tree algorithms in order to limit the aforementioned shortcomings, leading to several variations in classification and decision tree structures over the years. For instance, the CART algorithm (Classification and Regression Trees) (Breimanet al., 1984) was developed as an advanced version of the ID3 (Iterative Dichotomiser 3) (Quinlan, 1986) algorithm in order to improve on issues related to the ID3 structure such as overfitting, dealing with continuous variables and computational inefficiency. However, the CART algorithm could lead to unstable classification trees and also possess complex structures, particularly when dealing with non-linear data sets (Timofeev, 2004). Classification trees continue to evolve in order to increase prediction performance and computational efficiency; random trees and random forests are two recent examples. The following section describes the background of random trees and random forests as they pertain to solving classification problems.

4.3.1.2.1 Random Trees and Random Forests

Decision trees have evolved to overcome some of the shortcomings of the traditional versions of the algorithm as described in the previous section, such as overfitting and sensitivity to small changes in the training dataset. Two of the most successful classification tree variations are Random Trees and Random Forests (Breiman, 2001), have been repeatedly shown to offer the advantages of improved prediction accuracy and increased algorithm robustness. This section provides an overview of these two models.

Random Trees

A Random Tree (RT) is a randomly selected decision tree from an ensemble of trees constructed while considering a random subset of variables. The algorithm is constructed as follows (Thaseen & Kumar, 2013; Dhurandhar & Dobra, 2008; Vaidya & Fan, 2014).

1) From a total dataset ***S*** comprised of variable set **K**, with ***k*** variables, a decision tree $\boldsymbol{T_i}$ is constructed by considering a random subset of variables $\boldsymbol{k_i}$ (comprised of n random variables) at each node, instead of the entire variable set, K. Tree T_i is grown until no further splits are possible and no pruning is performed.
2) Step (1) is repeated for a sufficiently large number of trees N, producing an ensemble of trees in a Forest such that:

$$F = \{RT_1(x), RT_2(x), RT_3(x), \dots T_N(x)\} \text{…… } (4.1)$$

Where: F = ensemble of classification trees

RT_i = Random Decision Tree

3) Predictions from the RT algorithm are generated from a single tree drawn at random from forest F, such that

$$Y = RT(x) \text{ …… } (4.2)$$

Where: y = predicted class

RT = randomly selected tree from a set of possible trees in forest F

Since the tree is drawn at random, each tree in forest F has an equal chance of being selected (i.e. the distribution of random trees is uniform). Figure 4.5 illustrates the random tree procedure.

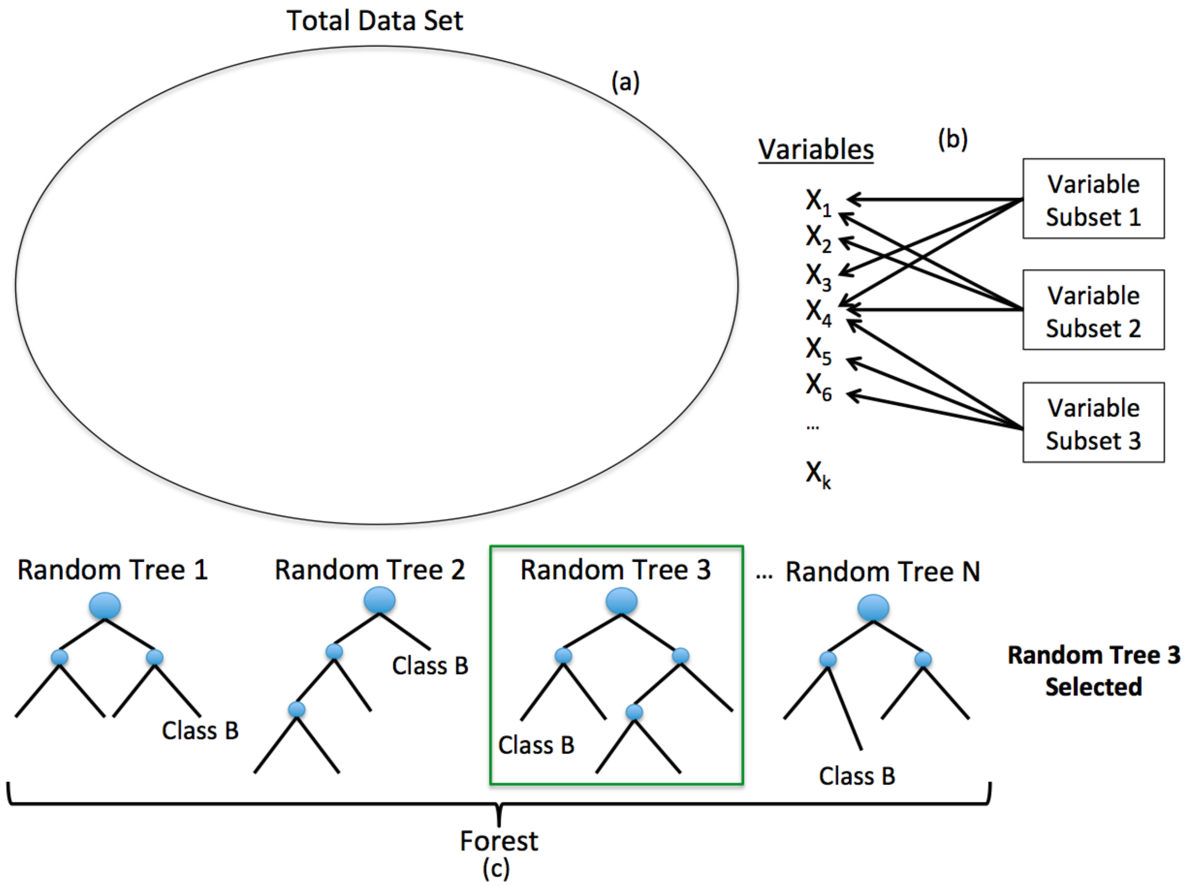


Figure 4.5 - Random Tree Process

Random trees (RTs) can be generated efficiently and while the process of randomly selecting a decision tree may seem counterintuitive, RTs generally lead to accurate models. Moreover, in comparative analyses, RTs have outperformed traditional decision tree classifiers with regards to prediction accuracy (in some cases by more than 50%) while exhibiting lower false alarm rates (Zhao & Zhang, 2008; Altınçay, 2007; Kalmegh, 2015; Thaseen & Kumar, 2013; Dhurandhar & Dobra, 2008; Vaidya & Fan, 2014; Fan et al., 2003).

Random Forests

A Random Forest (Breiman, 2001) is an ensemble of unpruned classification trees that produce a result based on the majority output from individual trees. For instance, if a random forest (RF) consists of 10 classification trees, with 8 trees predicting Class A and the remaining 2 trees predicting Class B, the RF output would be Class A since this is the most frequently predicted class by the classification forest. The process by which RF models operate is explained in detail below:

1) From a total dataset S with K variables, a bootstrapped sample $\boldsymbol{S_i}$ (random sampling with replacement) is drawn.

2) For each bootstrapped sample $\boldsymbol{S_i}$ a classification tree $\boldsymbol{T_i}$ is constructed by considering a random subset of variables $\boldsymbol{k_i}$ (comprised of n random variables) at each node instead of the entire variable set $\boldsymbol{K}$. Tree $\boldsymbol{T_i}$ is grown until no future splits are possible and no pruning is performed.

3) Steps (1) and (2) are repeated until a given number of trees $\boldsymbol{N}$ are constructed, producing a total of $\boldsymbol{N}$ bootstrapped samples during the classification process.

4) For new data, results are predicted by aggregating the predictions of the group of trees. In other words, prediction $\boldsymbol{y}$ for a given random tree is such that

$$y_i = T_i(x) \text{ ………. } (4.3)$$

Where: y_i = Predicted class

$T_i(x)$ = Decision tree grown from random data sample S_i

Since a random forest is comprised of N random trees, Random Forest prediction y_{RF} is given by,

$$y_{RF} = \{T_1(x), T_2(x), T_3(x), \ldots\ T_N(x)\} \text{ ……….. } (4.4)$$

If the predicted variable is categorical, the most frequently predicted class determines y_{RF}. However, if the predicted variable is continuous, y_{RF} is given by the average output of the tree ensemble. Figure 4.6 illustrates the Random Forest process.

5) Performance assessment is done internally. Recall that the random bootstrapped sample, S_i provides the training dataset for each tree T_i. This results in a separate remaining testing data set (S-S_i) that allows for reliable estimation of prediction error for each constructed sub-tree.

The prediction errors assessed from the random holdout data are aggregated similar to step (3), to generate an overall error estimate for the random forest out-of-bag error (out-of-bag refers to data not included in the tree building process).

With a sufficient number of trees in the random forest (RF), predictions tend to converge, resulting in a reliable algorithm without the concern of overfitting that occurs in traditional decision trees. RFs have also been tested and found to be relatively robust to outliers and noise, and their out-of-bag error has been hound to reliably indicate algorithm performance due to their generation from aggregated trees. Moreover, RF performance assessment does not require additional cross-validation due to its iterative bootstrapping procedure (Breiman, 2001; Liu et al., 2013; Svetnik et al., 2003; Liaw & Wiener, 2002).

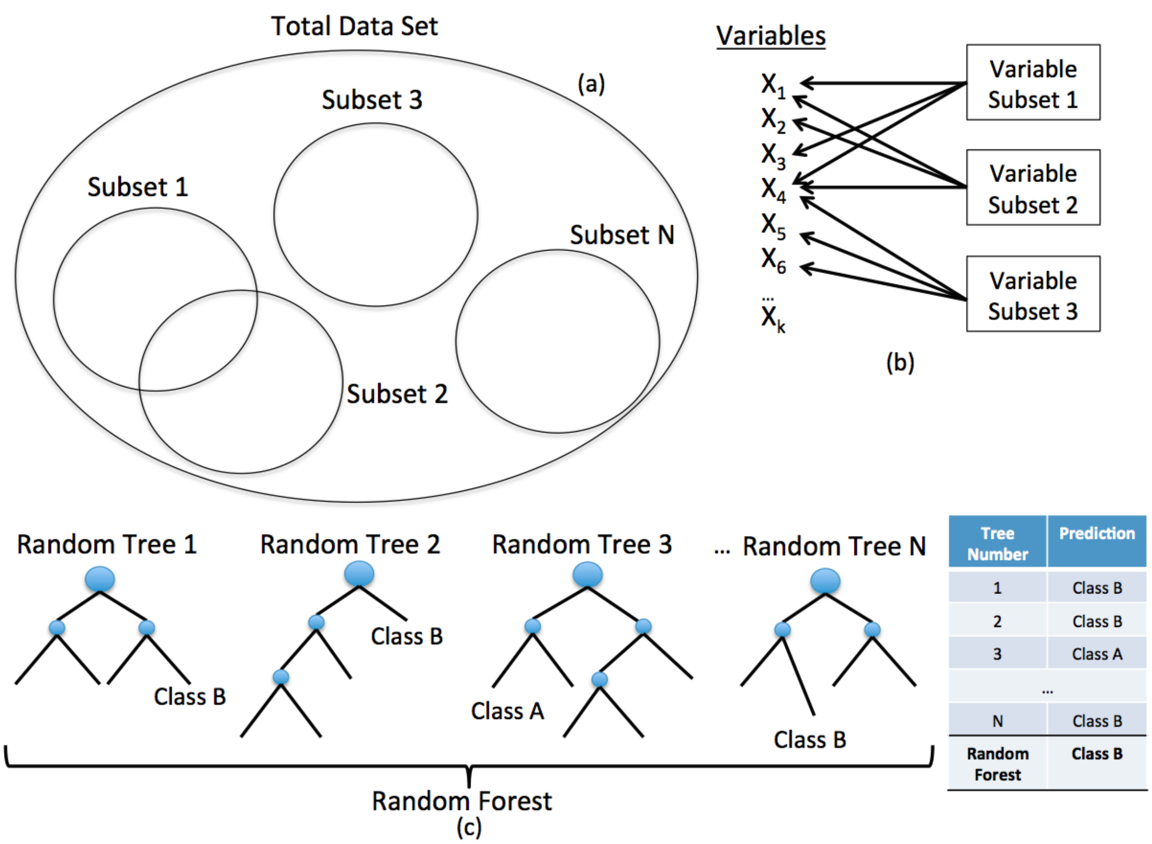


Figure 4.6 - Random Forest Process

4.4 Data Collection and Processing

4.4.1 Data Collection Overview

Data was collected from the test site described in Chapter 3. Images were captured from the smartphone system AVL-Genius, attached to four patrol vehicles and one dedicated mobile data collection unit. Additional data was collected from the RWIS stations located on the test route and the summary statistics are described in the section below.

4.4.1.1 Road Weather Information System Data

For each captured image, GPS data was used to identify the nearest RWIS station while timestamps were used to identify the applicable road and weather data. Model development and testing were performed using images with RWIS stations SW-25 and SW-13 as their corresponding data sources. In particular, data obtained from the dedicated MDCU significantly to the model data set, since images were captured at a high temporal frequency (10 continuous hours between each RWIS station) compared to the data obtained from patrol vehicles. Such high frequency has the potential to identify changing road and weather conditions over several hours at a time, allowing for recognition of any patterns that may assist in accurate RSC detection.

RWIS data is captured at a 20-minute frequency and the variables that are included in the connected vehicle model development described below as defined by the manufacturer.

- Relative Humidity – Percent of moisture in the air. A relative humidity of 0% shows that the air contains no moisture and 100% shows that the air is completely saturated and cannot absorb more moisture.
- Wind Speed – Average speed of the wind during an evaluation cycle
- 1hr Accumulation - Rainfall amount or snowfall liquid equivalent for the previous 1 hour period
- 3hr Accumulation - Rainfall amount or snowfall liquid equivalent for the previous 3 hour period
- 6hr Accumulation - Rainfall amount or snowfall liquid equivalent for the previous 6 hour period
- 12hr Accumulation - Rainfall amount or snowfall liquid equivalent for the previous 12 hour period
- 24hr Accumulation – Rainfall amount or snowfall liquid equivalent for the previous 24 hour period
- Precipitation Intensity – Intensity of the precipitation as derived from the precipitation rate

- Salinity – Salinity is roughly the number of grams of dissolved matter per kilogram of seawater. Units shown in parts per 100,000
- Surface Status - Condition of the pavement surface
- Air Temperature - Air temperature at the site.
- Surface Temperature - Temperature of the pavement sensor roughly 3 mm (1/8 inch) below the surface of the sensor.

Tables 4.1 and 4.2 show summaries of the variable features associated with the observations for RWIS stations SW-13 and SW-25. A total of 5363 samples were included for model development, representing a variety of road and weather conditions during 18 unique days of data collection.

Table 4.1 – Summary Statistics of Variables Used for Model Development

		Number of Observations = 5363			
Field Name	Unit	Min	Max	Mean	SD
Surface Temperature	°C	-19.2	18.6	-3.54	5.52
Air Temperature	°C	-25	5	-8.04	5.99
1hr Accumulation	cm	0	0.6	0.05	0.12
3hr Accumulation	cm	0	1.53	0.18	0.39
6hr Accumulation	cm	0	3.52	0.43	0.89
12hr Accumulation	cm	0	5.19	0.75	1.43
24hr Accumulation	cm	0	5.66	0.78	1.48
Wind Speed	Km/h	0	84	23.81	18.72
Relative Humidity	%	0	100	82.65	14.57
Salinity	parts/100,000	0	35830	7769.75	9283.62

Table 4.2 – Categorical Variable Sample Size

Field Name	Categories	Size	%
Surface Status	Chemically Wet	432	8.1%
	Dry	886	16.5%
	Ice Warning	1147	21.4%
	Ice Watch	439	8.2%
	Other	65	1.2%
	Snow Watch	1024	19.1%
	Trace Moisture	1303	24.3%
	Wet	67	1.2%
Precipitation Intensity	Moderate	539	10.1%
	None	4069	75.9%
	Slight	755	14.1%

4.4.2 Exploratory Analysis

Before starting any model calibration, an exploratory data analysis is often conducted to identify trends in the raw dataset. Exploring variable interaction and relationship help identify trends such as linearity and non-linearity, providing support for model selection. Moreover, a genuine understanding of conditions that contribute to particular outcomes allows model features to be explained in relation to intuition and existing physical phenomena. The following section details an exploratory analysis of the road and weather conditions associated with particular RSC types.

Figure 4.7, which illustrates the trends between RSCs and the road and weather conditions measured by nearby RWIS stations, reveals some interesting patterns. As the level of snow coverage increases, the salinity level also increases. There is also less variation observed with fully snow covered surfaces. This overall trend makes intuitive sense since salinity can indicate the quantity of deicing chemical present on the road surface and thus indicates that maintenance operations have been carried out. Typical practice involves applying salts in accordance with the amount of snow and ice present on the surface, and as a result, salinity levels could be used to discriminate RSC types. Surface temperatures are shown to decrease as RSCs worsen. Moreover, there are no instances of fully snow covered surfaces when the pavement temperature is at or above 0°C. Relative humidity is shown to have the opposite effect, as fully snow covered RSCs are generally associated with higher levels of humidity than bare and partly snow covered RSCs.

Figure 4.8 shows the proportion of RSCs captured by the images according to RWIS pavement status. The major observation is that fully snow covered surfaces tend to occur primarily when the surface status indicates snow watch, ice watch or ice warning. When the surface status indicates dry, the RSC is highly likely to be dry.

Figure 4.9 shows a scatter plot of salinity against pavement temperature with two distinct trends observed. When salinity is below approximately 3000 parts per 100,000 the RSC appears to be mostly bare regardless of the pavement temperature. Above this threshold, as salinity increases and pavement temperature decreases, partly and fully snow covered RSCs become more prevalent.

This exploratory analysis reveals some trends in the relationship between RSC types and other road and weather conditions. The variable associations identified seem to make intuitive sense, providing some important insights for model development, especially when machine learning algorithms are to be developed.

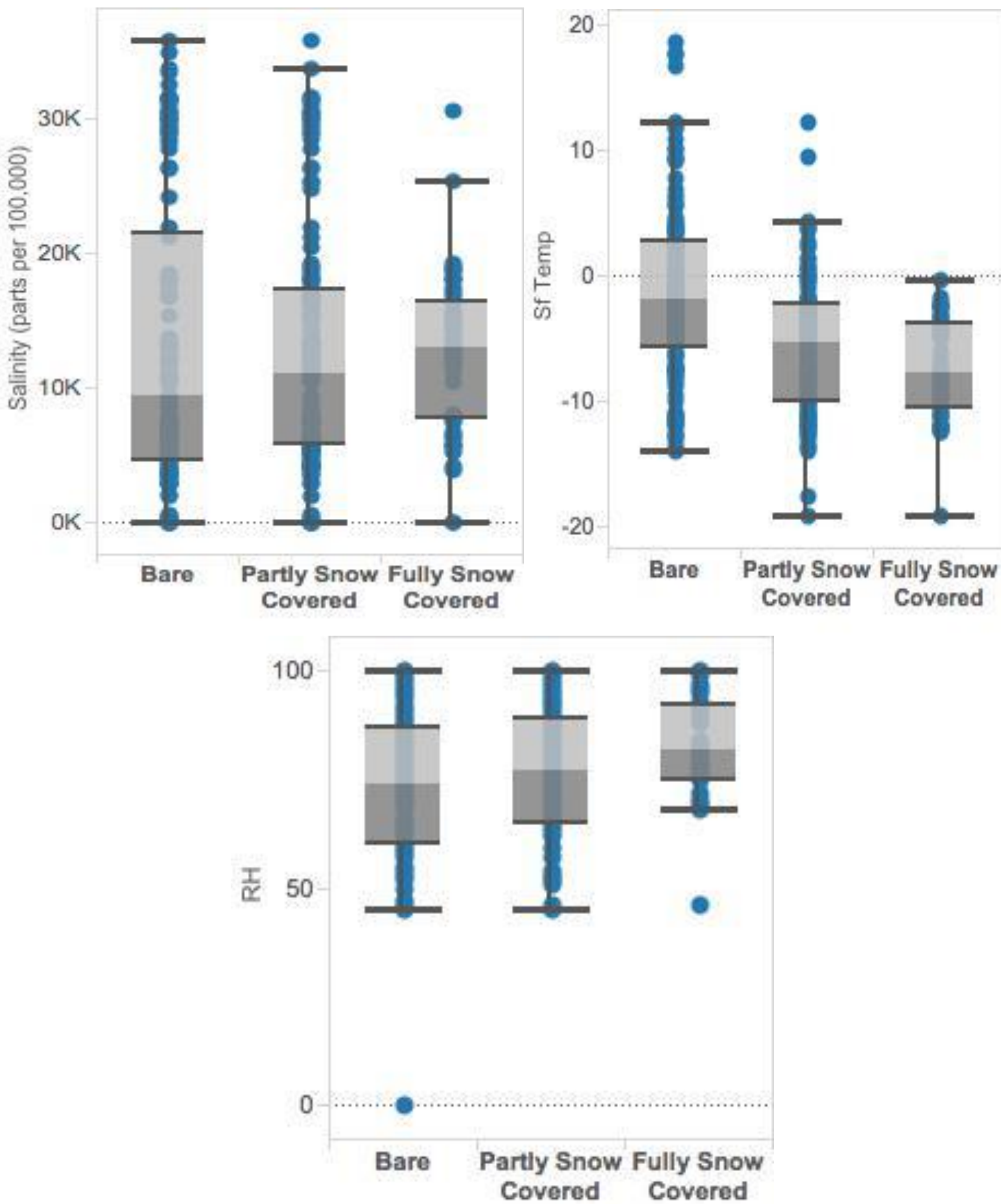


Figure 4.7 – Box Plot of RWIS Conditions by RSC Type

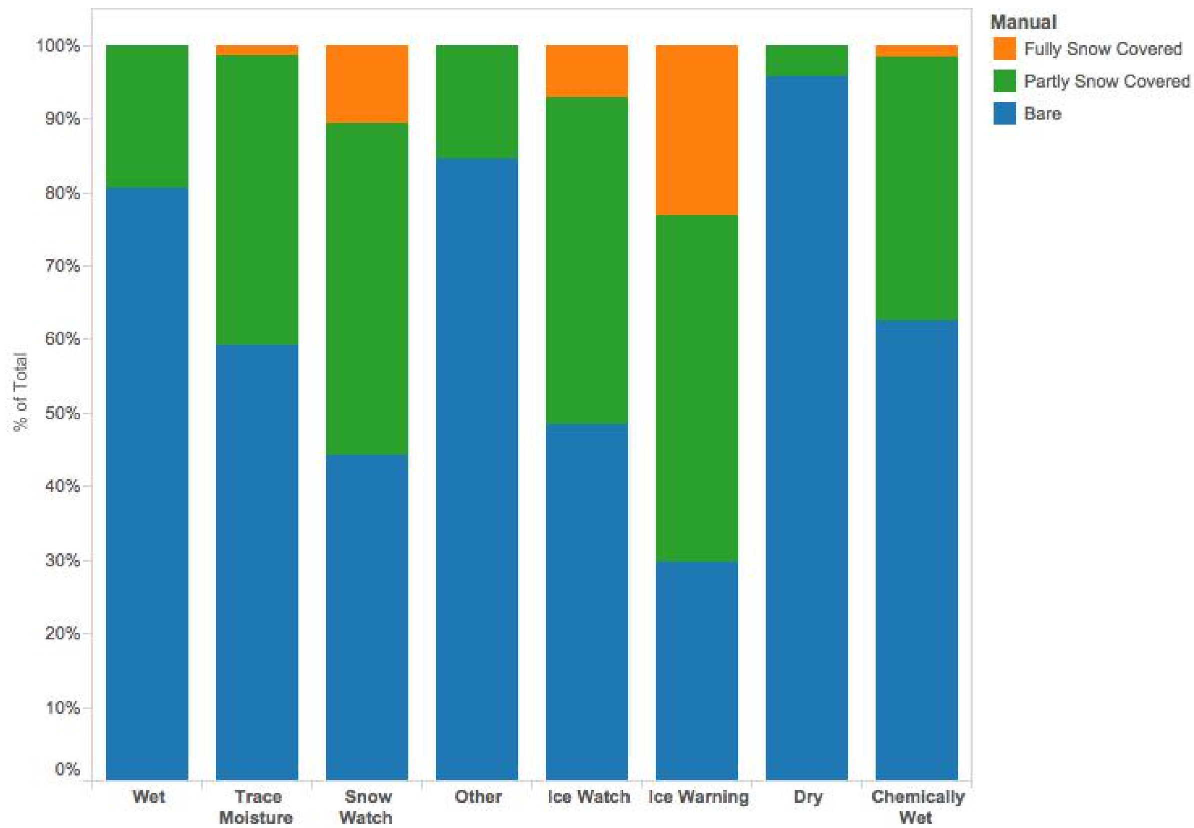


Figure 4.8 – Percentage of RSC classification by RWIS Surface Status

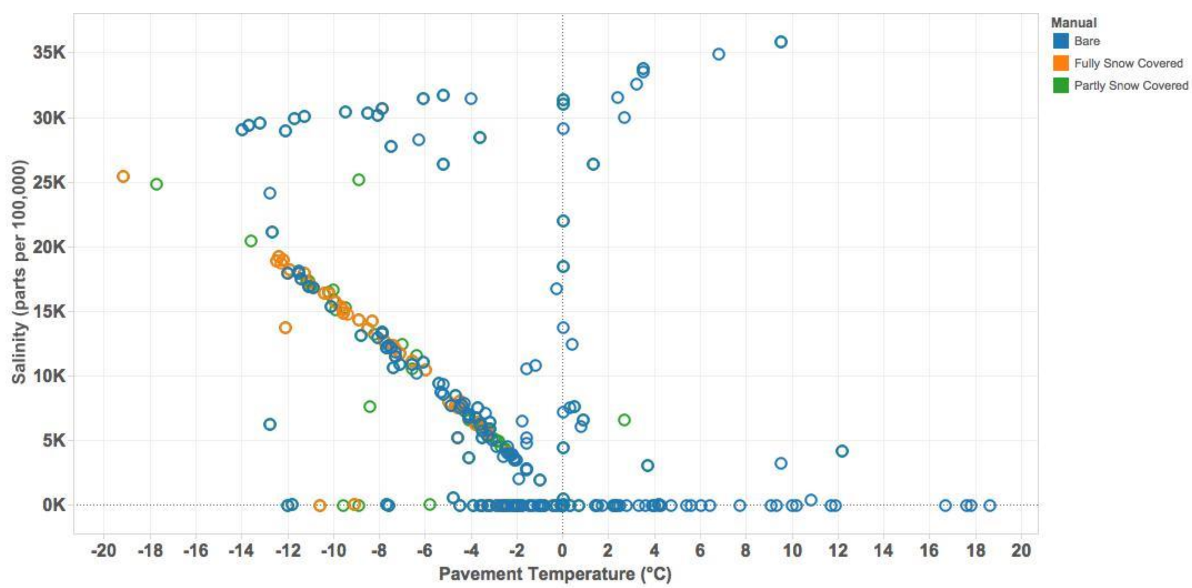


Figure 4.9 – Salinity vs. Pavement Temperature

4.5 Model Calibration and Validation

Statistical software Weka (Waikato Environment for Knowledge Analysis) was used to calibrate the ANN and decision tree models (Hall et al., 2009). As previously stated in the objectives of this chapter, the purpose of model development is to improve the classification accuracy of the current three-class system and to provide more detailed RSC classifications according to the five-class RSC classification scheme used to manually classify each image, as described in Table 3.3.

This section details the models and results obtained while using a 30% holdout rate for the dataset of 5363 observations. The resulting training and testing sets included 3754 and 1609 observations respectively. All model calibration was performed using 10 fold cross-validation in order to limit the effects of issues such as overfitting, which is a concern for classification tree models as noted in the preceding section. Moreover, training and testing data sets were identical for all calibration and validation of all models in order to maintain consistency in assessing model performance. Multiple configurations of neural networks, random trees and random forests were each calibrated and model configurations were selected for evaluation based on primary and secondary criteria of classification accuracy and processing time, respectively. For instance, when several neural networks are calibrated, the neural network configuration that results in highest classification accuracy for the calibration data set is selected as the "optimal" model variant. If multiple model configurations result in the same classification accuracy, the model configuration with the lowest processing time is selected as the "optimal" model, since processing time is a major consideration for real-time application.

Models will be compared using performance measures such as classification accuracy (as described in Chapter 3), processing time and Kappa statistic (or Kappa coefficient), which measures pairwise agreement between a set of outputs (observed and predicted) while making adjustments and correcting for expected chance of agreement (Carletta, 1996). Kappa statistic (K) is defined by Equation 4.5.

$$K = \frac{P(A)-P(E)}{1-P(E)} \quad \ldots\ldots\ldots (4.5)$$

where K = Kappa statistic

P(A) – proportion of times predicted and observed outputs match

P(E) – proportion of times predicted and observed outputs are expected to match by chance

When there is no agreement other than that which is expected by chance, K=0. When there is total agreement between observed and predicted, K equals 1. Kappa statistic is a measure that accounts for the fact that observed and predicted output may agree only by chance instead of accurate model calibration. While interpretation may vary according to researcher and application with respect to the threshold for acceptable model agreement, most statisticians typically prefer a minimum value of 0.6 (Carletta, 1996; Viera & Garrett, 2005). The following section describes model calibration and evaluation using the previously mentioned performance metrics for the RSC classification problem.

4.5.1 Three-Class Road Surface Condition Classification

4.5.1.1 Artificial Neural Networks

A three-layer ANN model was calibrated with an input layer that includes AVL-Genius image classification (bare, partly snow covered, fully snow covered) and the data received from the RWIS stations (relative humidity, precipitation intensity, 1hr, 3hr, 6hr, 12hr and 24hr accumulation, salinity, pavement surface status, air temperature, pavement temperature). More than 80 neural network models were investigated to obtain the optimal model configuration based on performance, with results included in Appendix E. Figure 4.10 illustrates the optimal calibrated ANN model with a learning rate of 0.1, momentum of 0.1 for the backpropagation algorithm and one hidden layer with 13 hidden neurons. The number of hidden neurons was determined using a commonly applied heuristic rule described in Equation. 4.6 (Hall et al., 2009).

$$\text{Hidden Neurons} = (\text{No. of Input Attributes} + \text{No. of Classes})/2 \quad \ldots.. \quad (4.6)$$

In this dataset, there are 3 output classes with 24 input attributes (Figure 4.10) resulting in 13 hidden neurons. Although numerous models were calibrated, more complex models often lead to a point of diminishing returns since computational times increased substantially when adding additional layers to the model. For instance, adding a second layer to the model in Figure 4.10 with the same number of neurons resulted in an 89% increase in computational time but only a 0.6% increase in overall classification accuracy.

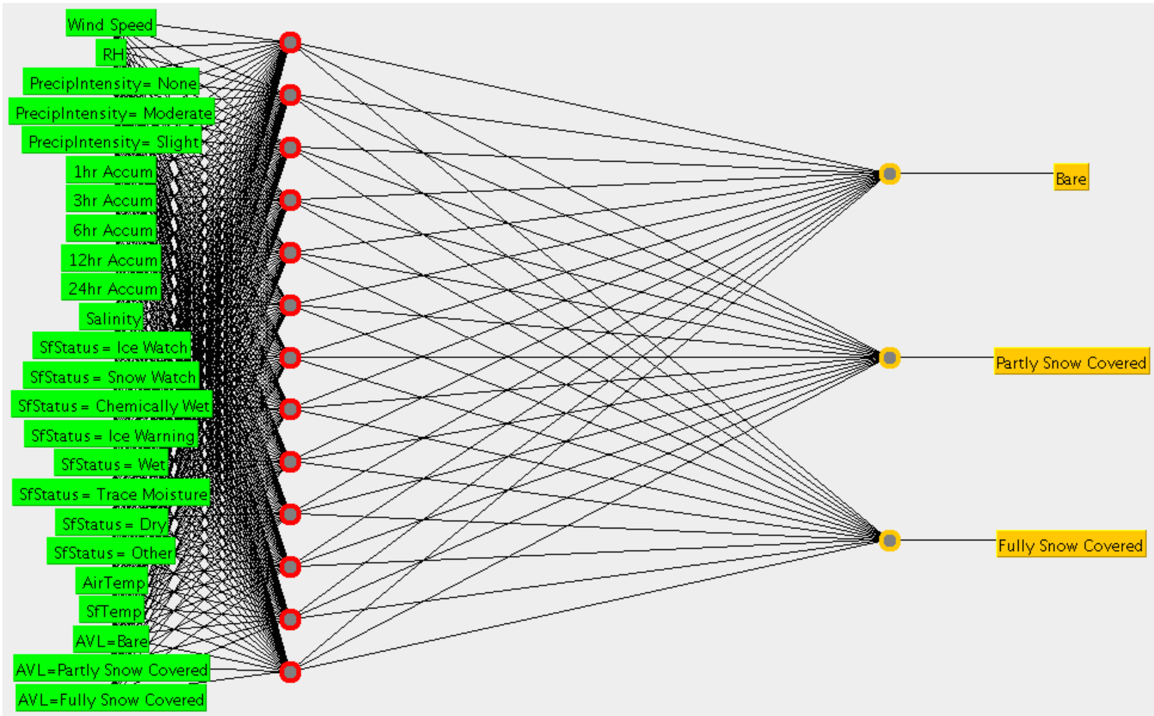


Figure 4.10 –Neural Network RSC Classification Model (Three-Class System)

Table 4.3– Neural Network Model Characteristics (Three-Class System)

Model Calibration Summary	
Correctly Classified Instances	3165 (84%)
Incorrectly Classified Instances	589 (16%)
Kappa Statistic	0.71
Total Number of Instances	3754

Table 4.4 – Neural Network Model Calibration Characteristics (Three-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.937	0.184	0.865	0.9	0.932
Partly Snow Covered	0.749	0.095	0.816	0.781	0.897
Fully Snow Covered	0.614	0.016	0.774	0.685	0.937
Weighted Avg.	0.843	0.138	0.84	0.84	0.92

Table 4.5 – Neural Network Model Calibration and Validation Results (Three-Class System)

	Calibration Data				
	Automatic				
		BP	PS	FS	Accuracy
Manual	BP	1963	121	10	93.7%
Manual	PS	295	1014	45	74.9%
Manual	FS	11	107	188	61.4%
	Overall Accuracy = 84.3%				

	Validation Data				
	Automatic				
		BP	PS	FS	Accuracy
Manual	Bare	859	54	1	91.2%
Manual	PS	108	422	40	74%
Manual	FS	3	31	91	72.8%
	Overall Accuracy = 85.3%				

Legend: Bare – BP; Partly Snow Covered – PS; Fully Snow Covered – FS

Tables 4.3 and 4.4 show the characteristics of the calibrated ANN model. A Kappa statistic of 0.71 indicates substantial agreement between model classification and expected output, based on the benchmarks established by several researchers in the past (Landis & Koch, 1977; Altman, 1991; Emam, 1998). Unfortunately, due to the black-box nature of the ANN, variable effects cannot be described easily.

Table 4.5 shows the calibration and validation performance of the ANN model, which demonstrated an overall hit rate of 84.3% and 85.3% for calibration and validation data, respectively, showing consistency in model performance. For calibration data, hit rates for bare, partly and fully snow covered conditions are 93.7%, 74.9% and 61.4%, respectively. Similarly, validation hit rates for bare, partly and fully snow covered conditions are 91.2%, 74% and 72.8%, respectively. This model demonstrates improvement in overall and individual RSC class estimation; however, further details on observed improvement are discussed in the model comparison section that follows.

4.5.1.1.1 Random Tree Model

During exploratory analyses, several types of decision trees were investigated to determine the most appropriate classification models. The models that showed the best performance were Random Trees and Random Forests.

Table 4.6 shows the calibration results of the random tree (RT) model for RSC classification. The RT structure consisted of a tree constructed while considering 4 random variables out of a possible 14 at each node. Alternate model structures were investigated but no substantial difference in classification accuracy was observed. Results of these alternate configurations are listed in Appendix F. The RT model was found to have a Kappa statistic of 0.74, which indicates strong agreement between actual and predicted RSC classes.

Table 4.6 – Random Tree Characteristics (Three-Class System)

Model Calibration Summary	
Correctly Classified Instances	3232 (86%)
Incorrectly Classified Instances	522 (14%)
Kappa Statistic	0.7429
Total Number of Instances	3754

At the root of the tree is the original AVL-Genius image classification, which makes sense because as noted in Chapter 3, image-based classification accuracy of bare surfaces was found to be as high as 82%. Therefore, an initial classification of a bare surface would require less additional data and modeling efforts to confirm, compared to AVL-Genius classification of fully snow surface, which demonstrated a much lower accuracy of 38%. The next sequential variables in the tree are salinity, surface status and relative humidity. As explained in the methodology section of this chapter, the variables closest to the root of the tree, considered to be most important to the classification process, are automatically determined during tree construction. Salinity level can indicate the amount of deicing chemicals on a pavement surface, thus suggesting that maintenance operations have been carried out. This variable could reflect the state of the RSC, since maintenance operations typically occur after accumulation on the road surface has begun. Surface status can give additional information about the spot RSC, particularly if the surface is wet, slushy or snow covered. Intuitively, salinity combined with surface status can provide details contributing to more accurate RSC estimation. For instance, high levels of salinity on a dry surface could indicate dried residual salt, a situation that was found to contribute to misclassifications by an image based system. According to an exploratory analysis, relative humidity was found to have a high correlation with accumulated snowfall, particularly when the surface status was listed as “snow watch”. It makes intuitive sense why these factors would be considered as important variables to the RT model since they can describe a likely RSC resulting from particular maintenance operations and weather conditions.

Table 4.8 shows classification results from both calibration data and 30% holdout validation data. Overall hit rates for calibration and validation data were 86% and 85.3% respectively. Calibration data demonstrated hit rates for bare, partly and fully snow covered conditions at 94.3%, 77.7% and 67.3% respectively. Similarly, validation data hit rates for bare, partly and fully snow covered conditions are 93.8%, 75.6% and 68%, indicating good consistency in model performance for the two data sets.

Table 4.7 – Random Tree Calibration Characteristics (Three-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.943	0.161	0.881	0.893	0.938
Partly Snow Covered	0.777	0.085	0.838	0.765	0.909
Fully Snow Covered	0.673	0.015	0.802	0.647	0.934
Weighted Avg.	0.861	0.122	0.859	0.827	0.927

Table 4.8 – Random Tree Calibration and Validation Results (Three-Class System)

	Calibration Data				
	Automatic				
		BP	PS	FS	Accuracy
Manual	BP	1974	112	8	94.3%
	PS	259	1052	43	77.7%
	FS	8	92	206	67.3%
	Overall Accuracy = 86.1%				

	Validation Data				
	Automatic				
		BP	PS	FS	Accuracy
Manual	BP	857	54	3	93.8%
	PS	108	431	31	75.6%
	FS	2	38	85	68.0%
	Overall Accuracy = 85.3%				

Legend: Bare – BP; Partly Snow Covered – PS; Fully Snow Covered – FS

4.5.1.1.2 Random Forest Model

A random forest (RF) comprising 10 classification trees was modeled with each tree including four random features (out of 14 features available). This RF structure was selected after a number of iterations on alternative combinations as found in Appendix G. Table 4.9 shows calibration results of the RF model. A Kappa statistic of 0.75 is considered to indicate good agreement strength between the model prediction and expected output.

Table 4.9 - Random Forest Characteristics

Model Calibration Summary	
Correctly Classified Instances	3238 (86%)
Incorrectly Classified Instances	516 (14%)
Kappa Statistic	0.7468
Total Number of Instances	3754

Table 4.10- Random Forest Calibration Characteristics (Three-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.943	0.149	0.889	0.914	0.948
Partly Snow Covered	0.782	0.087	0.853	0.808	0.919
Fully Snow Covered	0.68	0.017	0.776	0.725	0.949
Weighted Avg.	0.864	0.116	8.86	0.86	0.938

Table 4.11 shows classification results of calibrated and validated data for the RF RSC prediction model. The model was found to have an overall classification accuracy of 86% and 85.4% for calibration and validation data, respectively. Bare, partly and fully snow covered surfaces were classified with accuracies of 94.3%, 78.2% and 68%, respectively for calibration data. For validation data the RF model classified bare, partly and fully snow covered surfaces with accuracy of 93.8%, 75.1% and71.2%, respectively. In addition to showing consistency in model performance between the calibration and validation data, the RF model demonstrated the ability to improve the classification accuracy for each RSC class compared to the AVL-Genius results.

Table 4.11 – Random Forest Calibration and Validation Results (Three-Class System)

	Calibration Data						Validation Data				
	Automatic						Automatic				
		BP	PS	FS	Accuracy			BP	PS	FS	Accuracy
Manual	BP	1971	117	6	94.3%	Manual	BP	857	54	3	93.8%
	PS	241	1059	54	78.2%		PS	105	428	37	75.1%
	FS	6	92	208	68.0%		FS	3	33	89	71.2%
	Overall Accuracy = 86.4%						Overall Accuracy = 85.4%				

Legend: Bare – BP; Partly Snow Covered – PS; Fully Snow Covered – FS

4.5.1.2 Model Comparison

In the previous section, three models were described with respect to their parameters and classification results. To select the best performing model, a comparison must be made regarding their ability to perform the intended tasks (i.e., improving RSC classification described in Chapter 3) using a simulated connected vehicle system. Performance is evaluated from two main points of view: classification accuracy and model features. Overall classification accuracy as well as classification of individual RSC types will be compared across individual models. Additional model features such as prediction error and computation time are also evaluated as part of the model selection process. These features are compared below by applying the ANN, RT and RF models on an identical sample of 30% holdout data, in a similar manner to the spot-wise RSC monitoring accuracy evaluated in Chapter 3.

Classification Performance

One of the main objectives in this chapter is to improve the RSC classification accuracy of the AVL-Genius system. Table 4.12 shows the AVL-Genius classification results for the holdout data. For the holdout data, AVL-Genius yielded an overall hit rate/classification accuracy of 67%. Bare, partly and fully snow covered conditions were automatically classified with accuracies of 80.6%, 51.6% and 40%, respectively. Low hit rates for the classification of partly and fully snow covered surfaces compromise the reliability of the system to provide maintenance personnel and the travelling public with accurate RSC information. Therefore, any alternative models would be considered successful if they could increase the accuracies observed in Table 4.12

Table 4.12 – AVL-Genius Confusion Matrix for Holdout Data

		Automatic			
		BP	PS	FS	Accuracy
Manual	BP	747	156	24	80.6%
	PS	226	285	41	51.6%
	FS	7	71	52	40.0%
	Overall Accuracy = 67%				

BP - Bare Pavement; PS – Partly Snow Covered; FS – Fully Snow Covered

Figure 4.11 shows a comparison of individual and overall RSC classification between AVL-Genius, ANN, RT and RF models. Overall, the three models were successful in improving the RSC classification results of AVL-Genius. An approximately 13% increase in classification accuracy for bare surfaces was observed when any of the three models was applied. RT and RF models resulted in identical classification accuracy of 93.8% for bare surfaces. The ANN model resulted in marginally higher accuracy of 94%, showing the most improvement for this RSC type. While the original AVL-Genius classification was already high for bare surfaces, this improvement is still considered to be important, as bare pavement conditions are arguably the most important RSC type for winter road maintenance management.

An important feature of the bare pavement classification is the effect of a false positive, the incorrect prediction of a particular (poor) condition (e.g., snow covered) as an acceptable (good) condition (e.g., bare pavement). Therefore, for bare conditions, a false positive occurs when the model incorrectly classifies a snowy surface as bare. False positive rates for bare surfaces were approximately 15% for all three model types, which may be considered high for image recognition in transportation engineering applications (Belaroussi et al., 2010). The implication of a high false positive rate for bare conditions is a compromise in safety, since some partly or fully snow covered condition are classified as bare pavement. Conversely, there is less safety risk associated with high false positive values of other RSC types; the potential negative effect could be a waste of maintenance resources due to false responses to poor conditions predicted.

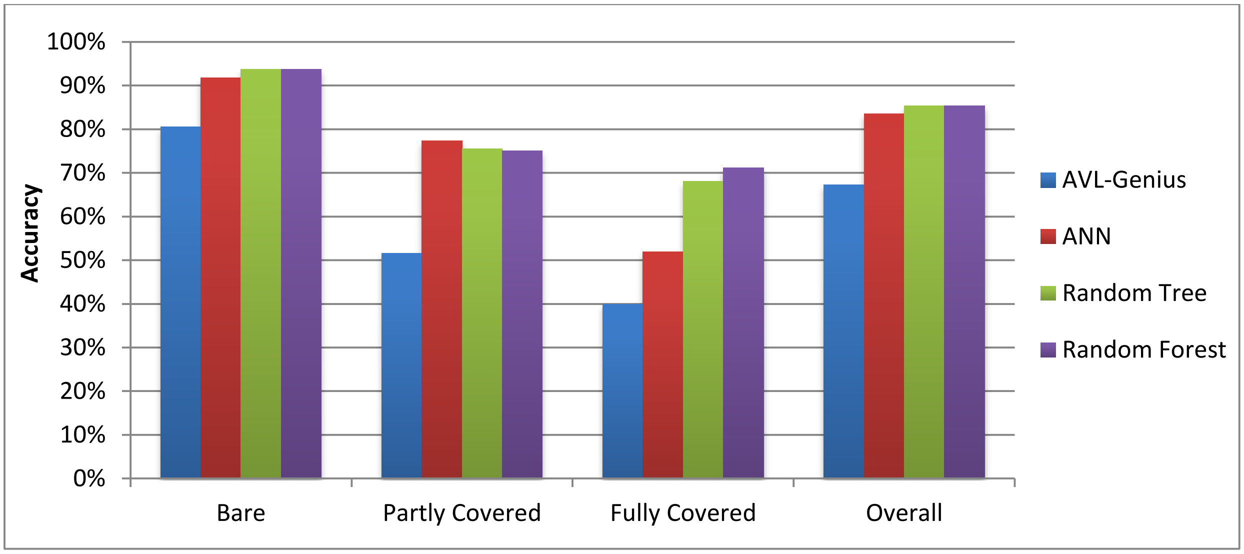


Figure 4.11 – Original Classification Accuracy vs. Proposed Connected Vehicle Models (Three-Class System)

Classification of partly snow covered conditions increased from the original accuracy of 51.6% with AVL-Genius to 74%, 75.6% and 75.1% with ANN, RT and RF models respectively. Although performance was similar across all models, RT demonstrated the highest classification accuracy for this RSC-type with the lowest false positive rate at 8.5% - resulting in increased reliability. Although there is a lower safety risk associated with false positive classification of partly snow covered conditions, there is a risk of resource wastage since maintenance equipment would typically be dispatched in order to regain bare pavement.

Similar to the original AVL-Genius classifications, fully snow covered detection was the least accurate of the three RSC classes. The ANN model increases the classification accuracy of fully snow covered surfaces by approximately 33%, resulting in classification accuracy of 72.8% for this class. Although the ANN model outperformed its counterparts, the RF model accurately classified these conditions at a marginally lower rate of 71.2%, with the RT model demonstrating the lowest accuracy at 68%. False positive rates for this category were observed to be lower than 3% for all three models. Classification of fully snow covered conditions was an initial concern as noted in Chapter 3 due to the relatively low accuracy demonstrated by AVL-Genius' image-based analysis. The proposed models therefore significantly increase the performance and reliability of the RSC monitoring system by increasing classification accuracy by up to 33%.

Model Characteristics

For any RSC classification system, accurate identification of winter road conditions is the primary concern for the end user. However, since the technology is intended to be a real time RSC monitoring tool, model features such as computation time and prediction error also play an important role. Computation time refers to the length of time required by a model to provide RSC classifications. For the holdout data, the processing times of the ANN, RT and RF models were recorded as 13.27s, 0.02s and 0.1s respectively on a computer with a 2.9GHz Intel Core i7 processor and 8 GB 1600 MHz DDR3 memory. The ANN model required a significantly longer time to provide improved RSC classifications than that required by the decision trees. Again, in considering model selection for any real-time system, computation time could be considered the most important feature after result accuracy. As a result, classification trees appear more suitable for this real-time classification task since they are significantly faster than neural networks for RSC estimation as outlined in this thesis. Figure 4.12 shows additional model features typically considered during a model selection process. RF model demonstrates the highest Kappa statistic at 0.729, however the ANN and RT models demonstrate only marginally lower values at 0.727 each.

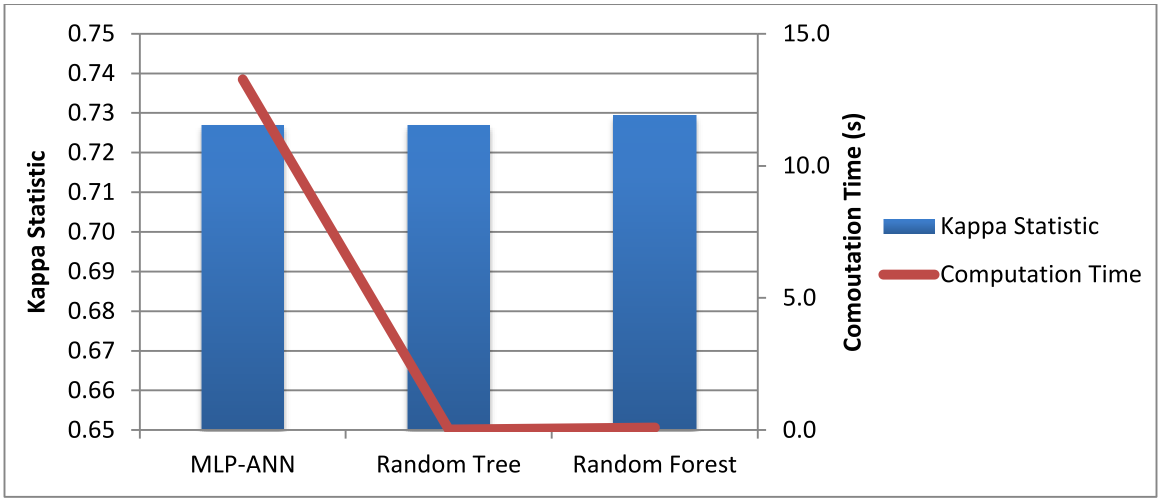


Figure 4.12 – Comparison of Kappa Statistic for Connected Vehicle Models (Three-Class System)

4.5.2 Five-Class Road Surface Condition Classification

One of the main goals of RSC monitoring for road authorities is to obtain accurate and detailed information about winter road conditions. The previous section describes RSC monitoring using the default three-class system of AVL-Genius. As mentioned, this level of detail is sufficient for motorist decision-making but may be insufficient to make informed maintenance decisions since maintenance operations can differ depending on the level of snow coverage on a highway. For instance, the three-class system categorizes conditions as either bare, partly or fully snow covered. With this three-class system, center-bare, both wheel tracks bare and one wheel track bare are all described as partly snow covered despite having very different safety and maintenance implications. The following section details the model calibration and validation process used to provide more detailed RSC descriptions. Conditions will be assigned into one of five classes of snow coverage (Bare, <25, 25 to 50, 50 to 75, fully snow covered) as described in Table 3.3.

4.5.2.1 Artificial Neural Network Models

The ANN model was calibrated with an input layer that includes the original AVL-Genius output (bare, partly snow covered, fully snow covered) and the data received from the RWIS stations (relative humidity, precipitation intensity, 1hr, 3hr, 6hr, 12hr and 24hr accumulation, salinity, pavement surface status, air temperature, pavement temperature). Multiple ANN configurations were evaluated for and Figure 4.13 illustrates the optimal calibrated ANN model with a learning rate of 0.1, momentum of 0.2 for the backpropagation algorithm and one hidden layer with 13 hidden neurons.

Table 4.13 shows the results of the ANN model calibration. The Kappa statistic was recorded as 0.57, lower than the equivalent model of the three-class output. This value is considered to indicate moderate agreement between observed and predicted RSC classifications. A lower agreement was generally expected since there are two additional output categories in this model. As with any black-box classifier, it is difficult to interpret the model and therefore variable interaction remains unexplained in the ANN model.

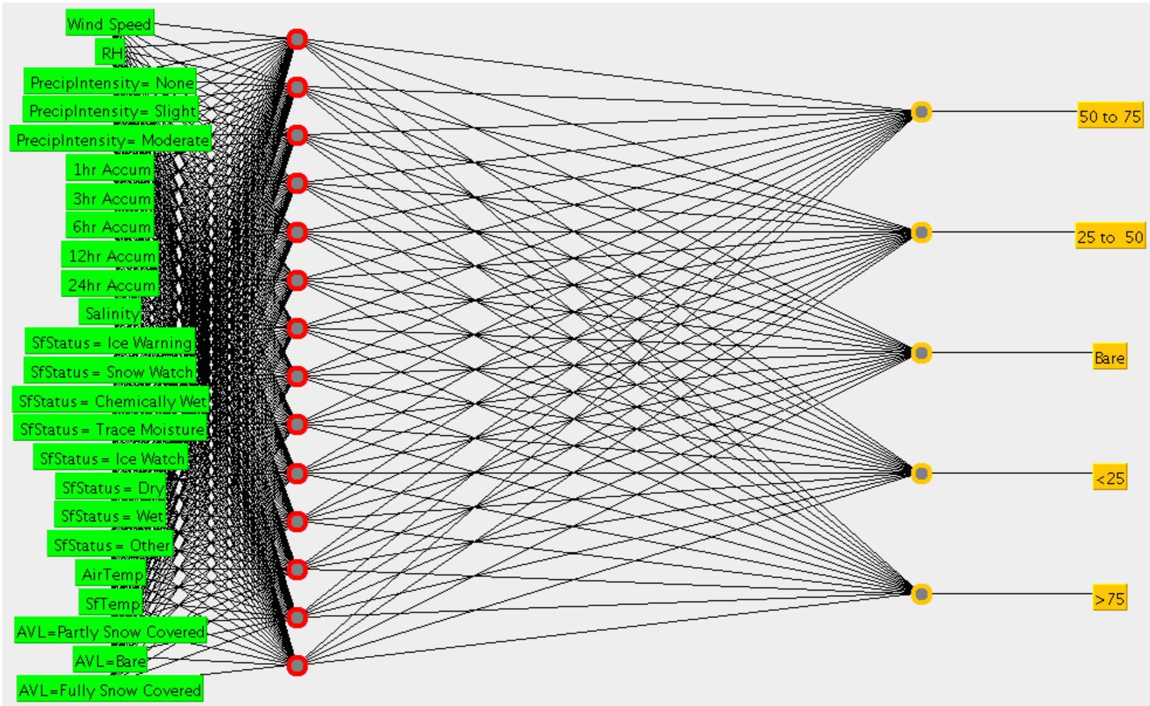


Figure 4.13 – Neural Network RSC Classification Model (Five-Class System)

Table 4.13 - Neural Network Characteristics (Five-Class System)

Neural Network Model Calibration Summary	
Correctly Classified Instances	2791 (74.3%)
Incorrectly Classified Instances	963 (25.7%)
Kappa Statistic	0.567
Total Number of Instances	3754

Table 4.14 - Neural Network Calibration Characteristics (Five-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.965	0.23	0.839	0.898	0.936
<25	0.099	0.019	0.347	0.155	0.737
25 to 50	0.174	0.03	0.366	0.236	0.822
50 to 75	0.761	0.115	0.598	0.7	0.914
Fully Snow Covered	0.545	0.017	0.732	0.625	0.934
Weighted Avg.	0.743	0.154	0.699	0.707	0.904

Table 4.15 shows the classification results of the calibration and validation of the ANN model. The overall accuracy was found to be 74% and 76% for the calibration and validation data, respectively. Classification accuracy of individual RSC types ranged from 9% to 97%, indicating difficulty in distinguishing particular RSC types with the proposed model. Bare images were classified with an accuracy of 96.5%, higher than when bare images were classified using the three-class system. The "<25" RSC class was poorly classified by the model with a low classification accuracy of 9.9%. However this class is visually similar to bare conditions, and as a result the model misclassified the majority of the "<25" class as bare.

Fully snow covered and "50 to 75" (single wheel track bare) surfaces were classified with accuracies of 54.5% and 76.1%, respectively. These two RSC types can be similar in appearance, especially when pavement contaminants were primarily brown slush. This observation makes sense since over 90% of fully snow covered surfaces were classified as either "50 to 75" or fully covered. A similar trend was observed with AVL-Genius' image-based classifications.

Surfaces with "25 to 50" snow coverage were classified by the model with 11.5% accuracy. The model was observed to overestimate the snow coverage for the "25 to 50" conditions since most of the misclassifications were categorized as "50 to 75".

Table 4.15 - Neural Network Calibration and Validation Results (Five-Class System)

		Bare	<25	25 to 50	50 to 75	FS	Accuracy
Calibration Data							
Manual	Bare	2008	28	14	29	2	96.5%
	<25	243	34	23	40	2	9.9%
	25 to 50	77	24	59	171	8	17.4%
	50 to 75	47	9	61	526	48	76.1%
	FS	17	3	4	113	164	54.5%
	Overall Accuracy = 74%						
Validation Data							
		Bare	<25	25 to 50	50 to 75	FS	Accuracy
Manual	Bare	891	7	3	26	0	96.1%
	<25	90	15	8	21	0	11.2%
	25 to 50	34	10	13	83	1	9.2%
	50 to 75	10	6	9	219	33	79.1%
	FS	12	1	1	37	79	60.8%
	Overall Accuracy = 76%						

4.5.2.2 Random Tree Model

A Random Tree (RT) model was calibrated in order to obtain more detailed RSC classifications using the five-class scheme. Table 4.16 shows the results of the RT calibration. The RT model was found to have a Kappa statistic of 0.63, indicating good agreement between actual and predicted RSC classes. Since the RT is considered to be a white-box classifier, we are able to understand variable interaction and assess intuitive effects.

At the root of the tree is salinity with a threshold value of 5305 parts per 100,000 (53,050 ppm). Salinity at the root makes intuitive sense because it was established in exploratory analysis that increased salinity generally indicates an increase in RSC severity. This is likely due to deicing chemicals being applied to the road surface according to the level of snow coverage observed by maintenance personnel.

Surface status and AVL-Genius classification were the next closest variables to the root. During exploratory analysis, it was observed that each RWIS pavement surface status type showed particular trends in association with other variables. For instance, a surface was more likely to be partly or fully snow covered with higher levels of salinity (>5000 parts per 100,000) accompanied by a pavement status of “Snow Watch”. Knowledge of the pavement condition at the RWIS station (if pavement is wet or snow covered) can therefore improve the decision making process by indirectly identifying the presence of tracks. AVL-Genius image classification then follows with regards to level of importance. The five-class RSC classification model appears to more heavily rely on RWIS data since the AVL-Genius classification input is frequently located near leaves of the tree. This makes intuitive sense and tends to mimic the thought process of first assessing the present environmental conditions in order to create a likely RSC scenario and then considering the results of automatic image classification to result in a detailed RSC result.

Table 4.16 – Random Tree Model Characteristics (Five-Class System)

Model Calibration Summary	
Correctly Classified Instances	2894 (77%)
Incorrectly Classified Instances	860 (23%)
Kappa Statistic	0.6275
Total Number of Instances	3754

Table 4.17- Random Tree Calibration Characteristics (Five-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.943	0.159	0.881	0.911	0.941
<25	0.278	0.031	0.473	0.35	0.757
25 to 50	0.31	0.043	0.415	0.355	0.803
50 to 75	0.755	0.09	0.655	0.702	0.907
Fully Snow Covered	0.694	0.019	0.763	0.727	0.919
Weighted Avg.	0.771	0.113	0.75	0.756	0.904

Table 4.18 shows classification results of calibrated data and 30% holdout validation data for the RT model. Overall hit rates for calibration and validation data were 77% and 74%, respectively. Calibration data demonstrated hit rates for bare, <25, ‘25 to 50’, ‘50 to 75’ and fully snow covered surfaces at 94.3%, 27.8%, 31.0%, 75.5%, 69.4% respectively. Validation data hit rates were very similar, indicating good consistency in model performance for the two data sets. RSC classes <25 and

25 to 50 again demonstrate the lowest hit rates, underscoring the difficulty in accurately predicting these classes with this proposed model structure.

Table 4.18 - Random Tree Calibration and Validation Results (Five-Class System)

Calibration Data							
		Bare	<25	25 to 50	50 to 75	FS	Accuracy
Manual	Bare	1963	63	16	31	8	94.3%
	<25	172	95	38	32	5	27.8%
	25 to 50	48	27	105	149	10	31.0%
	50 to 75	32	11	84	522	42	75.5%
	FS	14	5	10	63	209	69.4%
	Overall Accuracy = 77.1%						
Validation Data							
		Bare	<25	25 to 50	50 to 75	FS	Accuracy
Manual	Bare	875	30	9	12	1	94.4%
	<25	64	40	18	12	0	29.9%
	25 to 50	20	14	47	60	0	33.3%
	50 to 75	8	7	24	208	30	75.1%
	FS	4	1	3	38	84	64.6%
	Overall Accuracy = 77.9%						

4.5.2.3 Random Forest

A random forest (RF) comprised of 10 trees was modeled in order to provide detailed RSC classifications. Each tree was constructed while considering 4 random features. Table 4.19 shows calibration results of the RF model. A Kappa statistic of 0.63 indicates good agreement strength between the model prediction and expected output. As stated in the previous section, RFs are considered to be black-box models, making interpretation and assessment of variable importance problematic.

Table 4.19 - Random Forest model characteristics (Five-Class System)

Model Calibration Summary	
Correctly Classified Instances	2890(77%)
Incorrectly Classified Instances	864 (23%)
Kappa Statistic	0.6273
Total Number of Instances	3754

Table 4.20 – Random Forest Calibration Characteristics (Five-Class System)

Detailed Accuracy By Class					
Class	**TP Rate**	**FP Rate**	**Precision**	**F-Measure**	**ROC Area**
Bare	0.938	0.158	0.88	0.908	0.951
<25	0.322	0.04	0.449	0.375	0.784
25 to 50	0.313	0.042	0.422	0.359	0.823
50 to 75	0.737	0.08	0.676	0.705	0.927
Fully Snow Covered	0.708	0.022	0.74	0.723	0.931
Weighted Avg.	0.77	0.112	0.751	0.758	0.918

Table 4.21 shows classification results of calibrated and validated data for the RF detailed RSC prediction model. The model was found to have an overall classification accuracy of 77% and 74% for calibration and validation data respectively. Bare, <25, “25 to 50”, “50 to 75” and fully snow covered surfaces were classified with accuracies of 93.8%, 32.2%, 31.3%, 73.7% and 70.8% respectively for calibration data. For validation data, results were very similar, except for “50 to 75” and fully snow covered conditions, where a 3% decrease in accuracy was observed.

Table 4.21- Random Forest Calibration and Validation Results

Calibration Data							
		Bare	**<25**	**25 to 50**	**50 to 75**	**FS**	**Accuracy**
Manual	**Bare**	1952	84	18	18	9	93.8%
	<25	171	110	31	25	5	32.2%
	25 to 50	46	32	106	141	14	31.3%
	50 to 75	34	13	88	509	47	73.7%
	FS	14	6	8	60	213	70.8%
	Overall Accuracy = 77%						
Validation Data							
		Bare	**<25**	**25 to 50**	**50 to 75**	**FS**	**Accuracy**
Manual	**Bare**	870	37	10	7	3	93.9%
	<25	63	43	18	10	0	32.1%
	25 to 50	25	18	46	52	0	32.6%
	50 to 75	11	9	28	196	33	70.8%
	FS	5	1	5	31	88	67.7%
	Overall Accuracy = 74%						

4.5.3 Model Comparison

The previous section described the calibrated models and their statistical characteristics. Using the holdout data (30% for validation), these models are compared with respect to their ability to use RWIS data combined with smartphone images to classify RSCs into one of five classes.

Classification Performance

Figure 4.14 shows the comparison between the proposed ANN, RT and RF connected vehicle models for classifying winter RSCs. Bare surfaces were classified well with accuracies between 94% and 97% for all three models, with the ANN model demonstrating the highest classification accuracy. However, ANN model demonstrated a false positive rate of 22% for this class - approximately 7% higher than observed in the RT and RF models.

The RSC class of <25 was classified with an accuracy of 11.2% by the ANN model. The decision trees classified this RSC type with significantly higher accuracy, with RT and RF models demonstrating hit rates of 29.9% and 32.1%, respectively. In all three models, a significant portion of images with <25 snow coverage were misclassified as bare. In practice, it would not be uncommon for <25 to be described as "essentially bare" and if for classification purposes bare and <25 were combined into a single class, hit rates would increase tremendously, up to 96%.

For the "25 to 50" RSC class, performance was generally poor among the three models. Hit rates for the ANN, RT and RF models were 9.2%, 33.3% and 32.6%, respectively, demonstrating the comparatively superior performance of the classification trees. Although the classification accuracy of this RSC type was too low for practical purposes, the classification trees demonstrated the ability to increase classification accuracy by four times of that by the ANN.

Classification of the "50 to 75" RSC type showed a relatively high accuracy for all models compared to the remaining classes. ANN, RT and RF models classified "50 to 75" coverage with hit rates of 79.1%, 75.1% and 70.8%, respectively. The ANN model clearly outperformed the decision trees, but it also demonstrated the highest false positive rate of 12.5%, which is 5% higher than the RF model.

Fully snow covered surfaces were classified with moderate accuracy compared to those with the three-class RSC scheme. Hit rates for the ANN, RT and RF models were 60.5%, 64.6% and 67.7%, respectively. For this RSC type, decision trees again outperformed the neural network; however across all models, fully snow covered classification accuracy was lower than that using the three-class system.

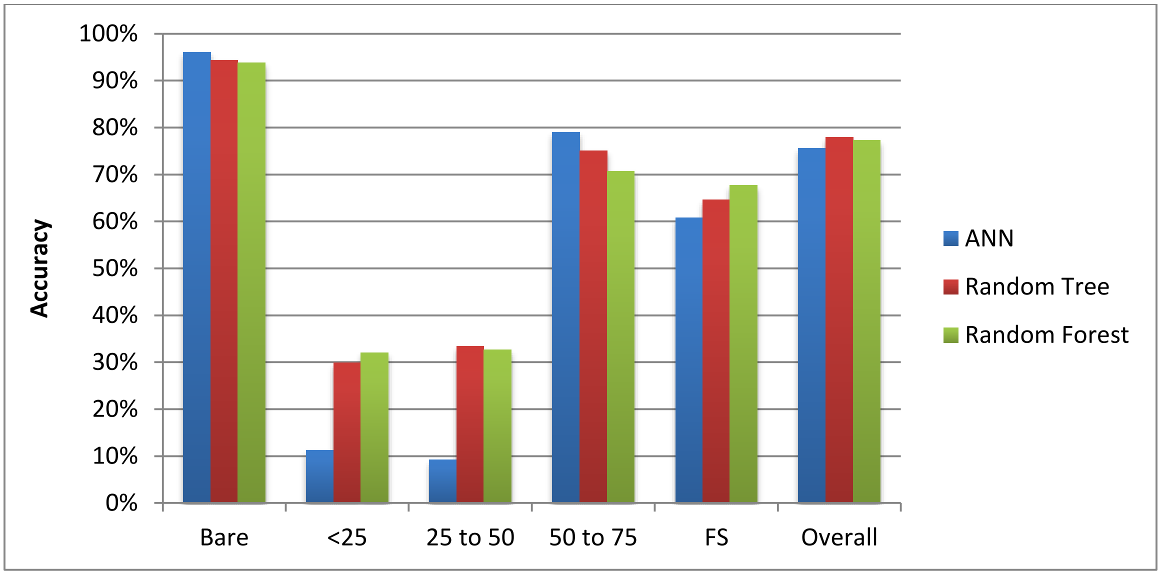


Figure 4.14 - Classification Accuracy of Proposed Connected Vehicle Models (Five-Class System)

Model Characteristics

In addition to classification accuracy, there are other model features to be considered when deciding on the most feasible RSC estimation solution. The ANN model showed a Kappa statistic of approximately 0.57, which is considered lower than ideal when assessing model performance. Moreover processing time, of high importance to a real-time classification system, was recorded at 15.35s. Both decision trees performed the same classifying task in less than 1s, with the random tree and random forest demonstrating processing times of 0.02s and 0.11s respectively as illustrated in Figure 4.15. This substantial difference in classification time as well as inferior performance compared to the decision trees makes the ANN model the least preferred for the proposed real time winter RSC classification using a connected vehicle structure. The two decision trees show similarities in classification accuracy and statistical characteristics. Specifically, the RT model demonstrated marginally better overall classification accuracy, lower false positive rates and significantly shorter processing time than all competing models. However, there is still a need for improvement when classifying “<25” and “25 to 50” classes.

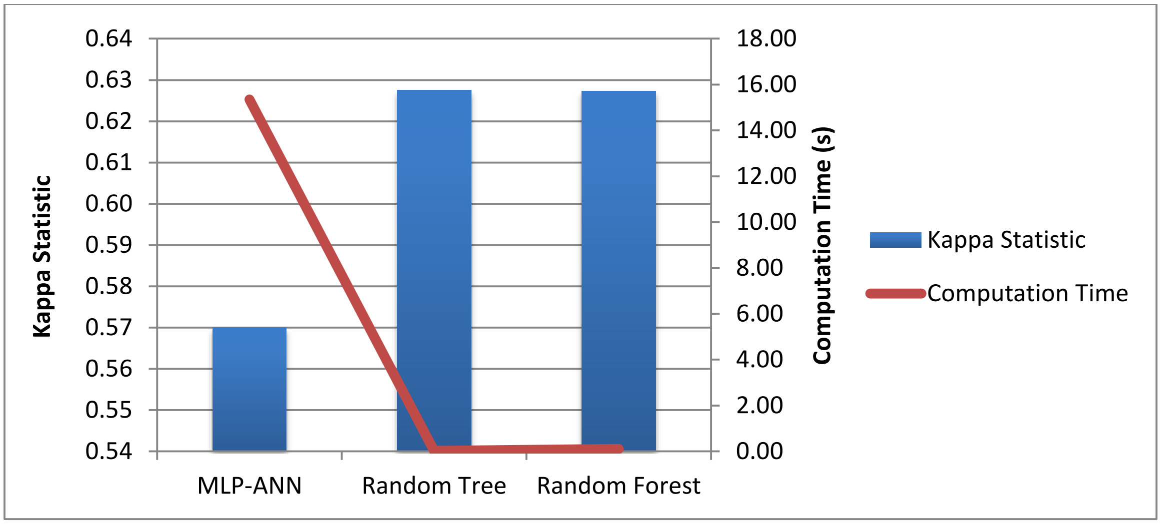


Figure 4.15 - Comparison of Kappa Statistic for Connected Vehicle Models (Five-Class System)

4.6 Transferability

Prior to the adoption of any data driven model, it is important to understand whether it can be generalized across all settings. The underlying models for the proposed connected vehicle RSC monitoring system depend heavily on RWIS data; therefore, it is important to validate their transferability across different locations. In this section, the selected models are evaluated with respect to their transferability (i.e., can the connected vehicle models be applied to new locations and perform similar to when they were initially calibrated?). This transferability assessment was determined based on data captured on a different test route during a subsequent winter of 2015. The following sections describe the test site, collected data and transferability assessment of the connected vehicle RSC monitoring solution.

4.6.1 Test Site

Field Tests were carried out in the winter of 2015 on a section of a two-lane, two-way Class 2 highway – Hwy 23 between Minto and Monkton, Ontario as shown in Figure 4.16. The test section is approximately 43 km long with a winter average daily traffic (WADT) volume of 5300 (Ontario Ministry of Transport, 2010). The site has uniform geometrical features, very few horizontal curves and a combination of open and sheltered agricultural fields and woodlots. The area also experiences an annual average of 59 days with snowfall of at least 0.2cm (Environment Canada, 2014). An Area Maintenance Contractor (AMC), Integrated Maintenance and Operations Services (IMOS) maintains the route and typical WRM activities include plowing, sanding and salting. One RWIS station (SW-26) is located on the route near Gowanstown and provides road and weather data for the study route.

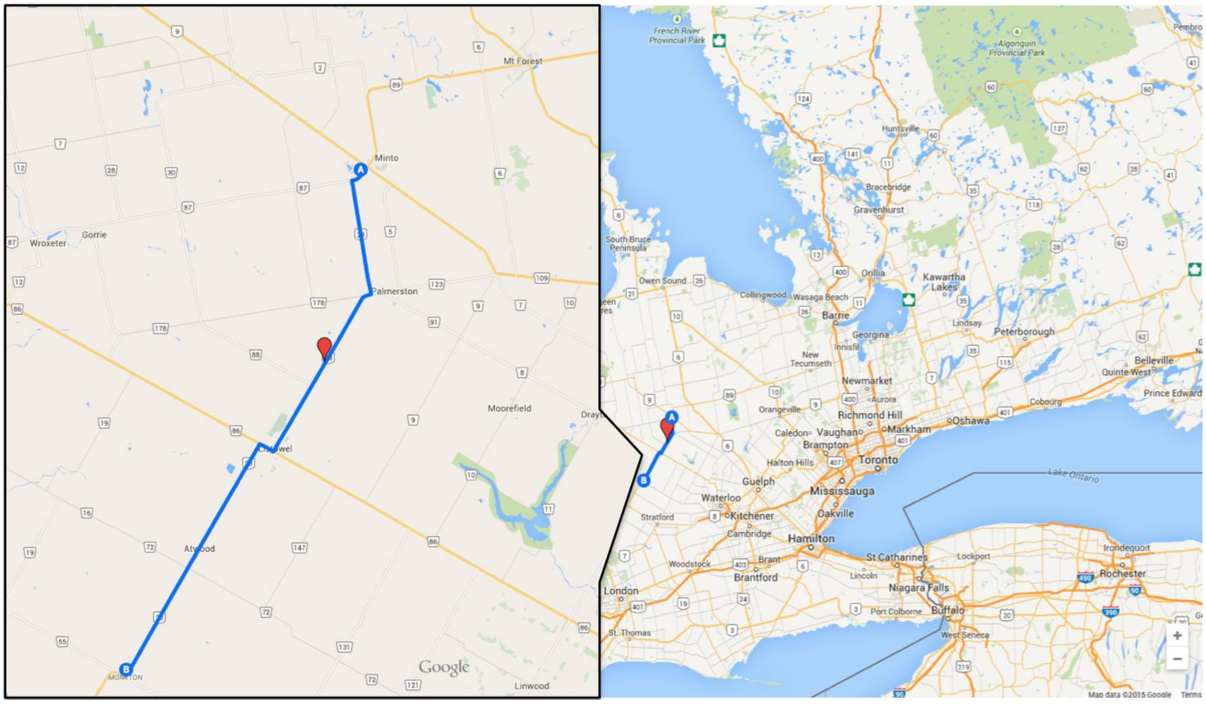


Figure 4.16 – Transferability Test Site and RWIS Stations

4.6.2 Data Collection Summary

In order to evaluate model transferability, over 1500 images were collected during six events between Feb 5, 2015 and Feb 21, 2015 on the test route. The events consisted of snowfall, drifting snow or a combination of both. Figure 4.17 and Table 4.22 show the summary statistics of the field tests and associated event characteristics.

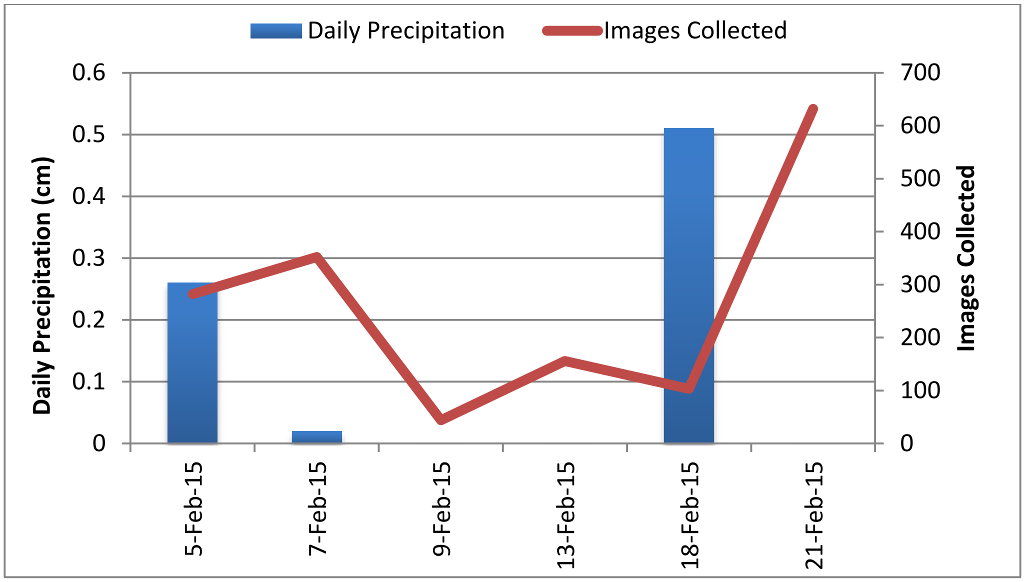


Figure 4.17 – Summary of Events and Images Collected for Transferability Test

Table 4.22 – Summary of Event Attributes and Road Surface Images Captured

	Min	Max	Mean	Std. Dev.
Events	6			
Total Precipitation (cm)	0	0.51	0.07	0.13
Pavement Temperature (°C)	-16.5	3.5	-7.2	5.52
Air Temperature (°C)	-19	-4	-11.6	4.15
Wind Speed (km/h)	2	23	9.5	4.63
No. Of Camera Images	44	632	262	214

4.6.3 Road Weather Information System Data

The RWIS data on weather conditions, as described in Section 4.3.2, was collected at a 20-minute frequency from station SW-26. Table 4.23 and Table 4.24 show the summary of variable features for a total of 1569 samples extracted from RWIS station SW-26.

Table 4.23 – Summary of Statistics of Variables Used for Model Transferability Test

		Number of Observations =1569			
Field Name	**Unit**	**Min**	**Max**	**Mean**	**SD**
Surface Temperature	°C	-16.5	3.5	-7.2	5.52
Air Temperature	°C	-19	-4	-11.6	4.15
1hr Accumulation	cm	0	0.18	0.01	0.03
3hr Accumulation	cm	0	0.48	0.02	0.09
6hr Accumulation	cm	0	0.51	0.03	0.10
12hr Accumulation	cm	0	0.51	0.06	0.12
24hr Accumulation	cm	0	0.51	0.07	0.13
Wind Speed	km/h	2	23	9.5	4.63
RH	%	54	98	86.06	9.48
Salinity	parts/100,000	0	33470	15734.90	9629.22

Table 4.24 – Transferability Data Categorical Variable Sample Size

Field Name	**Categories**	**Size**	**%**
Surface Status	Chemically Wet	352	22.4%
	Dry	86	5.5%
	Ice Warning	604	38.5%
	Snow Watch	2	0.1%
	Trace Moisture	481	30.7%
	Wet	44	2.8%
Precipitation Intensity	None	1544	98.4%
	Slight	25	1.6%

4.6.4 Evaluation

In the previous section random tree and random forest models were calibrated for RSC monitoring using three-class and five-class system, respectively. In order to evaluate the transferability of these models, they are applied to a dataset comprised of 1569 observations captured during the Winter 2015 season at a completely different test site (SW-26). These observations are separated into two segments: 70% (1098) for training and 30% (471) for testing. The purpose of this separation is to compare the prediction performance of the models calibrated in the previous section (SW-25 & SW-

13) with that of a model calibrated for the specific dataset and location (SW-26). This ensures a fair evaluation of the transferability of the previously calibrated models.

Table 4.25 and Figure 4.18 show the comparison of prediction accuracy for the random tree model calibrated from the previous dataset and location as well as the model calibrated from 70% holdout. The transferred model calibrated for the RWIS stations SW-25 and SW-13 in the previous section, demonstrated overall classification accuracy of 52.4% - significantly lower than that observed in the previous section and 14% lower than the smartphone-based system accuracy. Bare, partly and fully snow covered surfaces were classified with 44%, 68% and 13% accuracy, respectively. Moreover, false positive rates for bare classification were found to be 23.3%. The transferred model classified bare and fully snow covered surfaces 42% and 23% accuracy - more poorly than the AVL-Genius image-based classification. However, the transferred model classified partly snow covered surfaces with 34% higher accuracy compared to AVL-Genius.

The locally calibrated Random Tree Model demonstrated an overall classification accuracy of 82.6%. Bare, partly and fully snow covered surfaces were classified with 96%, 85.7% and 68.4% accuracy, respectively. The false positive rate for classification of bare surfaces was found to be 1.7% - significantly lower than observed in the previous section. The locally calibrated model also resulted in substantial increases in classification accuracy of individual RSC types compared to AVL-Genius' image-based classification algorithm.

The significant difference in classification accuracy between the two models (28% overall), indicates poor transferability with the proposed connected vehicle models. Classification of bare and fully snow covered surfaces experience the greatest discrepancy in performance between the two models with 52% and 55% difference respectively. When calibrated for their respective locations, the models classified each RSC type with similar performance accuracies, indicating the strength of the classification models; however, they also demonstrate poor transferability. Moreover, the transferred model classified bare and fully snow covered surfaces much more poorly when compared to the AVL-Genius smartphone system, possibly because different model criteria exist for different geographical locations, particularly for areas susceptible to microclimatic data. Variation in maintenance practices and traffic behaviour across highways may also contribute to poor transferability, since salinity thresholds for RSC discrimination appear to differ (difference in residual salt due to initial application and subsequent traffic dispersion). Overall it appears that, due to poor transferability, models would have to be calibrated for individual or closely grouped RWIS stations along highway routes for reliable and accurate RSC data. This would result in each zone-of-influence having dedicated RSC connected vehicle models.

Table 4.25- Transferability Assessment Results

	Transferred Model						Local Model				
	Automatic						Automatic				
		BP	PS	FS	Accuracy			BP	PS	FS	Accuracy
Manual	BP	22	28	0	44.0%	Manual	BP	48	2	0	96.0%
	PS	88	210	9	68.4%		PS	7	263	37	85.7%
	FS	10	89	15	13.2%		FS	0	36	78	68.4%
	Overall Accuracy = 52.4%						Overall Accuracy = 82.5%				

Legend: Bare – BP; Partly Snow Covered – PS; Fully Snow Covered – FS

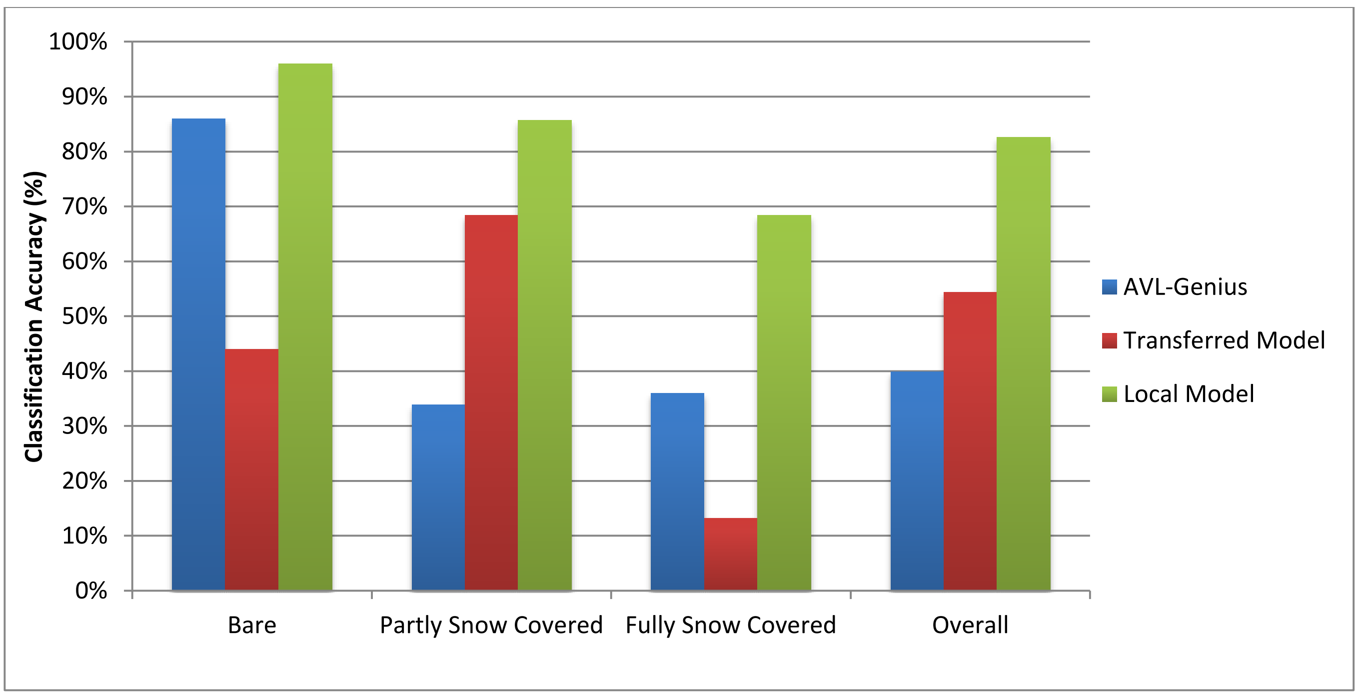


Figure 4.18 – Model Transferability Comparison

4.7 Summary

This chapter described a modeling effort used to combine image classifications with RWIS data in order to simulate a connected vehicle road surface classification (RSC) system to address the shortcomings of the AVL-Genius RSC classification system noted in Chapter 3.

Firstly, a series of artificial neural network (ANN) and classification trees were calibrated in order to directly improve the AVL-Genius' three-class spot-wise classification accuracies for bare, partly and fully snow covered conditions. Secondly, a series of decision tree and ANN models were calibrated in order to provide more granular results using a five-class RSC description system.

The models tested were multilayer ANNs, random trees and random forests. In all modeling efforts, classification trees were found to outperform ANN, according to overall classification accuracy, processing time and prediction errors. Random forests were found to improve the overall classification accuracy observed in Chapter 3 by 18% and improve the classification accuracy of fully snow covered conditions by 30%, using the three-class system. Random trees were found to provide more granular RSC classifications with overall accuracy of 78% using the five-class system. Conditions essentially bare (<25) and both wheel tracks bare (25 to 50) were classified more poorly with approximately 30% accuracy. Model transferability was tested with data obtained from a new test site in a subsequent winter season and it was found that the proposed models are not easily transferable, suggesting that the proposed models need to be trained with site-specific data.

The results of the simulation conducted in this chapter lead to three main conclusions:

1) The combination of road and weather data with road surface image processing results in a more accurate RSC monitoring tool than the solution based on image recognition alone.
2) A connected vehicle system has the potential to be a timely, reliable, RSC monitoring tool that provides more detailed RSC information at high spatial resolution to winter maintenance personnel.
3) The proposed models are not transferable; therefore for practical implementation they need to be trained with local data.

Chapter 5
Conclusions and Future Work

This research has been motivated by the need by transportation agencies to provide reliable winter RSC information to the travelling public and winter maintenance personnel. The thesis research includes two main components. The first component focuses on evaluating the performance of a smartphone-based automated road surface condition (RSC) monitoring system called AVL-Genius. AVL-Genius integrates an image recognition system for classifying RSC. The second effort focuses on the development and evaluation of a connected vehicle based solution, which was proposed to address the limitations of AVL-Genius. Both efforts made use of real world data collected from a series of field tests conducted along a 70km stretch of Highway 6 near Owen Sound, Ontario and a 43km stretch of Highway 23 near Gowanstown, Ontario during the winter seasons of 2013/2014 and 2014/2015. This section summarizes the main findings of each research effort and provides recommendations for future research.

5.1 MAJOR FINDINGS

5.1.1 Evaluation of a Smartphone-based RSC Monitoring System

A smartphone system (AVL-Genius) that classifies road images into three categories including bare, partly or fully snow covered was evaluated during the winter of 2014, with a total of 15,913 images collected during 23 test runs covering a variety of weather conditions. Results from the smartphone system were compared to manual classifications, patrol reports and MTO's Traveller's Road Information Portal (TRIP) system. The main findings from the field test are summarized below:

- AVL-Genius was first evaluated for its spot-wise RSC monitoring accuracy. AVL-Genius RSC classifications of individual images were compared to the "ground-truth" from manual classifications of the same images. It was found that the system achieved an overall classification accuracy of 72% matching for the conditions detected. Detailed analyses found the primary causes of mismatched classifications were poor visibility, glare from sunlight and dried residual salt on the road surface.
- Route-level conditions generated by AVL-Genius were found to be highly consistent with those reported by field personnel via routine patrol reports. The main difference between these two RSC monitoring methods was that AVL-Genius provided quantitative details on individual RSC

types occurring along a route as opposed to the descriptive nature of the patrol reports. By providing information about the proportion of RSC types observed along a maintenance route, maintenance operators can make more informed decisions since they can see the dominant conditions and the extent to which RSCs change during winter events and maintenance operations.

- Compared to the TRIP system, AVL-Genius provides more timely and spatially detailed RSC information along a maintenance route. RSC information of high temporal and spatial resolution is invaluable to highway agencies, maintenance personnel and the travelling public. Transport agencies can use this information to improve performance monitoring while maintenance personnel can deliver safe winter roads with better-targeted treatments in order to reduce operating costs and salt use and to improve level of service.

- The field study also identified key issues and areas of improvement for the current version of the smartphone-based system. The system needs to be extended in order to include night-time monitoring. This is especially important for practical consideration.

- AVL-Genius currently classifies RSCs into three major classes: bare, partly snow covered or fully snow covered. Although this may be sufficient for visualization and public reporting, maintenance personnel typically require more detailed information about level of snow coverage and contaminant type in order to make informed maintenance decisions. Snow coverage level could refer to wheel paths bare or to centre or lane bare; contaminant type could refer to packed snow or slush. The algorithm could generally be improved for better classification accuracy.

5.1.2 Connected Vehicle Based RSC Monitoring System

A connected vehicle RSC monitoring system was proposed as an attempt to address the shortcomings of the previously evaluated smartphone-based system. The system is proposed to make use of the classification results of an image based system as well as data from nearby RWIS stations through connected vehicle technology. New machine learning models were developed and their performance was evaluated and compared using field data. The main findings from this research are summarized below:

- A comparison between neural networks, random trees and random forests showed that random trees were the best models for a connected vehicle RSC monitoring system that categorizes winter roads as bare, partly snow covered or fully snow covered. Further comparisons showed the random forest model improved classifications from the smartphone-based system by 18% overall – with individual improvements of 13%, 24% and 28% for classification of bare, partly and fully snow covered surfaces, respectively.

- Models were calibrated and validated in order to provide more detailed classifications for use by winter maintenance personnel. RSC types were bare, center of lane bare (<25), both wheel tracks bare (25 to 50), single wheel track bare (50 to 75) and fully snow covered. Neural networks, random trees and random forests were developed and compared; however, random trees were found to provide the highest classification accuracy, with an overall matching of 78%. The Random trees model classified bare, <25, (25 to 50), (50 to 75) and fully snow covered conditions with accuracies of 94%, 30%, 33%, 71% and 65% matching, respectively.

5.2 RECOMMENDATIONS

5.2.1 Smartphone-based System

- A more extensive pilot study should be carried out to evaluate the robustness and reliability of AVL-Genius. In particular, its real-time RSC monitoring capability should be evaluated in order to assess its ability to function effectively (i.e., evaluating the time between data collection and obtaining system results). Further research efforts should also test data communication between AVL-Genius and the TRIP system in order to evaluate the system's ability to reach its intended audience (travelling public and maintenance personnel).

- In order to be considered a practical RSC monitoring tool, AVL-Genius must be further developed to include night-time classifications. Following the inclusion of this feature, similar evaluations should be conducted to assess its performance during the night-time. Moreover, an important future effort would be to investigate the optimal patrolling frequency required to obtain sufficient spatial and temporal resolution for decision-making.

- The smartphone-based platform allows for cost effective crowdsourcing due to its scalability. This can result in the public receiving RSC information with denser and more extensive coverage of the road network. This potential should be explored during future efforts and should include several different types of vehicles such as commercial or surface transit.

5.2.2 Connected Vehicle System

- A more extensive field investigation that includes different locations and climates should be conducted. Further exploration should include additional variables such as proximity and direction to RWIS station. Future modelling efforts should also consider simultaneously processing extracted image features with road and weather information, instead of conducting image classification separately for input.
- Future research efforts should investigate the optimal system performance within the RWIS zone-of-influence (i.e., comparing the extent to which RWIS station proximity affects classification accuracy). It was observed that model transferability was poor, therefore future investigations should examine whether a modifier variable could improve model transferability.
- Decision trees and neural networks were the primary models evaluated for the connected vehicle RSC monitoring solution. Future research should investigate other model types and configurations in order to maximize system performance. Moreover, alternate system structures should be investigated in order to maximize classification accuracy for each RSC type (i.e., investigate the most appropriate models for classifying particular RSC types).
- In this research, a connected vehicles system via vehicle-to-infrastructure has been hypothesized and tested. Future research should consider the inclusion of maintenance vehicle data. Maintenance vehicles are already equipped with GPS technology and mechanical sensors that indicate operation type and material deposited. A system comprised of both RWIS data and maintenance vehicle data may provide maximum accuracy and details for RSC monitoring as proposed.
- Future studies should focus on developing a connected vehicle RSC monitoring prototype as described in this research. In the field tests reported in this study, images were captured via a smartphone and paired with RWIS road and weather data for model calibration and testing. While this represents a simulation of the data obtained through the described connected vehicle system, future efforts should aim at testing the system as proposed using front facing vehicle cameras and dedicated short range communications to connect to RWIS stations. This testing would provide the opportunity to address any challenges that may arise during the implementation of the connected vehicle system, including its ability to provide RSC information in real-time.

Successful implementation of such a system could prove important to both human drivers and autonomous vehicles.

5.3 Concluding Remarks

This research represents an important step towards reliable road surface condition monitoring. By leveraging the latest technologies, the development of a connected vehicle RSC monitoring tool has the potential to contribute to safer and more environmentally friendly winter roads. At the writing of this thesis, this work represents the first known connected vehicle model for winter RSC monitoring. Moreover, this research can serve as a quantitative guide and point of reference for Intelligent Transportation Systems (ITS) and connected vehicle RSC monitoring solutions, as the transportation industry moves towards a more technology-driven future.

Appendix A

Patrol Reports

Winter Patrol Records - MTO

Ontario **Winter Patrol Record** Yard Location: AMC oversite Page # ______ of ______

Date: Mon. Feb. 24/14

Patrolled by: S. Arnold

Shift Start: Time 13:54 Mileage 119314

Shift End: Time 15:50 Mileage

Hwy. Closure Information

Hwy	Time Closed	Time Opened

Patrol Route

Highway	Time	From	To	Time	Air Temperature	Pavement Temperature	Clear	Partly Cloudy	Overcast	Rain (L, M, H)	Snow (L, M, H)	Freezing Rain	Fog (L, M, H)	Visibility (G, F, P)	Wind (L, M, S)	Wind Direction	Bare and Dry	Bare and Wet	Track Bare	Center Bare	Snow Covered	Snow Packed	Drifted Sections	Icy Sections	Frost	Slushy	Patrolling	Plowing	Sanding	Salting	Snow Blowing (hwy)
6	13:55	Owen Sound	Mt. Forest	14:45	-10				✓					G	M	W		✓	✓		✓		✓								
6	14:55	Mt. Forest	Owen Sound	15:45	-10				✓					G	M	W		✓	✓		✓		✓					✓			

Weather Conditions: Clear through Wind Direction. Highway Conditions: Bare and Dry through Slushy. Operations: Patrolling through Snow Blowing (hwy).

Comments:

* Viaesys Web-Cam Recording

Owen Sound – Mt. Forest

- Some Drifted Sections

- No Operations Noticed.

Mt. Forest → O.S

Plow SBL 1 Km N/Mt. Forest

- Plow Down 15:00

Plow SBL N/Dornoch

- Plow Down 15:24

- Plow S/Chatsworth SBL 15:34

- Drifting Causing Some Snow Covered Sections

SA

Patroller's Signature:

Visibility reported as **G**ood, **F**air or **P**oor Rain, Snow or Fog reported as **L**ight, **M**oderate or **H**eavy Wind reported as **L**ight, **M**oderate or **S**trong

Winter Patrol Records – IMOS

OWEN SOUND HIGHWAY MAINTENANCE LIMITED

Winter Patrol Record

Yard Location: CHATSWORTH 24

Page # 1 of 2

Date: MAR. 4/2014

Patrolled by: TIM DRAPER

Shift Start: 7:00

Shift End: 19:00

Hwy. Closure Information

Hwy	Time Closed	Time Opened

Patrol Route

Highway	Time	From	To	Time	Air Temperature	Pavement Temperature	Clear	Partly Cloudy	Overcast	Rain (L, M, H)	Snow (L, M, H)	Freezing Rain	Fog (L, M, H)	Visibility (G, F, P)	Wind (L, M, S)	Wind Direction	Bare and Dry	Bare and Wet	Track Bare	Center Bare	Snow Covered	Snow Packed	Drifted Sections	Icy Sections	Frost	Slushy	Plowing	Sanding	Salting	Snow Blowing (hwy)
#6	8:00	DURHAM	MOUNT FOREST	8:25	-15	-14			✓		L			G	L	SW	✓	✓											SPOT ✓	
#6	8:25	MOUNT FOREST	DURHAM	8:50	-14	-13			✓		L			G	L	SW	✓	✓											SPOT ✓	
	9:05	DURHAM.	FLESHERTON	9:30																										
#10	9:30	FLESHERTON	DUNDALK	9:45	-14	-14			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#10	9:45	DUNDALK	CHATSWORTH	10:30	-13	-10			✓		L			G-F	L	S-W		✓	✓	✓	✓								✓	
#6	10:45	CHATSWORTH	OWEN SOUND	11:00	-12	-9			✓		L			G-F	L	S-W		✓	✓	✓	✓								✓	
#26	11:00	OWEN SOUND	MEAFORD	11:30	-11	-10			✓		L			G-F	L	S-W		✓	✓	✓	✓								✓	
#26	12:10	MEAFORD	COLLINGWOOD	12:50	-8	-7			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#26	12:50	COLLINGWOOD	MEAFORD	13:15	-9	-6			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#26	13:20	MEAFORD	OWEN SOUND	13:45	-11	-9			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#6	13:45	OWEN SOUND	CHATSWORTH	14:10	-11	-10			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#6	14:45	CHATSWORTH	OWENSOUND	15:00	-9	-8			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#26	15:00	OWEN SOUND	IRISH BLOCK Rd. HWY 26	15:30	-8	-9			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#26	15:30	IRISH BLOCK Rd AND HWY 26	OWEN SOUND	15:50	-9	-10			✓		L			G-F	L	S-W		✓	✓		✓								✓	
#6	15:50	OWEN SOUND	CHATSWORTH	16:05	-9	-10			✓		L			G-F	L	S-W		✓	✓		✓								✓	

Comments: SNOW BLOWER WORKING ON HWY. #26 FROM OWEN SOUND TO MEAFORD

Visibility reported as Good, Fair or Poor Rain, Snow or Fog reported as Light, Moderate or Heavy Wind reported as Light, Moderate or Strong

Winter Patrol Records – Road and Weather Information Sheet

Ontario

Ministry of Transportation
London Traffic Operations Centre
659 Exeter Road
London, Ontario N6E 1L3
Telephone: (519) 873-4223
Facsimile: (519) 873-4443

Ministère des Transports
Le Centre des Opérations de London
659, chemin Exeter
London, Ontario N6E 1L3
Téléphone: (519) 873-4223
Télécopieur: (519) 873-4443

Report Information

Report Time: 03:00 09:00 13:00 15:00 21:00

Special Report Time: ____________

Date: SAT, MAR 22/14

Operator: L'COLLINS

Road and Weather Information Sheet

NORTH

Time Entered into RCS: ________

Precipitation Conditions

- **AN** - No Precipitation
- **Rain** - Rain [L-M-H]
- **Fr Rain** - Freezing Rain [L-M-H]
- **Snow** - Snow [L-M-H]
- **Ice Pel** - Ice Pellets [L-M-H]
- **Mix** - Mix (Rain/Snow) [L-M-H]

Atmospheric Conditions

- **Temp** - Air Temperature [Reported in °Celcius]
- **Wind** - Wind Direction [8 point compass]
- **Speed** - Wind Speed [Light - Moderate - Strong]
- **Vis** - Visibility [Good - Fair - Poor]
- **Sky** - Cloud Condition [Clear - Partly Cloudy - Overcast]
- **Fog** - Fog [L-M-H]
- **Drift** - Drifting [Y / N]

Road Conditions

- **RB** - Roads Bare and Dry
- **WET** - Roads Bare and Wet
- **PSC** - Partly Snow Covered
- **RSC** - Snow Covered
- **PSP** - Partly Snow Packed
- **RSP** - Snow Packed
- **PIC** - Partly Ice Covered
- **RI** - Roads Ice Covered

Maintenance Operations

- **PAT** - Patrolling
- **OP** - Plowing
- **OSA** - Sanding
- **OSG** - Salting
- **DLA** - Anti-Icing
- **CU** - Clean Up

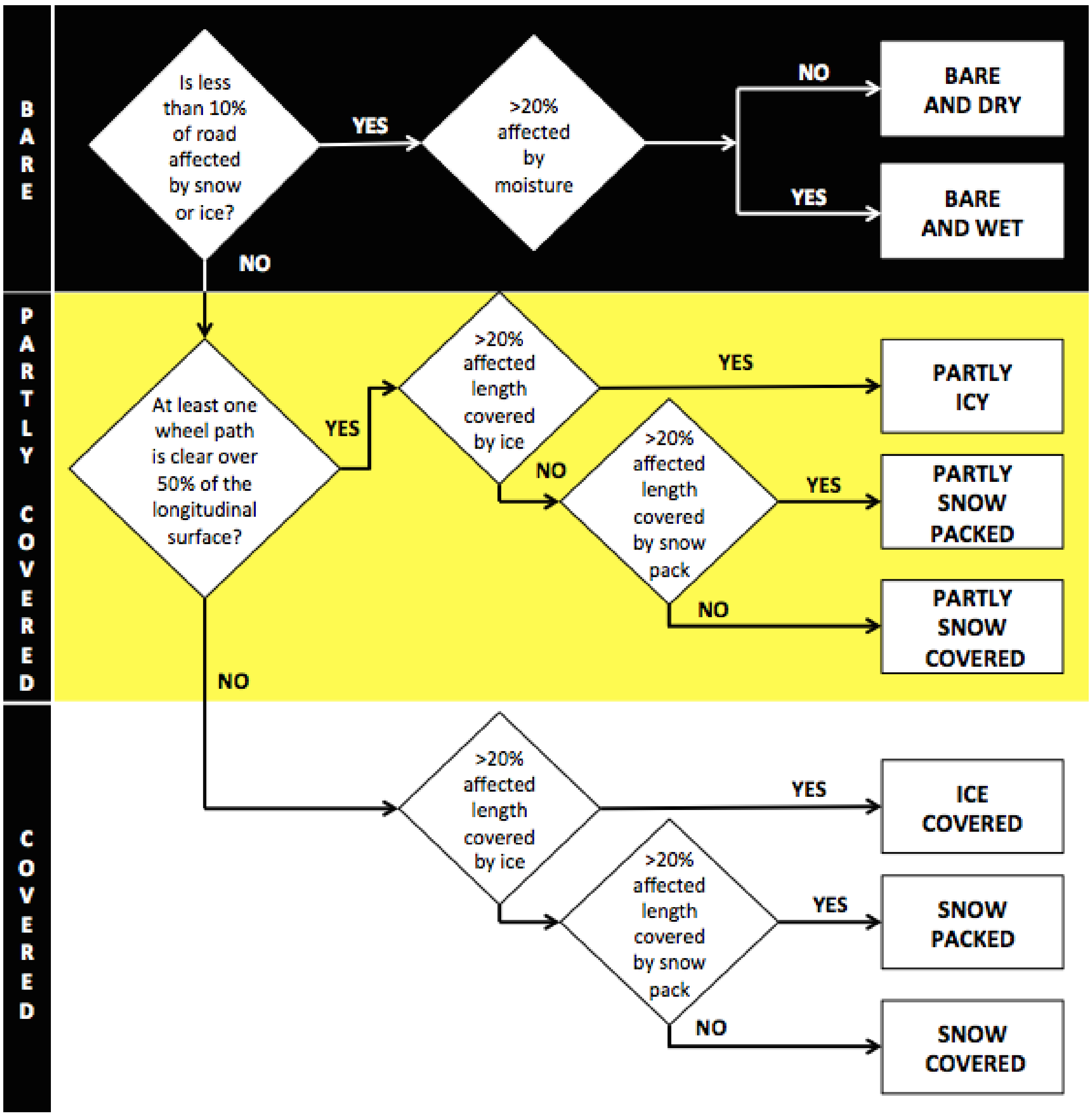
Patrol / Highways: Patrol Name	Patrol #	HWYS	Precipitation Conditions: AN No Precip	Rain L-M-H	Fr Rain L-M-H	Snow L-M-H	Ice Pel L-M-H	Mix L-M-H	Atmospheric Conditions: Temp °C	Wind Direction	Wind Speed	Visibility G-F-P	Sky Clear	Sky Partly Cloudy	Sky Overcast	Fog L-M-H	Drifting Y / N
Mitchell East	21	23	✓						-1	NW	L	G			✓		Y
		7															Y
		8															N
Mitchell West	22	21	✓						0	NW	L	G		✓			N
		4															
		8															
Shelburne	23	6	✓						-1	N	L	G			✓		Y
(Orangeville to Shelburne)	23A	10&10/89															
(Shelburne to Dundalk)	23B	10															N
	23	89															
Chatsworth	24	10	✓						-2	NW	L-M	G			✓		Y
		26															N
		6 & 6/10															
Kincardine	25	21	✓						0	W	L	G			✓		N
		9															
Clavering	26	21	✓						-1	NW	L-M	G		✓			
		6															

Patrol Name	Patrol #	HWYS	Road Conditions: Roads Bare RB Bare/Dry	Roads Bare WET Bare/Wet	Snow Covered PSC Partly	Snow Covered RSC Covered	Snow Packed PSP Partly	Snow Packed RSP Packed	Ice PIC Partly	Ice RI Covered	Maintenance Operations: PAT Patrolling	OP Plowing	OSA Sanding	OSG Salting	DLA Anti-icing	CU Clean Up
Mitchell East	21	23	✓	✓	✓	✓					✓	✓		✓		
		7	✓	✓	✓						✓	✓		✓		
		8	✓								✓					
Mitchell West	22	21	✓								✓					✓
		4														
		8														
Shelburne	23	6		✓	✓						✓	✓		✓		
(Orangeville to Shelburne)	23A	10&10/89									✓					
(Shelburne to Dundalk)	23B	10	✓								✓					
	23	89		✓	✓						✓	✓		✓		
Chatsworth	24	10			✓						✓	✓		✓		
		26	✓								✓					
		6 & 6/10			✓						✓	✓		✓		
Kincardine	25	21	✓	✓							✓					✓
		9									✓					
Clavering	26	21	✓	✓							✓					
		6														

Page 3 of 3

Appendix B

TAC RSC Classification Scheme



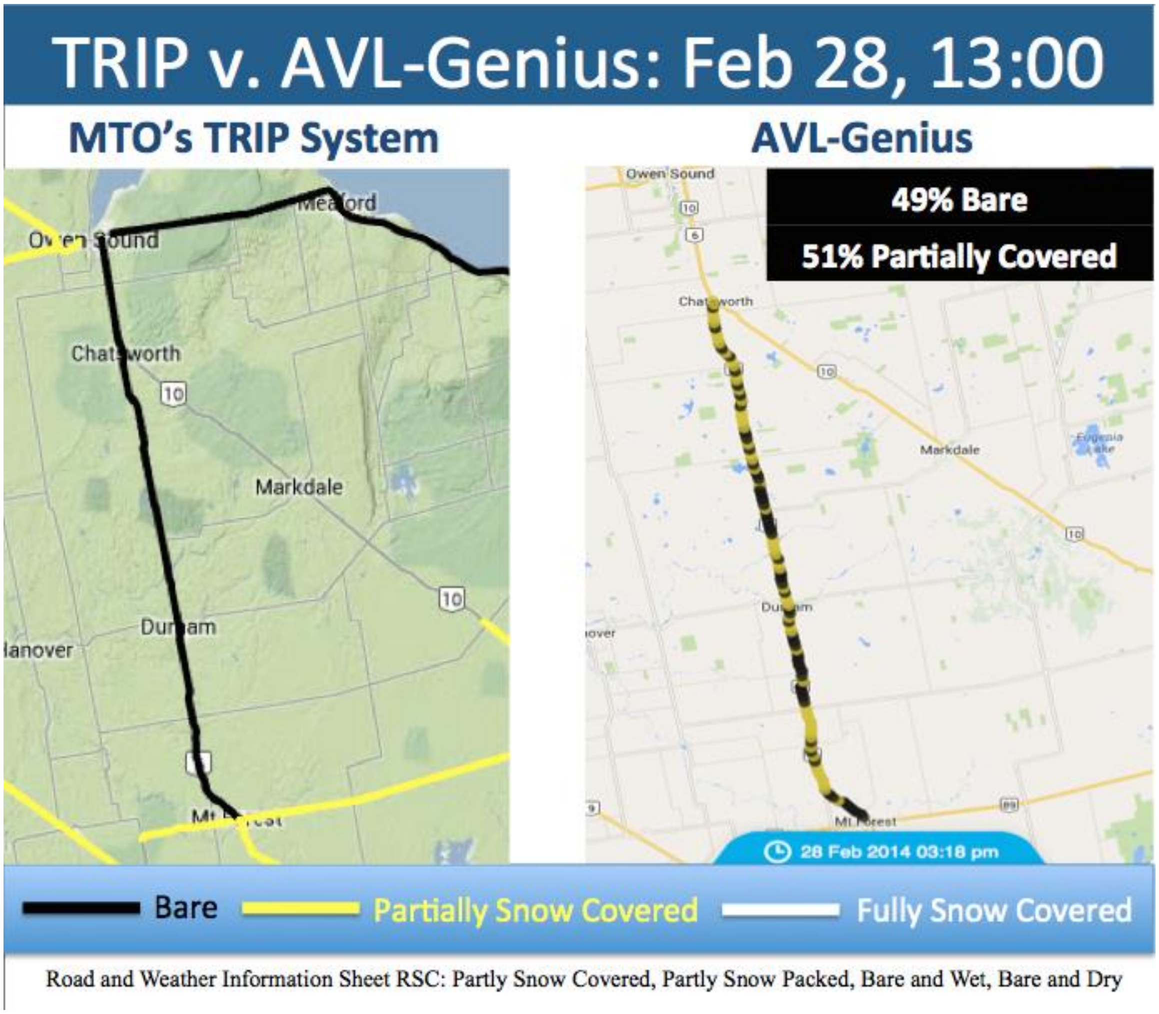
Appendix C

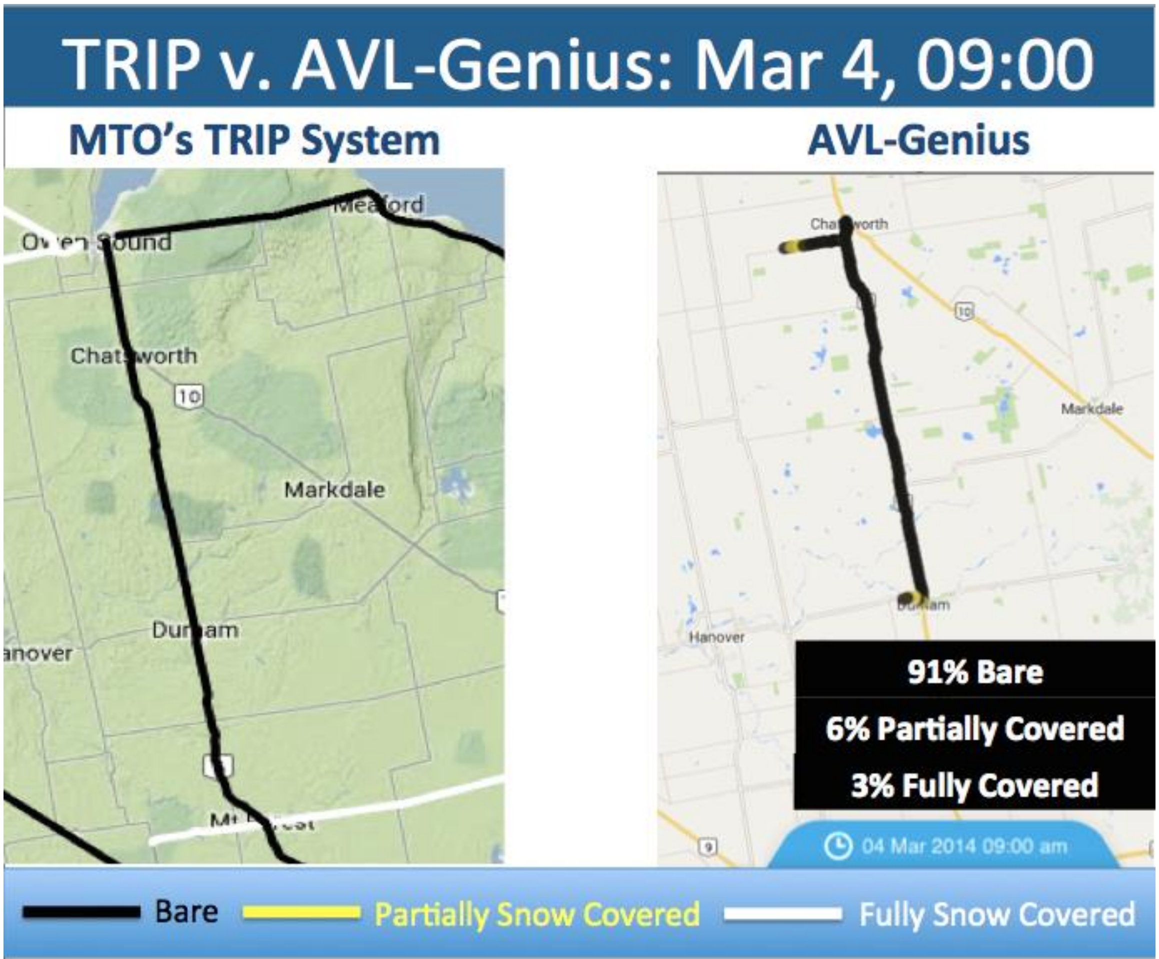
Image Classification Summary

Event Date	Matching Images	Non-Matching Images	Classification Accuracy
24-Feb-14	1001	381	72.4%
25-Feb-14	414	136	75.3%
26-Feb-14	293	249	54.1%
27-Feb-14	213	70	75.3%
28-Feb-14	532	471	53.0%
01-Mar-14	479	180	72.7%
04-Mar-14	847	359	70.2%
05-Mar-14	613	194	76.0%
10-Mar-14	499	44	91.9%
12-Mar-14	439	349	55.7%
13-Mar-14	267	236	53.1%
14-Mar-14	457	209	68.6%
15-Mar-14	686	490	58.3%
16-Mar-14	42	58	42.0%
19-Mar-14	183	84	68.5%
20-Mar-14	519	261	66.5%
21-Mar-14	547	115	82.6%
22-Mar-14	622	100	86.1%
23-Mar-14	477	69	87.4%
25-Mar-14	468	78	85.7%
26-Mar-14	371	44	89.4%
27-Mar-14	978	47	95.4%
02-Apr-14	532	210	71.7%
Total	**11479**	**4434**	**72.1%**

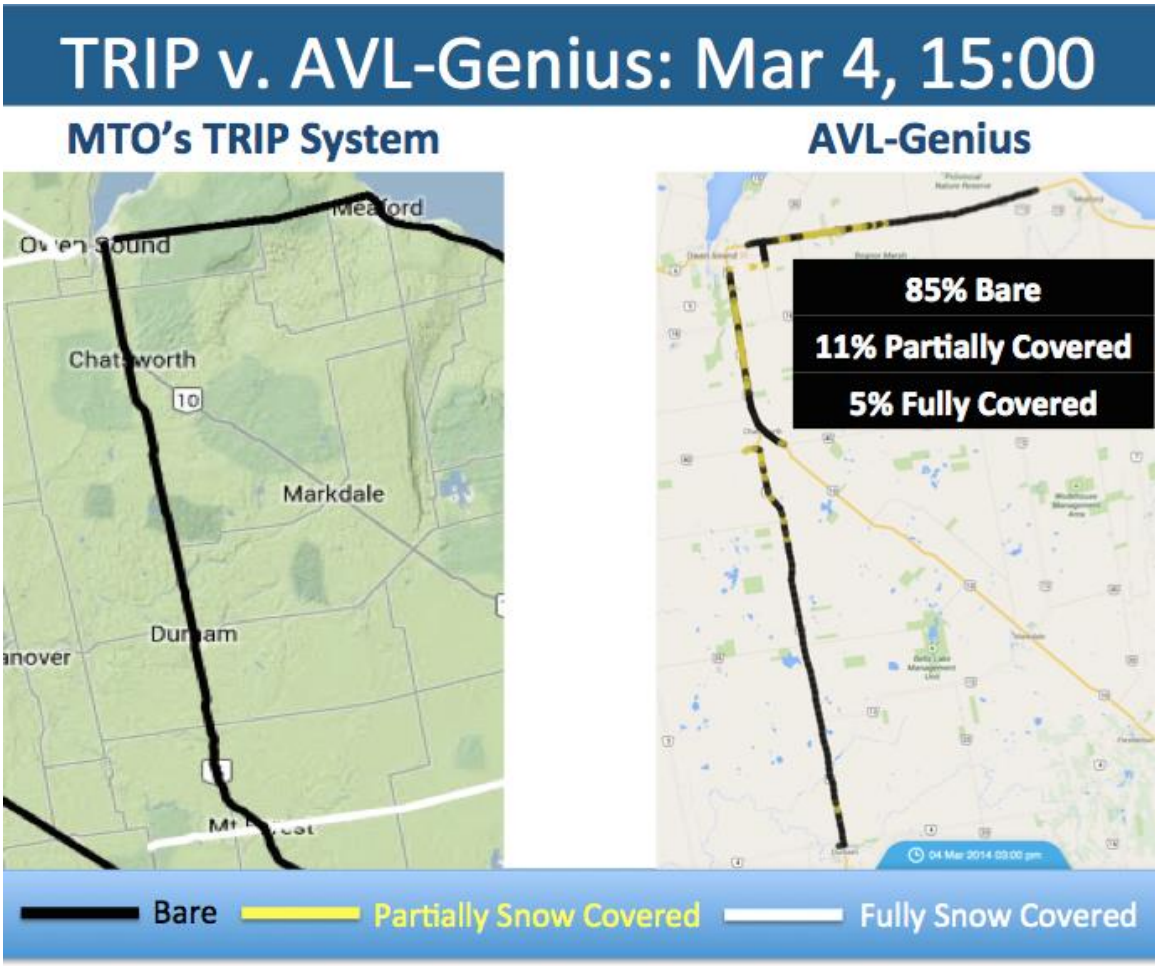
Appendix D

Comparison between AVL-Genius and TRIP System

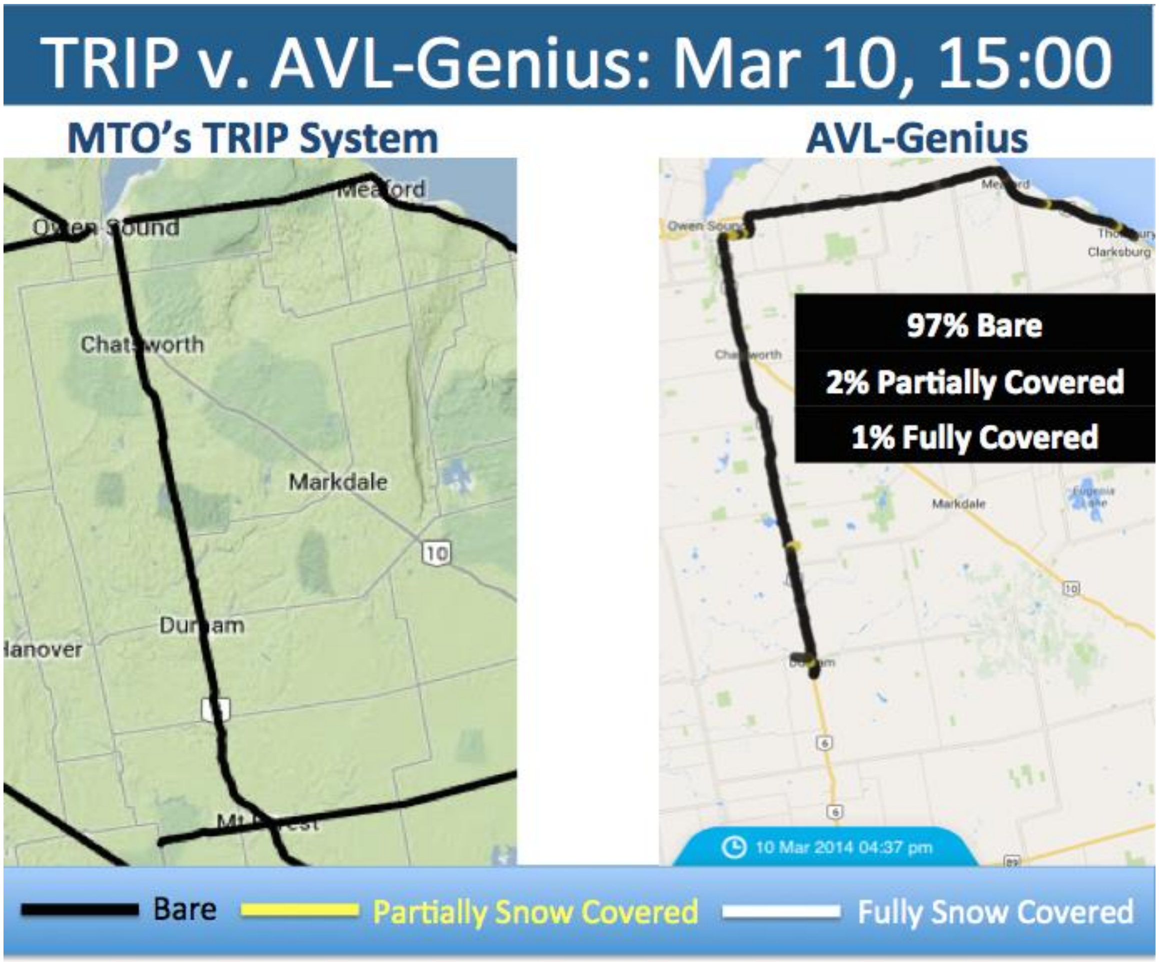




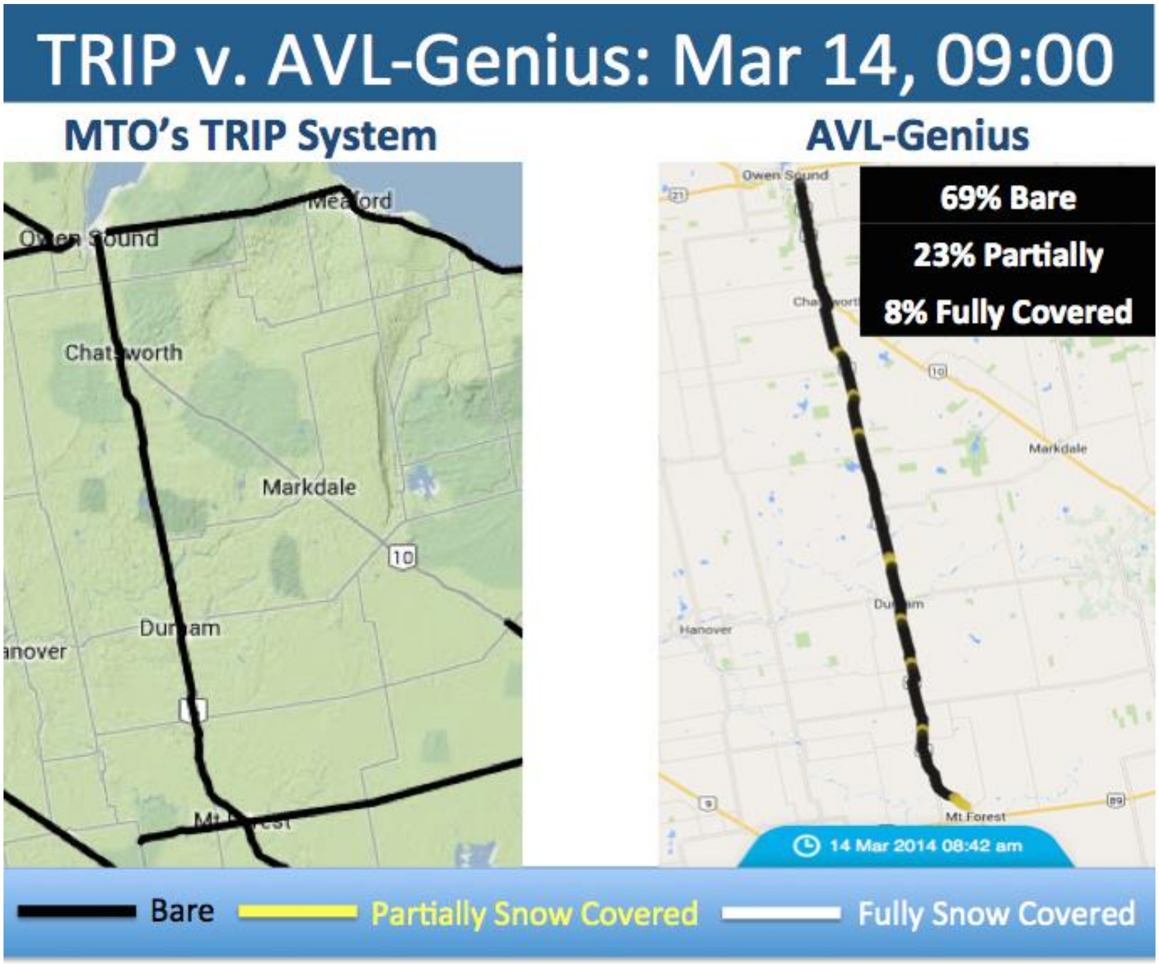
Road and Weather Information Sheet RSC: Partly Snow Covered, Bare and Wet, Bare and Dry



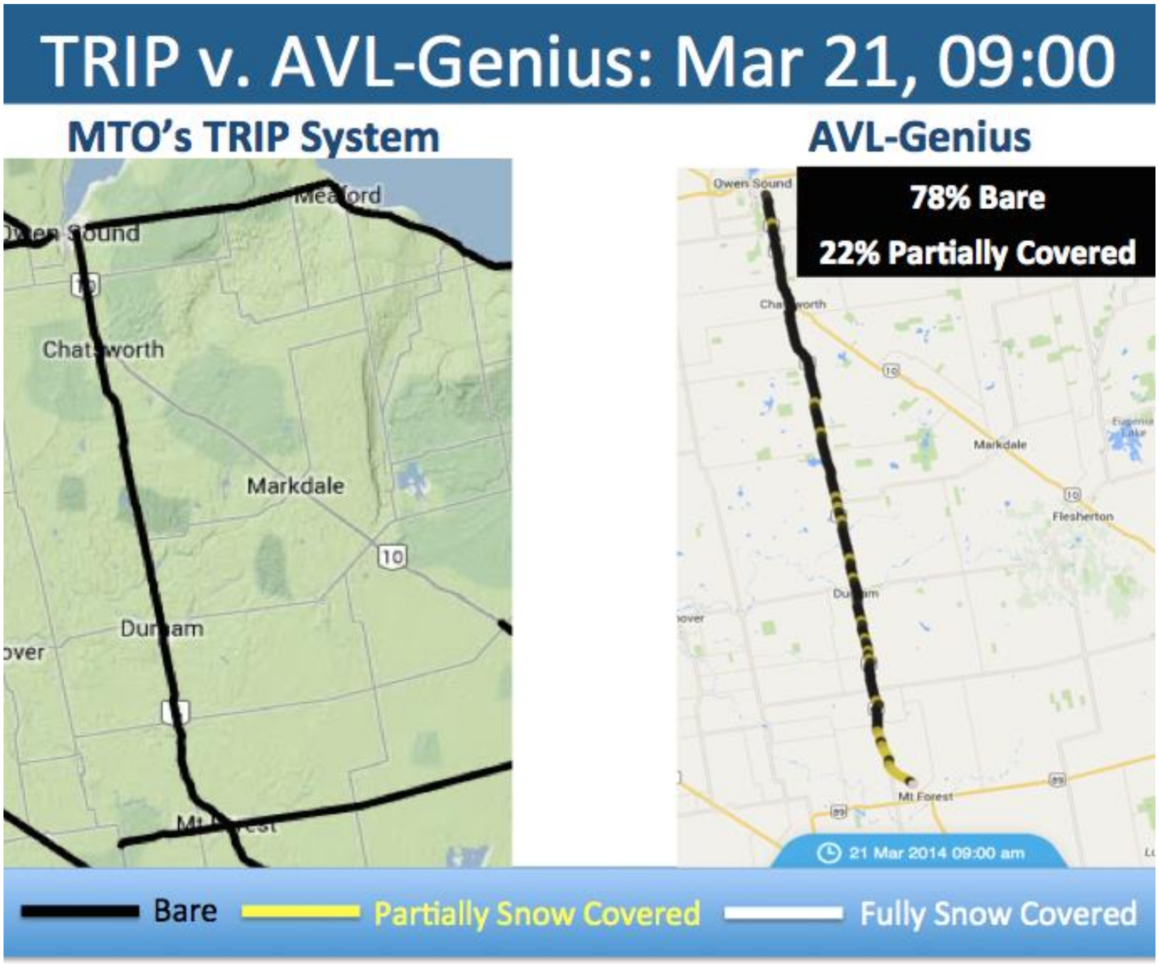
Road and Weather Information Sheet RSC: Partly Snow Covered, Bare and Wet, Bare and Dry



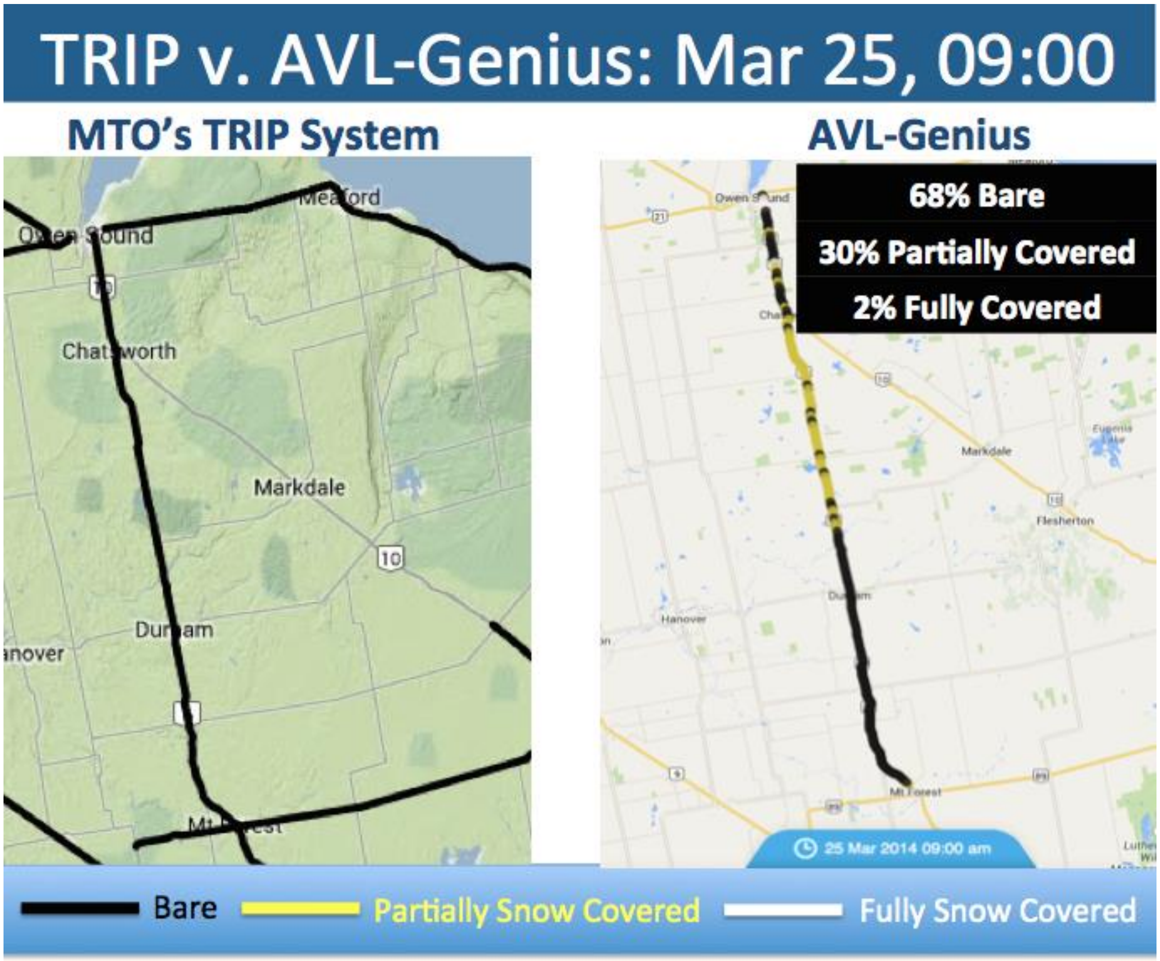
Road and Weather Information Sheet RSC: Bare and Wet, Bare and Dry



Road and Weather Information Sheet RSC: Partly Snow Covered, Bare and Wet, Bare and Dry



Road and Weather Information Sheet RSC: Snow Packed, Partly Snow Covered, Partly Snow Packed, Bare and Wet, Bare and Dry



Road and Weather Information Sheet RSC: Partly Snow Covered, Partly Snow Packed, Bare and Wet, Bare and Dry

Appendix E

Neural Network Model Results Sample – Three-Class System

Learning Rate = 0.1				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	93.7	74.9	61.4	84.3
0.2	93.4	75.3	58.5	84
0.3	92.8	74.7	60.1	83.6
0.4	93.2	75.6	52.9	83.6
0.5	94	75	55.6	84
0.6	94.2	74.6	54.6	83.9
0.7	93.6	74.9	51.3	83.4
0.8	93.6	74.2	56.2	83.5
0.9	93.3	75.8	29.7	83.4

Learning Rate = 0.2				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	93.8	76	55.6	84.3
0.2	93	73.9	58.2	83.3
0.3	92.9	75.7	51	83.3
0.4	93	74.2	53.3	83
0.5	92	76	52.9	83
0.6	93.4	74.3	55.6	83.4
0.7	93.4	74.8	52.3	83.4
0.8	91.5	75	43.5	81.6
0.9	90.4	69.5	47.4	79.4

Learning Rate = 0.3				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	93.7	74.7	55.9	83.8
0.2	92	74.2	58.2	83
0.3	92.6	75.4	58.5	83.6
0.4	93.4	74.3	50	83
0.5	92.9	72.8	51.3	82.3
0.6	93.2	75.4	43.1	82.7
0.7	91.4	77.4	39.2	82.1
0.8	91	72.1	44.1	80.4
0.9	88.6	69.2	37.9	77.5

Learning Rate = 0.4				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	92.1	74.9	52.6	82.7
0.2	93.7	75.6	50	83.6
0.3	93.8	72.4	57.8	83.2
0.4	92.6	75.2	48.7	82.8
0.5	92.1	74.9	48.7	82.4
0.6	90.2	75.9	49	81.7
0.7	92.7	74.8	44.1	82.3
0.8	90.6	71.3	42.8	79.8
0.9	85.7	68.7	32.4	75.2

Learning Rate = 0.5				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	93.5	72.9	54.9	82.9
0.2	92.9	72.8	54.9	82.9
0.3	93.3	75.4	44.4	82.9
0.4	91.9	75	54.2	82.8
0.5	92.7	75.8	47.7	83
0.6	90.5	75.3	46.1	81.4
0.7	90.7	70.4	48.4	79.9
0.8	88	73.7	39.2	78.8
0.9	85.6	57.8	36.6	71.6

Learning Rate = 0.6				
	Classification Accuracy (%)			
Momentum	**BP**	**PS**	**FS**	**Overall**
0.1	93.3	75.4	46.1	83
0.2	92.6	76.2	44.8	82.8
0.3	92.6	75.8	52	83.2
0.4	91	75.3	48.7	81.9
0.5	90.2	76.4	45.8	81.6
0.6	92.6	73.7	41.8	81.6
0.7	91.5	72.8	41.2	80.7
0.8	86.6	75.7	38.9	78.8
0.9	89.1	39.5	29.4	66.4

Appendix F

Random Tree Results – Three-Class System

Number of Random Features	Classification Accuracy			
	BP	PS	FS	Overall
1	94.7	77.5	67	86.2
2	94.4	77.8	66.7	86.1
3	94.5	77.3	68	86.1
4	94.3	77.7	67.3	86.1
5	94.5	77.8	67.6	86.2
6	94.2	78	67.3	86.2
7	94.4	77.8	68	86.3
8	94.4	77.7	68.3	86.2
9	94.3	77.9	67.6	86.2
10	94.5	77.6	68.3	86.3
11	94.5	77.9	67.6	86.3
12	94.4	77.8	67.6	86.3
13	94.4	77.8	68	86.3
14	94.4	77.8	68	86.3

Appendix G

Random Forest Results Sample – Three-Class System

Number of Random Features = 2					Number of Random Features = 5				
Trees in Random Forest	**Classification Accuracy (%)**				**Trees in Random Forest**	**Classification Accuracy (%)**			
	BP	**PS**	**FS**	**Overall**		**BP**	**PS**	**FS**	**Overall**
2	93.9	76.5	67.6	85.5	**2**	93.9	77	68.6	85.7
3	93.6	77.8	69	85.9	**3**	93.8	77.8	69	86
4	94	77.9	68.6	86.1	**4**	94.1	77.9	68.3	86.1
5	94.1	77.7	68.6	86.1	**5**	94.2	78	68.3	86.3
6	93.8	78.2	68.6	86.1	**6**	94.1	78.5	67.6	86.3
7	94.1	77.9	69.3	86.3	**7**	94	78.4	69	86.3
8	94.4	77.8	68.3	86.3	**8**	94.4	78	68	86.3
9	94.3	77.5	68.3	86.1	**9**	94.1	77.8	68	86.1
10	94.3	78.4	68.6	86.3	**10**	94.2	78.3	68.6	86.4
11	94.2	77.5	69	86.1	**11**	94.2	77.6	68.3	86.1
12	94.2	77.5	69	86.1	**12**	94.2	77.8	68.3	86.2
13	93.9	77.6	70.3	86.1	**13**	94.1	78.1	69.6	86.3
14	94	77.8	70.6	86.3	**14**	94.1	78.1	69.6	86.3
15	94	77.8	70.6	86.3	**15**	94.2	77.8	69.6	86.3
16	94.1	77.8	70.3	86.3	**16**	94	78.1	69.3	86.3
17	94	78.1	69.9	86.3	**17**	94	78.3	69.3	86.3
18	94	77.5	70.3	86.1	**18**	94.1	78	69.6	86.3
19	94.1	77.4	69.9	86.1	**19**	94.1	77.8	69.3	86.2
20	94.1	77.3	70.3	86.1	**20**	94.2	77.6	69.6	86.2

Number of Random Features = 8					Number of Random Features = 11				
Trees in Random Forest	**Classification Accuracy (%)**				**Trees in Random Forest**	**Classification Accuracy (%)**			
	BP	**PS**	**FS**	**Overall**		**BP**	**PS**	**FS**	**Overall**
2	93.9	77	67.6	85.6	2	93.9	77	67.6	85.6
3	93.9	77.6	68.6	86	3	94	77.8	69.6	86.2
4	93.9	77.7	68	85.9	4	94	78.1	68	86.1
5	93.9	78	67.6	86	5	93.9	78.1	68.6	86.1
6	93.7	78.4	67	86	6	93.7	78.4	68.6	86.1
7	94	78.3	68.6	86.3	7	94	78.4	70.6	86.4
8	94.3	78.1	68.3	86.3	8	94.2	77.8	69.3	86.3
9	94	77.8	68	86	9	94.1	77.7	68.3	86.1
10	94.2	78.4	68.3	86.4	10	94.1	78.1	69	86.3
11	94.1	77.5	68	86	11	94.1	77.5	69.3	86.1
12	94.1	77.6	68	86	12	94	77.6	69.3	86.1
13	93.9	77.5	69	86	13	93.9	77.8	70.6	86.2
14	94	77.7	69.3	86.1	14	93.9	77.8	70.6	86.2
15	94	77.8	69.3	86.1	15	93.9	77.9	70.6	86.3
16	93.9	77.6	68.6	86	16	93.8	77.9	70.3	86.1
17	94	78.1	69	86.2	17	93.8	78.2	70.3	86.3
18	94	77.7	69.3	86.1	18	93.9	77.5	70.6	86.1
19	93.9	77.7	69.3	86.1	19	93.9	77.7	70.3	86.1
20	94	77.6	69.6	86.1	20	93.9	77.6	70.6	86.1

Appendix H

Neural Network Sample Results – Five-Class System

Learning Rate = 0.2						
	Classification Accuracy (%)					
Momentum	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
0.1	96.2	11.7	17.4	76.1	51.8	74.1
0.2	96.3	11.7	14.7	76.3	54.2	74.2
0.3	96.5	9.4	15	77.4	51.5	74.1
0.4	96.1	10.5	14.7	78.7	50.8	74.1
0.5	96.4	9.4	15.3	77.7	52.5	74.2
0.6	96.4	10.8	14.2	75.4	53.2	73.8
0.7	95.9	9.9	15.9	76.4	51.5	73.7
0.8	96.1	6.7	13.9	76.7	49.2	73.2
0.9	94.9	7.6	18.6	72.4	41.9	71.7

Learning Rate = 0.3						
	Classification Accuracy (%)					
Momentum	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
0.1	96.3	7.6	12.1	79.2	46.2	73.4
0.2	96.4	7.3	11.5	77.4	49.8	73.4
0.3	96.7	8.8	15	77	50.8	74
0.4	96.2	9.9	13.6	78	50.5	73.8
0.5	96	13.2	14.2	74.4	48.5	73.3
0.6	96	8.2	20.4	74.2	51.2	73.5
0.7	96.6	6.7	14.7	71.8	54.2	73
0.8	96.2	7	20.9	73.5	49.2	73.3
0.9	95.2	5.3	10	73.2	35.9	70.5

Learning Rate = 0.5						
	Classification Accuracy (%)					
Momentum	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
0.1	95.5	11.1	19.5	76	47.5	73.5
0.2	96.3	8.8	18.3	77.1	47.5	73.9
0.3	96	8.5	20.1	77	49.5	73.9
0.4	96.3	11.1	13	77.1	49.8	73.8
0.5	95.9	8.5	16.2	73.1	45.2	72.5
0.6	95.9	7.3	21.8	70.9	44.2	72.4
0.7	95.3	5.3	18.6	69.9	45.8	71.5
0.8	93.8	8.8	18.3	72.1	43.9	71.2
0.9	72.8	9.4	14.5	30	60.1	52.9

Learning Rate = 0.6						
	Classification Accuracy ($)					
Momentum	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
0.1	96	10.5	13	77.3	46.2	73.3
0.2	95.8	10.2	15.6	75.7	49.5	73.4
0.3	95.1	12	15	73.7	53.2	73
0.4	96.2	9.4	17.7	74.2	45.2	73.1
0.5	95.9	5.6	17.7	76.3	48.2	73.2
0.6	96	9.4	18.3	67.9	51.2	72.3
0.7	95.6	9.1	18.3	70.6	46.2	72.2
0.8	93.8	6.4	20.9	64.1	43.2	69.7
0.9	88.2	0.3	3.8	32	39.2	58.3

Appendix I

Random Tree Results – Five-Class System

Number of Random Features	Classification Accuracy					
	BP	<25	25 to 50	50 to 75	FS	Overall
1	94.6	28.7	30.1	76.3	68.4	77.3
2	94.8	27.8	30.7	76	68.8	77.3
3	94.4	28.1	30.4	75.8	68.8	77.1
4	94.3	27.8	31	75.5	69.4	77.1
5	94.5	28.4	30.7	76	69.8	77.3
6	94.3	27.8	30.7	76	69.4	77.1
7	94.1	28.1	30.4	76.3	69.8	77.1
8	94.1	28.4	31	76.1	70.1	77.2
9	94.2	28.4	30.7	75.7	70.1	77.2
10	94.1	28.7	31	75.8	69.8	77.1
11	94.3	28.4	30.4	75.8	70.1	77.2
12	94.2	28.1	31	76.3	70.1	77.3
13	94.3	28.4	30.4	75.8	69.8	77.2
14	94.3	28.4	30.4	75.8	69.8	77.2

Appendix J

Random Forest Sample Results – Five-Class System

Number of Random Features = 3						
	Classification Accuracy					
Trees in Random Forest	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
1	93.7	24.9	32.4	72.5	65.8	75.7
2	94.3	27.8	31.3	71.6	68.8	76.3
3	93.6	30.7	31.9	72.4	67.8	76.3
4	94.2	31	33	72.2	71.1	77
5	93.8	31.6	30.1	73.7	72.4	77
6	93.7	32.2	31.9	72.9	70.4	76.8
7	94	31.6	31	73.5	71.8	77.1
8	94.1	30.4	31.9	73.4	72.1	77.1
9	93.7	31.9	32.2	73.5	71.8	77
10	93.8	31.6	31.3	74.2	70.8	77
11	93.6	316	31	74.8	70.8	77
12	93.6	31.3	32.4	74.2	71.8	77.1
13	93.9	31.6	31.3	74.1	71.8	77.1
14	93.9	31	32.7	73.7	71.4	77.1
15	93.8	31.9	33.6	72.5	71.4	77
16	93.7	32.2	33.6	72.8	72.4	77.1
17	93.8	31.9	33.3	72.9	72.1	77.1
18	93.8	32.5	32.2	73.5	71.8	77.2
19	93.8	31.9	33.6	72.6	72.1	77.1
20	94	31.6	32.4	73.2	71.4	77.1

Number of Random Features = 4						
	Classification Accuracy					
Trees in Random Forest	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
1	93.6	24.9	32.2	72.4	65.1	75.6
2	94.1	28.1	31.6	71.3	66.8	76.1
3	93.9	30.4	31.6	71.6	66.8	76.2
4	94.2	31	33.6	71.8	70.8	77
5	93.9	31.3	31	73.1	72.4	77
6	93.6	31.6	32.2	72.5	70.1	76.6
7	94	31.6	31.6	73.1	71.4	77
8	94.1	29.8	32.2	72.6	72.1	76.9
9	93.8	31.6	32.4	72.8	71.4	76.9
10	93.8	32.2	31.3	73.7	70.8	77
11	93.7	31.3	31.3	74.2	70.4	76.9
12	93.7	31	32.7	74.1	70.8	77
13	94	31.3	31	74.2	70.8	77.1
14	93.9	31.6	32.7	73.5	70.4	77.1
15	93.7	31.9	34.2	73.1	70.1	77
16	93.7	32.2	33.9	72.9	71.1	77.1
17	93.8	31.6	33.6	73.1	71.1	77.1
18	93.8	31.6	32.2	73.4	70.8	77
19	93.8	31.3	33.9	72.8	71.1	77
20	93.9	31.6	32.7	73.5	70.4	77.1

Number of Random Features = 8						
	Classification Accuracy					
Trees in Random Forest	**BP**	**<25**	**25 to 50**	**50 to 75**	**FS**	**Overall**
1	93.7	24.9	31.9	72.1	66.4	75.7
2	94.2	28.7	31.3	71.2	68.8	76.3
3	93.8	30.7	31.3	71.6	67.8	76.3
4	94.1	31.3	32.7	71.6	71.8	76.9
5	93.8	31.9	30.7	72.8	72.8	76.9
6	93.5	31.9	32.2	72.5	70.4	76.6
7	93.8	31.6	31.9	73.1	71.8	76.9
8	93.8	30.1	32.2	72.6	72.4	76.8
9	93.7	31.9	32.2	72.9	71.1	76.9
10	93.6	32.2	31	73.5	70.8	76.8
11	93.4	31.6	30.7	74.1	70.4	76.7
12	93.4	31.6	31.9	74.4	70.8	76.9
13	93.8	31.6	30.4	74.1	71.4	77
14	93.7	31.6	31.9	73.2	70.8	76.8
15	93.5	32.2	33.6	72.5	70.1	76.7
16	93.4	32.5	33.9	72.2	71.8	76.9
17	93.6	32.2	33.9	72.8	71.4	77
18	93.6	31.9	32.2	72.9	71.1	76.8
19	93.6	31.3	33.6	72.8	71.1	76.9
20	93.7	31.3	32.7	73.2	70.8	76.9

Bibliography

Al-Qadi, I., Loulizi, A., Flintsch, G., Roosevelt, D., Decker, R., Wambold, J. C., et al. (2002). *Feasibility of Using Friction Indicators to Improve Winter Maintenance Operations and Mobility*. National Cooperative Highway Research Program (NCHRP). Transportation Research Board (TRB).

Altınçay, H. (2007). Decision Trees Using Model Ensemble-Based Nodes. *Pattern Recognition , 40* (12), 3540-3551.

Altman, D. (1991). *Practical Statistics for Medical Research.* Chapman and Hall.

Belaroussi, R., Foucher, P., Tarel, J.-P., Soheilian, B., Charbonnier, P., & Paparoditis, N. (2010). Road Sign Detection in Images: A Case Study. *2010 20th International Conference on Pattern Recognition (ICPR)* (pp. 484-488). IEEE.

Bengio, Y. (2009). Learning deep architectures for ai. *Foundations and trends in Machine Learning , 2* (1), 1-127.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Springer.

Breiman, L. (2001). Random Forests. *Machine Learning* , 5-32.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees.* Chapman and Hall.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees.*

Buchanan, F., & Gwartz, S. E. (2005). Road Weather Information Systems. *Annual Conference of the Transportation Association of Canada.* Alberta: Transportation Association of Canada.

Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics , 22* (2), 249-254.

Casselgren, J. (2007). *Road Surface Classification Using Near Infrared Spectroscopy*. Thesis, Department of Applied Physics and Mechanical Engineering.

Casselgren, J., Sjödahl, M., & LeBlanc, J. (2007). Angular Spectral Response From Covered Asphalt. *Appl Opt. , 46* (20), 4277-4288.

Chapman, L., & Thornes, J. (2005). The Influence of Traffic on Road Surface Temperatures: Implications for Thermal Mapping Studies. *Meteorological Applications , 12* (4), 371-380.

Chapman, M., & Drobot, S. (2012). Connected Vehicle Observations in Weather Applications for Various Highly Impacted Users of the Roads. *Proceedings of the 16th SIRWEC Conference.* Helsinki, Finland: SIRWEC.

Cloutier, M., & Donaldson, L. (2007). *Applicability of Winter Friction Measurements in Ontario.* Eastern Region Contracts Office. Ontario Ministry of Transport.

Conrad, P., & Foedisch, M. (2003). Performance evaluation of color based road detection using neural nets and support vector machines. *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop.* IEE.

Cortes, C., & Vapnik, V. (1995). Support Vector Networks. *Machine Learning , 20*, 273-297.

Danilo, C. (n.d.). *Decision Tree Algorithm Weka Tutorial*. Retrieved March 4, 2015 from http://art.uniroma2.it/basili/MLWM09/002_DecTree_Weka.pdf

Deng, L., Hinton, G., & Kingsbury, B. (2013). New types of deep neural network learning for speech recognition and related applications: An overview. *38th IEEE International Conference on coustics, Speech and Signal Processing.* Vancouver: IEEE.

Dhurandhar, A., & Dobra, A. (2008). Probabilistic Characterization of Random Decision Trees. *Journal of Machine Learning Research , 9*, 2321-2348.

Emam, K. E. (1998). *Benchmarking Kappa for Software Process Assessment Reliability Studies.* Fraunhofer Institute for Experimental Software Engineering.

Environment Canada. (2014, 05, 15). *Canadian Climate Normals 1981-2010 Station Data.* Retrieved 10 21, 2014 from http://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?stnID=4486&lang=e&dCode=1&province=ONT&month1=0&month2=12

Erpicum, M., Frederic, M., Mabille, G., Nyssen, T., & Litt, S. (2005). Real-Time Regional Automatic Mapping of Nocturnal Winter Road Temperatures: Application of A Methodology of Analogous Situations, The Case of the Wallon Region (Belgium). *15th Conference on Applied Climatology.* American Meteorological Society.

Erdogan, G., Alexander, L., & Rajamani, R. (2008). Friction Coefficient Measurement System for Winter Maintenance Vehicles. *American Control Conference* (pp. 4585-4590). IEEE.

Fan, W., Wang, H., Yu, P., & Ma, S. (2003). Is random model better? On its accuracy and efficiency. *Third IEEE International Conference on Data Mining, 2003* (pp. 51-58). IEEE.

Feng, F. (2013). *Winter Road Surface Condition Estimation and Forecasting.* University of Waterloo, Civil and Environmental Engineering. University of Waterloo.

Feng, F., & Fu, L. (2008). *Evaluation of Two New Vaisala Sensors for Road Surface Condition Monitoring.* Ministry of Transport Ontario.

Feng, F., & Fu, L. (2009). Spectral Analysis of Continuous Friction Measurements for Winter Road Surface Condition Discrimination. *Transportation Research Board 88th Annual Meeting.* Transportation Research Board.

Feng, F., Fu, L., & Perchanok, M. (2010). Comparison of Alternative Models for Road Surface Condition Classification. *TRB Annual Meeting 2010.* Transportation Research Board.

Finnish Road Administration. (2008). *Winter Maintenance Policy*. Hensinki: Finnish Road Administration.

Foedisch, M., & Takeuchi, A. (2004). Adaptive Real-Time Road Detection Using Neural Networks. *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems.* Washington, DC: IEEE.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge University Press.

Fu, L., & Linton, M. A. (2014). *Field Test and Evaluation of a Mobile Automated Winter Road Condition Reporting System* . Ontario Ministry of Transportation.

GripTester. (n.d.). *GT MK2 | GripTester*. Retrieved February 17, 2015 from http://griptester.us/gt-mk2/

Haavasoja, T., Nylander, J., & Nylander, P. (2012). Experiences of Mobile Road Condition Monitoring. *SIRWEC 2012 Conference Proceedings.* Helsinki: SIRWEC.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. (2009). The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter , 11* (1), pp. 10-18.

Halliday Technologies Inc. (n.d.). Retrieved July 6, 2014 from http://www.hallidaytech.com/rt3grip.html

Handa, H., Chapman, L., & Yao, X. (2007). Robust Salting Route Optimization Using Evolutionary Algorithms. *Studies in Computational Intelligence , 51*, 497-517.

Hill, C. (2013). *Concept of Operations for Road Weather Connected Vehicle Applications.* Federal Highway Administration. U.S. Department of Transportation.

Hinton, G. E., Osindero, S., & Teh, Y. (2006). A fast learning algorithm for deep belief nets. *Neural Computation* .

Hinton, G. (2007). Learning multiple layers of representation. *Trends in Cognitive Sciences , 11* (10), 428-434.

Hinton, G., Deng, L., Yu, D., Dahl, G., M, A., Jaitly, N., et al. (2012, November). Deep Neural Networks for Acoustic Modeling in Speech Recognition - The shared views of four research groups. *IEEE Signal Processing Magazine* , pp. 82-97.

Hong, L., Jun, L., & Yanhui, F. (2009). Road Surface Condition Recognition Method Based on Color Models. *2009 First International Workshop on Database Technology and Applications.* IEEE.

Iowa Department of Transportation. (n.d.). *Iowa Department of Transportation Traveler Information*. Retrieved February 18, 2015 from http://hb.511ia.org/main.jsf

Jin, P., Walker, A., Cebelak, M., & Walton, C. (2015). Determining Strategic Locations for Environmental Sensor Stations with Weather-Related Crash Data. *Transportation Research Record: Journal of the Transportation Research Board , 2440*, 34-42.

Jonsson, P. (2011a). Classification of Road Conditions From Camera Images and Weather Data. *Computational Intelligence for Measurement Systems and Applications (CIMSA)* (pp. 1-6). IEEE.

Jonsson, P. (2011b). Road Condition Discrimination using Weather Data and Camera Images. *2011 14th International IEEE Conference on Intelligent Transportation Systems* (pp. 1616-1621). Washington, DC, USA: IEEE.

Jonsson, P., Casselgren, J., & Thörnberg, B. (2015a). Road Surface Status Classification Using Spectral Analysis of NIR Camera Images. *IEEE Sensors , 15* (2), 1641-1656.

Jonsson, P., Vaa, T., Dobslaw, F., & Thörnberg, B. (2015b). Road Condition Imaging - Model Development. *94th Annual TRB Meeeting.* Transportation Research Board.

Kalmegh, S. (2015). Comparative Analysis of WEKA Data Mining Algorithm RandomForest, RandomTree and LADTree for Classification of Indigenous News Data . *International Journal of Emerging Technology and Advanced Engineering , 5* (1), 507-517.

Kuehnle, A., & Burghout, W. (1998). Classification, Winter Road Condition Recognition Using Video Image. *Transportation Research Record* (1627), 29-33.

Kwon, T., & Fu, L. (2013). Evaluation of Alternative Criteria For Determining the Optimal Location of RWIS Stations. *Journal of Modern Transportation , 21*, 17-27.

Kwon, T., Fu, L., & Jiang, C. (2014). RWIS Stations - Where and How Many to Install: A Cost Benefit Approach. *Proceedings of the 93rd Annual TRB Conference.* Washington D.C.

Landis, J., & Koch, G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics , 33*, 159-174.

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News: The Newsletter of the R Project 2/3* , pp. 18-22.

Linton, M. A., & Fu, L. (2015). Winter Road Surface Condition Monitoring: Field Evaluation of a Smartphone-based System. *Transportation Research Board 94th Annual Meeting.* Washington: Transportation Research Board.

Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical* , 970-980.

Lufft. (n.d.). *Non Invasive Road Sensor NIRS31-UMB | Road Sensors - Lufft*. Retrieved February 17, 2015 from http://www.lufft.com/en/products/road-sensors/non-invasive-road-sensor-nirs31-umb-8710ut01/

Marchetti, M., Chapman, L., Khalifa, A., & Buès, M. (2014). New Role of Thermal Mapping in Winter Maintenance with Principal Components Analysis. *Advances in Meterology* .

Mitchell, T. (1997). *Machine Learning.* McGRAW Hill.

Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. *Computer Vision and Pattern Recognition* .

North Dakota Department of Transportation. (n.d.). *NDDOT - Road Weather Information Systems (RWIS)*. Retrieved February 18, 2015 from http://www.dot.nd.gov/divisions/maintenance/road-weather-information-system.htm

Öberg , G., & Gregersen, N. (1991). *MINSALT: An Experiment With Unsalted Road in Västerbotten County.* Swedish National Road and Transport Research Institute (VTI).

Omer, R. (2011). *An Automatic Image Recognition System for Winter Road Surface Condition Monitoring.* Department of Civil and Environmental Engineering. University of Waterloo.

Omer, R., & Fu, L. (2010). An Automatic Image Recognition System for Winter Road Surface Condition Classification. *13th Internattional IEEE Annual Conference on Intelligent Systems* (pp. 19-22). Portugal: IEEE.

Ontario Ministry of Transport. (n.d.). *Ontario 511 - Interactive Map*. Retrieved February 18, 2015 from http://www.mto.gov.on.ca/english/traveller/trip/map.shtml?ll=50.148746,-85.605468&z=5

Ontario Ministry of Transport. (2010). *Provincial Highways Traffic Volumes 1998-2010.* Highways Standards Branch. Ontario Ministry of Transport.

Perchanok, M. (1998). Friction Based Quality Assurance Measures for Snow Clearing Operations. *Xth PIARC International Winter Road Congress , 3*.

Pilli-Sihvola, Y., Toivonen, K., Haavasoja, T., & Haavisto, V. (2006). New Approach To Road Weather: Measuring Slipperiness. *Proceesings of the 13th International Road Weather Conference.* Italy: SIRWEC.

Quinlan, J. (1986). Induction of Decision Tree. *Machine Learning , 1*, 81-106.

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks.* Cambridge University Press.

Rokach, L., & Maimon, O. (2007). *Data Mining with Decision Trees: Theory and Applications. Series in Machine Perception and Artificial Intelligence: Volume 69.* World Scientific Publishing Co. Pte Ltd.

Roy, K., Chaudhuri, C., Kundu, M., Nasipuri, M., & Basu, D. (2005). Comparison of the Multi Layer Perceptron and the Nearest Neighbor Classifier for Handwritten Numeral Recognition. *Journal of Information Science and Engineering , 21*, 1247-1259.

Salimi, S., Nassiri, S., & Bayat, A. (2014). Using lateral coefficient of friction to evaluate effectiveness of plowing and sanding operations. *Canadian Journal of Civil Engineering , 41*, 977-985.

Sukuvaara, T., & Nurmi, P. (2012). Connected Vehicle Safety Network and Road Weather Forecasting - The WiSafeCar Project. *Proceedings of the 16th SIRWEC Conference.* Helsinki, Finland: SIRWEC.

Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan, R., & Feutson, B. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of Chemical Information and Computer Sciences , 43*, 1947-1958.

Takahashi, N., Kiriishi, M., & Tokunaga, R. (2013). A Study on Reproducibility of Friction Data Collected Using a Continuous Friction Tester. *ISCORD 2013: Planning for Sustainable Cold Regions,* (pp. 343-354).

Teconer. (n.d.). *Winter Maintenance RCM411*. Retrieved February 17, 2015 from http://www.teconer.fi/en/winter.html

Thaseen, S., & Kumar, C. (2013). An Analysis of Supervised Tree Based Classifiers for Intrusion Detection System . *Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME)* .

The Institution of Civil Engineers. (2000). ICE Design and Practice Guides: A Practical Guide. *ICE Design and Practice Guides:* .

Timofeev, R. (2004). *Classification and Regression Trees (CART) Theory and Applications.* Centre of Applied Statistics and Economics. Berlin: Humboldt University.

Transportation Association of Canada. (2011). *Winter Road Condition Terminology User Guide.* Transportation Association of Canada.

TWO. (n.d.). *TWO | Traction Watcher One*. Retrieved February 17, 2015 from http://www.two-friction.com

U.S. Department of Transportation. (2014). *National Connected Vehicle Field Infrastructure Footprint Analysis.* American Association of State Highway and Transportation Officials.

U.S. Department of Transportation. (n.d.). *RITA - Interlligent Transportation Systems*. Retrieved May 19, 2015 from United States Department of Transportation: http://www.its.dot.gov/press/mobile_weather_station.htm

Vaidya, J., & Fan, W. (2014). A Random Decision Tree Framework for Privacy-Preserving Data Mining. *IEEE Transactions on Dependable and Secure Computing , 11* (5), 399-411.

Vaisala. (n.d.). *DSC111 Remote Surface State Sensor - Vaisala*. Retrieved February 17, 2015 from http://www.vaisala.com/en/products/surfacesensors/Pages/DSC111.aspx

Viera, A., & Garrett, J. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine , 37* (5), 360-363.

Yamamoto, I., Kawana, M., Yamazaki, I., Tamura, H., & Ookubo, Y. (2005). The Application of Visible Image Road Surface Sensors to Winter Road Management. *Society of Automotive Engineers of Japan* (3888).

Ye, Z., Shi, X., Strong, C., & Larson, R. (2012). Vehicle-based sensor technologies for winter highway operations. *IET Intelligent Transportat Systems , 6* (3), 336-345.

Zhao, Y., & Zhang, Y. (2008). Comparison of Decision Tree Methods for Finding Active Objects. *Advances in Space Research , 41,* 1955-1959.