Winter Road Surface Condition Estimation and Forecasting

by

Feng Feng

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2013

© Feng Feng 2013

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

This thesis research has attempted to address two challenging problems in winter road maintenance, namely road surface condition (RSC) estimation and forecasting. For RSC estimation, the goal of the research was to develop models to discriminate RSC classes based on continuous friction measurements (CFM) and other available data such as temperature and precipitation history. A systematic exploratory study was conducted on an extensive field data set to identify the categorical relationship between RSC and various aggregate CFM measures, such as those related to probability distribution and spatial correlation. A new multi-level model structure was designed, under which binary logistic regression models were calibrated and validated utilizing several carefully chosen aggregate measures to classify major RSC types. This model structure was found to be effective in capturing the overlapping nature of CFM ranges over different RSC types – a problem which has not been addressed adequately in the past studies. An alternative model with support vector machine (SVM) was also developed for benchmarking the performance of the proposed logit model. It was found that the two types of models are comparative in performance, confirming the high performance of the proposed multi-level model.

For road surface condition forecasting, a novel conceptual framework for short-term road surface condition forecasting is proposed, under which the short-term changing process of surface temperature, friction level and contaminant layer depths, is comprehensively explored and analyzed. This study framework is designed to consider all important conditional factors, including weather, traffic and maintenance operations. The maintenance operations, especially salting, are handled by loosening the strict Markovian assumption, i.e., a history instead of one single time interval of salting operations is considered. In this way, the variation of snow/ice melting speed caused by both residual salt amounts and salt/contaminant mixing states is incorporated in the forecasting model, which enables accurate short-term forecasting for contaminant layers. This approach practically circumvents a major limitation of previous studies, making the post-salting RSC forecasting more reliable and accurate.

Under the proposed model framework, several advanced time series modelling methodologies are introduced into the analysis, which can capture the highly complex interactions between RSC measures and conditional factors simultaneously. Those methodologies, especially the univariate and multivariate ARIMA methods, are for the first time applied to the winter RSC evolution process. The forecasting errors of surface temperature, friction level and contaminant layer depths are all found to be small, implying that both the proposed study framework and the resulting solutions closely match the real-world observations.

The proposed forecasting models are simple in structure, easy to interpret and mostly consistent with physical knowledge. Compared to the existing models, the proposed models provide extra flexibility for refactory, tuning and deployment. Furthermore, all the modelled RSC measures are numerical and the forecast errors are relatively small, suggesting empirical models could be an efficient alternative to physical models. With the well-designed modelling methods, the resulting empirical models as calibrated in our study can be implemented into a decision support and simulation tool with high temporal resolution and accuracy.

Acknowledgements

Completing my doctoral thesis is probably the most challenging and rewarding activity in my life as I approach middle age. I have learnt a lot from writing this thesis searching for the answers to both research problems and life. The best and worst moments of this journey have been shared with many people. It has been a great privilege to spend these years in the transportation research division of our department.

Foremost, I would like to express my sincere gratitude to my advisor Prof. Liping (Lee) Fu for his motivation and patience throughout my whole study. His guidance in every respect and at every stage of my research has been outstanding and indispensable to me. I could not imagine having a better mentor and role model in my academic career.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Carl Haas, Prof. Bruce Hellinga, Prof. Zhou Wang, and Prof. Ming Zhong, for their encouragement and insightful comments.

My sincere thanks go to Ministry of Transportation Ontario (MTO) for providing financial support and collecting field data for my study. Special thanks to Max Perchanok of MTO for all the guidance, communication, assistance and coordination.

I thank my fellow colleagues in our research group: Dr. Chaozhe Jiang, Romona Mirtorabi, Jonathan T.J. Kwon, Raqib Omer Mian, Luchao Cao, Dr. Taimur Usman and Bezhad Hashemloo, for the stimulating discussions and for all the fun we have had.

I wish to thank my parents, Enyou Feng and He Huili. Their love is my inspiration and driving force each step of the way. My wife, Nan Zhou, is the best companion in this long journey. I have always been surprised at the love, encouragement and patience she has given to me.

Dedication

This thesis is dedicated to Elena and Alex, my daughter and son. Their smiling faces would have been its most beautiful front cover.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Winter is a beautiful season. But in countries with severe cold weather, heavy snowfall and low temperatures could pose significant challenges to various aspects of people's lives. Among these challenges, automobile transportation is probably the most significant. Road safety and mobility could be significantly compromised during snow events accompanied by high wind, poor visibility, and cold and slippery road surface conditions.

The impact of winter weather on the safety and mobility of a road network has long been the subject of research from different perspectives. Ontario Road Safety Annual Reports (1993-2009) indicated that crashes happening on wet and snowy/icy road surface respectively accounted for 17.6%∼23.8% and 9.4%∼19.9% of total car accidents. In a study on identification of dangerous highway locations in Quebec, Brown & Baass (1997) found the winter season is associated with low rates of mortality-and-serious-injury accidents and high rates of property-damage-only accidents. Wallman & Åström (2001) estimated that collision risk increases by 50%∼100% during snowfall. The increase in risk is more related to the friction reduction of the road surface. Goodwin (2002) found that over 22% of total crashes in the United States are weather-related with 13% happening during snowfall. Wallman (2004) also found that accident rates could be 16 times higher on black ice road surfaces than on bare dry road surfaces. In their meta-analysis about the findings of thirty-four papers between 1967 and 2005 examining the interaction of weather and traffic safety, Qiu & Nixon (2008a) concluded that snow events could increase the crash rate by 84% and injury rate by 75%. Usman et al. (2010) were trying to quantify the empirical relationship

between accident occurrence and road surface conditions on event-based data and they found that their correlation is highly significant.

The adverse impact of snow events on roadway mobility can be more straightforwardly observed. Hanbali & Kuemmel (1993) and Knapp & Smithson (2000) reported snow events could lead to as high as 50%∼60% reduction in traffic volume. Wallman & Åström (2001) found the speed reduction due to wet/snowy/icy roadway conditions could be as large as 20%. A study based on 350 permanent traffic counter stations on the Alberta provincial highway network showed a reduction of 0.5%∼3% in traffic volume for each centimeter of snowfall during severe cold conditions (Datla & Sharma, 2008).

In summary, snow storms and the resulting adverse road surface conditions could have a significant impact on road safety and mobility, increasing the crash probability and reducing service accessibility, highway capacity and traffic speed. To minimize these impacts of winter weather, systematic snow and ice control programs, involving plowing, salting, sanding or their combinations, are commonly implemented in cold countries. These winter road maintenance operations play a critical role in keeping highway transportation safe and fast, but they have significant financial and environmental implications. For instance, the monetary cost of winter maintenance activities in Canada is in the range of $1.3 billion per year (Jones, 2003). Morin & Perchanok (2003) estimated that approximately five million tons of sodium chloride are used each winter season across Canada for snow and ice control. Road maintenance chemicals may induce corrosion of vehicles and highway infrastructures, contaminate ground and surface water in close proximity to highways by runoff and seepage, and damage roadside vegetation. The severity of these adverse effects on the environment is proportionate to the amount of salt used (Perchanok et al., 1991; Transportation Association of Canada, 2008; Burtwell, 2001). Therefore, maintaining a satisfactory level of service (LOS) for road users using less salt has become a significant challenge for road maintenance authorities and practitioners.

As for improving the efficiency of maintenance work, a large body of knowledge and experience has been accumulated, summarized and transformed into maintenance guides, manuals and policies. Meanwhile, a lot of research work is ongoing, new thoughts and innovative ideas are being proposed and tested, and advancing technologies are continuously employed. The next several sections of this chapter serve as an introduction of relevant knowledge about winter road maintenance.

1.2 Winter Road Maintenance

Many methods and techniques are available for roadway snow and ice control, and they generally fall into two categories, namely, chemical and mechanical (Minsk, 1998). Chemical methods involve applications of chemical agents, such as Calcium Chloride or Sodium Chloride, to lower the freezing-point of the water on a pavement surface. Mechanical methods include plowing, scraping, and air blowing (scouring a snow covered pavement using high-velocity air).

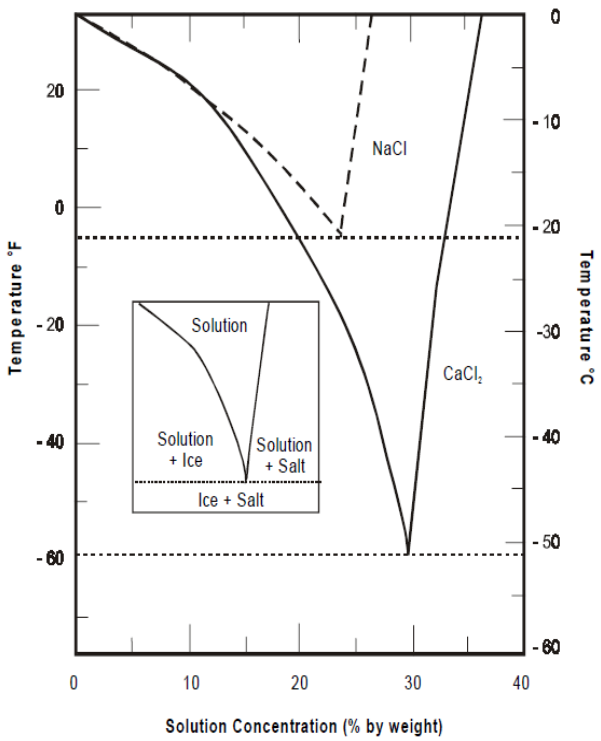


Figure 1.1: Phase Diagrams for $NaCl$ and $CaCl_2$
(Canadian Strategic Highway Research Program, 2000)

Each maintenance method has its own working mechanism, and thus it is only effective under certain road weather and surface conditions. This is especially true for chemical based methods. The freezing point of a chemical solution, also called the eutectic temperature, could vary markedly, depending on the type and concentration of the applied chemical. Figure 1.1 shows the phase states of Calcium Chloride and Sodium Chloride

by temperature and concentration. If there is no chemical in a solution (i.e. zero concentration), the freezing point of the solution is 0 °C. When the chemical concentration increases, the freezing point decreases until reaching the tipping point beyond which adding more chemical will have no effect on the freezing point. An important implication of this pattern is that any decisions on chemical application, such as type and rate, must ensure the effective temperature range of the chemical matches the actual road surface temperature. Some field studies have confirmed the large variation in freezing-point depressing capability of different chemicals (Ketcham et al., 1996; Minsk, 1998; Fu et al., 2006).

Deicing chemicals may be applied in either solid, pre-wetted solid or liquid form, and the decision on which form to apply usually depends on many factors such as weather, road surface condition, availability of salting equipment, and costs. Solid chemicals require moisture to become effective and are therefore most suitable when the road surface is wet. On the other hand, pre-wetted salt or liquid salt solutions become more effective when the road surface is dry. Salt can be applied in advance of a snow storm event to prevent snow and ice from being bonded to the pavement surface. This proactive strategy, called "anti-icing", makes the subsequent maintenance operations, especially plowings, much easier. When this anti-icing strategy is adopted, the amount of chemicals needed is usually significantly less than if the chemical is applied to existing compacted snow/ice (Canadian Strategic Highway Research Program, 2000). Regardless of the form of application, accurate information about the current and future conditions of road surface temperature and precipitation is critical to determine time, location, and chemical amount.

When the road surface temperature drops below the effective temperature range of any available deicing chemicals, sand is usually applied as a direct means of increasing road surface friction level. On the other side, when snow and ice accumulate to a certain depth, applying a mechanical process to remove pavement contaminants becomes a more effective approach than salting and sanding.

1.3 Road Weather and Surface Condition Monitoring

With the availability of a wide variety of snow and ice control methods, each having unique characteristics and optimal working conditions, timely and accurate information about the road weather and surface conditions of maintenance routes is a basic requirement in the development of efficient maintenance strategies and operation plans. In other words, the delivery of an effective winter road maintenance program largely depends on the level of knowledge about the current and future road weather and surface conditions, which is the primary concern of this research.

Real-time road surface conditions are traditionally monitored by human observation. This practice is changing around the world as new technologies are being developed and implemented. Some of these technologies have shown great promise in improving the monitoring of winter road surface conditions and thus the efficiency and cost-effectiveness of maintenance operations. However, applications of most of these technologies are still at an experimental stage due to incomplete understanding of their effectiveness, reliability, and costs/benefit tradeoff. A detailed discussion about these technologies is given in Chapter 2.

1.4 Policy and Performance Measure

In order to promote consistent service delivery and optimal resource planning, most highway agencies have established their own winter road maintenance standards or policies that designate a minimum LOS that must be maintained for different classes of highways. LOS is a concept used to represent the driving quality of the road surface from the drivers' perspective and is commonly defined on the basis of road surface conditions such as snow cover, snow depth and friction level. Once a service standard is established, the road surface condition (RSC in short) is monitored accordingly to measure the performance of maintenance operations and thus evaluate how well the service standard is met. Winter road maintenance policies are generally classified into two major types in terms of RSC measure:

1. bare pavement (BP) policy
2. friction-based (FB) policy

A BP policy sets a service goal of recovering snow/ice covered roadways to a bare state within a specific time frame. In general, a set of LOS classes are defined for different types of roadways on the basis of bare pavement recovery requirements in combination with some other RSC measures, like snow depth. For example, Ministry of Transportation Ontario (MTO) specifies that the Class I roads should be restored to essentially bare condition no later than eight hours after a storm ends (http://www.mto.gov.on.ca/english/engineering/winter-highway-maintenance.shtml).

BP policies are currently adopted by most provinces and states in Canada and the US. One major advantage of BP policies is that they are easy to implement and monitor. Bare pavement status and snow depth are usually reported by maintenance or quality assurance personnel based on periodic visual inspection during and after snow events. Because of the descriptive nature of the measure, it conveniently circumvents the need to account

for the random variations of RSC. On the other hand, the same nature also suggests the drawbacks of this method, mainly, the lack of objectivity and repeatability. Also, BP policies do not consider the winthin-storm conditions. Actually, maintaining RSC to meet certain minimum requirement during a snow event could become very important if the snow event lasts a long time.

FB maintenance policies represent an alternative to BP policies with the intention of addressing the drawbacks of BP policies. Friction is the level of the resistive force to the movement between road pavement and vehicle tires and is therefore by definition the most accurate and objective measure of driving quality of a road surface. Many field and laboratory studies have confirmed the significant correlation between RSC types and friction level, and have suggested that using friction measurement as a tool for decision making and performance measure is highly feasible. But field studies have also shown that different RSC types could produce very similar friction levels. This could be caused by the uncertainties associated with friction measurement itself or the fact that friction measurement is a line measurement and poor in lateral coverage. As a result, using friction levels solely as the maintenance policy standard could cause serious problems under certain circumstances. This is probably the major obstacle hindering a wide adoption of FB policies around the world. In any implementation of FB policies, technically, the following issues inevitably need to be addressed:

- How to consistently and reliably infer RSC based on friction measurement? The literature review in Chapter 2 indicates that existing friction-based performance measure methods vary significantly in terms of classification schemes. Furthermore, all existing approaches use mean friction measurement as the only RSC discriminator, and ignore other factors that could potentially influence the friction measurement. Thirdly, different friction measuring devices give quite different measuring results. In order to give consistent measurement, they have to be inter-calibrated, which is very challenging due to their different measuring mechanisms.
- How a friction measuring device should be operated to maximize its effectiveness as a road condition monitoring tool, e.g. what are the best sampling and reporting intervals and methods?

1.5 Research Objectives

As discussed in previous sections, effective winter road maintenance requires timely and reliable monitoring, estimation and forecasting of RSC. The latest developments in sensor

and information technologies have significantly advanced the means for RSC monitoring and data collection, making sheer volumes of real-time data available for maintenance management and traveller information systems. How to efficiently turn these data into models and tools giving information that is directly useful for the decision-making of maintenance service providers and travellers is the major concern of this thesis study. The goal of this research is to address this concern with the following specific objectives:

1. To conduct a systematic study on the characteristics of RSC types, their representation schemes and measures available in literature.

2. To investigate the correlation patterns between alternative measures of RSC. The research will primarily focus on two condition measures that are critical to the current state of maintenance practice, which are friction level and types of RSC.

3. To show the possibilities of inferring RSC by using vehicle driving data, especially speed. This type of data can be ubiquitously collected at very low cost and with negligible measurement error.

4. To systematically explore the relationship between RSC changing processes and maintenance operations under the influence of environmental factors, such as weather and traffic.

5. To develop models that make short-term RSC forecasting during and/or after snow events. These models should recognize the need to account for their inherent uncertainties and variations, and the availability of a variety of real time data such as road weather and surface conditions, maintenance operations, and traffic.

As indicated above, the modelling work of this thesis study consists of two major parts: RSC estimation and forecasting. The main task is to develop models for both parts, and the models should be easy to interpret and reconcile with common physical knowledge. Of particular importance, the modelling methodologies should make the resulting models be more *actionable* and *agile*. Actionable models can be calibrated, validated, and applied *in situ* under current technology conditions at reasonable cost. Agile models can be promptly tuned and refactoried with continuously collected data.

The focus of this study is not to build solid general *first principles*. Instead, the overall philosophy of the modelling work is data driven, result oriented and cost conscientious from an engineering point of view; notwithstanding, the calibrated models and their implications should not contradict physical common sense and widely agreed judgments.

1.6 Overview of Chapters

The main body of this thesis comprises five chapters and a set of supporting appendices. Chapter 1 introduces the background, addressed problems, study objectives and scope of this thesis. Chapter 2 details an extensive literature review on RSC monitoring and estimation, friction measurement and RSC forecasting. Chapter 3 describes the process and results of a systematic analysis of RSC estimation using friction measurement and vehicle speed data. Chapter 4 comprehensively explores the RSC changing process. A modelling framework tailored to the study objective is proposed, under which a set of calibrated short-term RSC forecasting models is calibrated, validated and discussed. Chapter 5 summarizes the major findings, implications and conclusions, and provides recommendations for future study.

Chapter 2

Literature Review

The primary objective of this research is to develop models that can be used to estimate and forecast RSC during/after winter storm events based on the data collected from various automatic monitoring technologies. This chapter reviews recent developments in relevant research fields including road weather monitoring technologies, traditional and state-of-the-art de-icing methods, RSC estimation methodologies, RSC forecasting models and related decision support systems.

2.1 RSC Monitoring

The modelling efforts related to RSC estimation and forecasting are mostly based on and limited by availability of data relevant to the corresponding physical processes. The basis for this study is the RSC measure data and related monitoring instruments/systems. Therefore, this literature review starts from RSC monitoring technologies and the measures they provide.

As mentioned in Section 1.3, winter road surface conditions (RSC) are traditionally monitored by human observations; however, the rapid development in sensor technologies is likely to make this practice obsolete. A variety of popular RSC monitoring instruments/ systems are introduced below. Some of them are under active testing, while the others have already been extensively studied and reliably used in practice.

(a) Ontario RWIS Sites

(b) RWIS Station and Pavement Sensor

Figure 2.1: RWIS Stations

2.1.1 Road Weather Information System (RWIS)

RWIS is a system consisting of meteorological stations strategically located alongside the roadway that monitor roadside weather and/or pavement conditions. An RWIS station is a system of sensors connected together and commonly configured to provide accurate real-time site-specific pavement conditions and weather data. These data usually include air and pavement temperature, relative humidity, wind speed and direction, precipitation rate and so on (http://www.aurora-program.org/rwis.cfm). The observational data from RWIS stations is usually communicated in real time to a central computer server where it is archived, organized and then presented on web pages. The real-time data can reduce the need and dependency for road patrolling in specific locations by providing timely and accurate information to those responsible for directing winter maintenance operations.

RWIS technology has been adopted widely around the world (Toivonen & Kantonen, 2004). Ontario started to implement RWIS in the mid-1990s and the Ministry of Transportation Ontario (MTO) had installed more than 100 RWIS stations across the province by 2005 (Buchanan & Gwartz, 2005). Figure 2.1 shows the map of RWIS sites across Ontario and the images of a typical RWIS station tower and an embedded pavement sensor, which is a part of an RWIS station.

The technologies used on RWIS sensors are under fast development and MTO is very active in testing and adopting such innovative winter maintenance technologies (Perchanok et al., 1991; Perchanok, 1994; Ministry of Transportation Ontario, 2003; Pilli-Sihvola et al., 2006; Feng & Fu, 2008). The installation cost of an RWIS station with basic functions is

more than $50,000, not including the cost for maintenance and installation of additional pavement sensors (Buchanan & Gwartz, 2005). With such a high cost, it is economically impossible to install RWIS stations with a high spatial density along the roadway. Therefore, the sparsely distributed spot-wise measurements at RWIS stations can only draw an incomplete picture about the RSC over the whole roadway network. The lack of information at locales not close to the stations is a major limitation of RWIS systems. During winter snowstorms, the RSC is affected by environmental factors, such as snowfall, sunlight, temperature, plowing, salting, traffic, etc., which naturally vary widely across different areas of a large road network. This fact usually causes the RSC to significantly vary along the roadway. So using spot-wise RWIS measurements to represent the RSC of one section or multiple sections of roadways is intrinsically questionable.

2.1.2 Thermal Mapping

Thermal mapping is a process by which the spatial variation pattern of road surface temperature is measured, archived, modelled and mapped. Usually the road surface temperature is remotely measured with an infrared thermometer mounted on the operating vehicle. As the distribution pattern of warm and cold sections within a specific road section is determined by its local environmental factors and weather conditions, quantifying the spatial variation of the road surface temperature is a process of generating a unique temperature map (or fingerprint) for this road section. With thermal mapping, it is possible that salt application schedules and rates can be optimized to match the variation pattern of the pavement temperature along a route, reducing the chance of over-application or under-application of salt.

Achieving reliable thermal mapping is, however, a challenging task. First, years of pavement temperature data need to be collected and analyzed in order to identify and establish reliable condition-dependent temperature profiles, as RSC is affected by snow precipitation and maintenance operations. Another limitation of thermal mapping is that pavement temperature is only one RSC measure; thus, it does not consider other important RSC measures, like snow coverage and depth.

2.1.3 Continuous Friction Measurement (CFM)

Friction measurement is the measurement of the friction coefficient between the vehicle tires and the roadway surface. The friction coefficient is measured with some specially designed testing tires mounted on a normal vehicle in touch with the pavement surface. Most recent

popular friction measuring devices can continuously collect data by certain spatial intervals along the test road; therefore, they are called continuous friction meters. The friction measurement they make is called continuous friction measurement (CFM). As the friction measurements are relatively easy to make and can cover long stretches of routes, CFM has been used in Nordic countries as a powerful tool to monitor winter RSC and evaluate maintenance performance (Norwegian Ministry of Transport and Communication, 2003).

It has been widely accepted that CFM is a technology that enables an objective, rapid and reliable assessment of RSC. As the snow coverage on road surface is negatively correlated to its friction coefficient, measuring the friction coefficient can give the most direct information on RSC. With real-time on-the-spot CFM, the quality of maintenance operations can be objectively evaluated, timely location-specific information on driving conditions can be relayed to the users, and salt application can be directed to where it is most needed (Nixon, 1998b; Perchanok, 1998; Wallman & Åström, 2001; Al-Qadi et al., 2002).

But CFM technology has several major limitations. First, characteristics of tire and pavement affect CFM, causing a large degree of uncertainty. Repeated CFMs on the same patch of roadway surface could be quite different suggesting a measurement variance that cannot be ignored. Second, different CFM meters usually give different value ranges for the same RSC type; therefore, CFM meters have to be inter-calibrated to allow their measurements to be meaningfully integrated/compared. Lastly, it is highly challenging to get a reliable mapping from CFM value ranges to RSC types due to the reasons detailed in Section 2.2.

2.1.4 Web-Based Surveillance Video and Automatic Road Surface Image Recognition

RSC can be monitored using remote video technologies such as web cameras and CCTV (Closed-Circuit Television) and transferred in real time to maintenance staff and road users via the Internet. As the Internet has become a ubiquitous media, web-based surveillance is a cost-effective alternative for remote monitoring during winter storms. It can provide timely warning of snowy or slippery conditions and reduce response time in directing plowers and spreaders to the monitored site. An example of such a system is shown in Figure 2.2. The image is extracted from MTO Traveller's Road Information Portal web page, on which the interactive map enables users to select a CCTV camera at certain location to show RSC (http://www.mto.gov.on.ca/english/traveller/trip/map.shtml). The CCTV cameras automatically capture and update the video image at a pre-configured time interval.

While RSC images can be manually monitored using CCTV cameras by maintenance

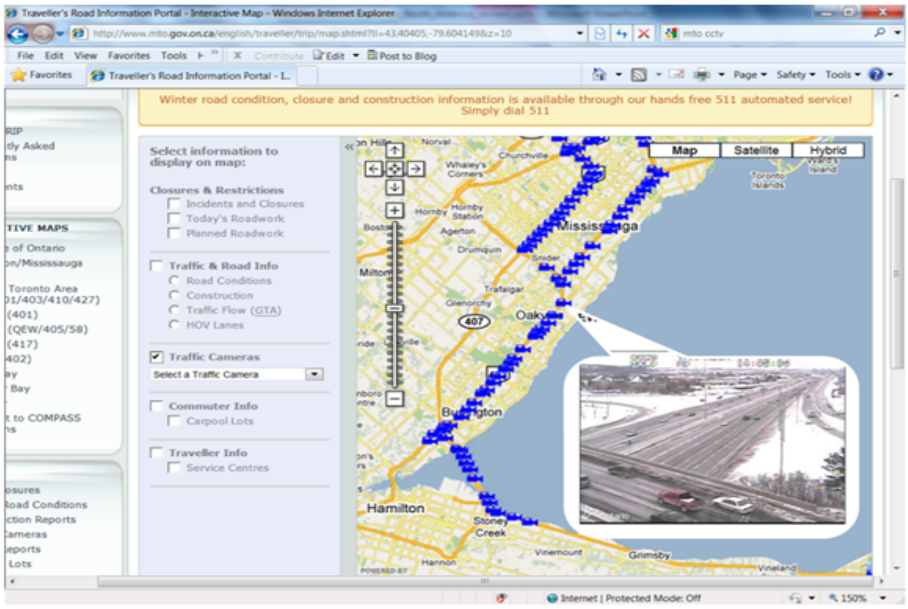


Figure 2.2: MTO Traveller's Road Information Portal

staff, the development of image recognition technology has made it possible to quickly classify RSC, based on video images using advanced computer algorithms. Image recognition technology has been a subject of intensive research for a couple of decades and has been applied to a variety of engineering problems. As common commercial web cameras are much cheaper, and the video image can be transferred and analyzed in real time, this monitoring technology has shown huge application potentials (Omer & Fu, 2010; Kawai et al., 2012).

While the technologies discussed previously have shown great promise in improving the monitoring of winter road surface conditions and thus the efficiency and cost-effectiveness of maintenance operations, they have limitations in terms of not only the types of condition parameters being monitored but also spatial and temporal coverage. For example, RWIS sensors are capable of providing temporally continuous monitoring of road weather and RSC such as temperature, precipitation, snow/ice but are limited to specific locations. In contrast, CFM can map out the surface slipperiness of a whole maintenance route or road network, but only at a few limited patrol observation runs. Image recognition accuracy is significantly affected by the changing ambient light, which is always a concern during snow events.

2.2 RSC Measures

Because of the limitations of these monitoring technologies, human observation is still the most important approach for RSC monitoring. Taking Ontario for example, maintenance staffs are sent out during and after snow storms to inspect highway surface conditions such as snow depth, snow coverage and BP status (Ministry of Transportation Ontario, 2007). Patrolling frequency may vary from every half an hour to six hours depending on the nature of the storm and the time of the day.

The reported RSC based on human observations are usually defined by contaminant types, contaminant coverage patterns, contaminant depths or a combination of them. The semantics used in the report are mostly descriptive, qualitative and fuzzy, sharing some major properties of human natural language. For example, NCHRP (2004) proposed a systematic RSC representation scheme that covers the following RSC types:

- center-line bare
- wheel path bare
- loose snow covered (percent area and depth)
- packed snow covered (percent area and depth)
- bare (percent area)
- thin ice covered (percent area)
- thick ice covered (percent area)
- dry
- damp
- slush (percent area and depth)
- frost
- wet

Note that the NCHRP classification is not completely categorical as additional quantifiers (e.g. percent area and depth) are required to define some of the categories, such as loose snow covered and slush. Those quantifiers are manually estimated thus are inherently subjective and inaccurate. Recognizing the challenge of establishing a comprehensive description of RSC, this classification system is relatively complete and free of ambiguity, and can be conveniently applied to almost all types of highways by the personnel with minimum training. MTO uses a much simpler RSC monitoring scheme in their weather/operation report sheets, which classifies RSC of a maintenance route into the following categories (Ministry of Transportation Ontario, 2007):

- bare dry
- bare wet
- slushy
- partially snow covered
- snow covered
- snow packed

Transport Quebec uses a very similar video-based RSC classification scheme including the following categories (http://www.quebec511.gouv.qc.ca/en/index.asp):

- bare (dry or wet)
- partly (snow covered,snow packed or ice covered)
- snow covered, snow packed or ice covered

A guide has been developed by TAC to define common terminologies that describe winter road conditions across Canadian provinces and territories (Hodgins et al., 2011). The RSC classification scheme in the guide is as follows:

1. bare
    - bare and wet
    - bare and dry
2. partly covered
    - partly ice covered
    - partly snow packed
    - partly snow covered
3. covered
    - ice covered
    - snow packed
    - snow covered

The major problem with this descriptive approach of representing RSC is that there is no clear distinction between some of the categories. As a result, people could give different descriptions about the same RSC, which stems from the fuzzy nature of human languages. In addition, the snow coverage and depth could have a wide interpretation due to their

significant spatial variation along a maintenance route. Similarly, a highway section, even a very short one, is likely to show obviously different RSCs, especially during snow event periods. Although the monitoring personnel can make multiple-choice style reports on RSC to address this spatial variation problem, the geospatial boundaries of different RSC types cannot be quantitatively identified; thus it helps very little in improving the efficiency of maintenance operations.

To address the drawbacks of categorical RSC monitoring schemes, CFM of winter RSC is used in some countries (Al-Qadi et al., 2002; Nixon, 1998b; Perchanok, 1998; Wallman & Åström, 2001). Sweden, Norway and Finland have their own CFM-based roadway maintenance standards and LOS definitions. Friction measurements can be used as the main basis for the decisions of sanding and salting operations (Norwegian Ministry of Transport and Communication, 2003). It is also used as a benchmark to evaluate the efficiency of the maintenance activities (Toivonen & Kantonen, 2004).

Besides categorical representation of RSC and numerical CFM, RSC measures could also include pavement surface temperature, freezing point and residual salt, both of which are measured by specific types of sensors (Mechler et al., 2000; Hunt et al., 2004). Table 2.1 summarizes RSC measures with respective monitoring approaches.

Table 2.1: RSC measures

RSC Measure	Measuring Method
contaminant types snow depth snow coverage	human observation pavement sensors
coefficient of friction	CFM meters
pavement surface temperature contaminant freezing point residual salt	pavement sensors

2.3 RSC Estimation

Among the RSC measures listed in Table 2.1, friction level and contaminant types have close correlations. They could be considered as two perspectives of the same thing: one is the visual appearance, and the other is the resulting slipperiness level. The visual

appearance of RSC is the most important performance measure for maintenance operations, especially in North America. On the other hand, the objective and quantitative nature of friction measurement can provide highly reliable support for decision making when combined with visual RSC observations. In fact, even in the countries using friction-based (FB) policies, human visual observation and descriptive RSC reporting are still used as an indispensable monitoring means. One major objective of this study is to explore the mapping relationships between friction level ranges and visually different RSC types, find ways to improve the reliability of this mapping, and numerically model this mapping. In this section, previous studies in this aspect are reviewed, and important findings are discussed.

2.3.1 Road Surface Friction Measuring Devices

The term *"friction"* used in this study refers to the friction coefficient, which is a measure of the resistive forces to movement between two opposing object surfaces, mathematically expressed by the equation:

$$\mu = \frac{F}{N} \tag{2.1}$$

where μ is the friction coefficient, F is the resistive force between two surfaces, and N is the normal force applied on the surfaces, which is usually the weight of the moving object.

The friction coefficient between moving vehicle tires and winter road surfaces is the major concern in our study. Thus, in this context, F is the slippery resistive force between the moving tire surface and the winter road surface, and N is the force put on the moving tire.

Based on Equation 2.1, the characteristics of the tire and pavement have a significant effect on the amount of friction that can develop (Kummer, 1996; Sandberg, 1997). The use of standardized friction measuring devices helps reduce the effect of tires on the variation of the friction measurement (Andresen & Wambold, 1998). The effect of pavement contaminants on friction measurement can therefore be isolated by using standardized friction measurement devices for further RSC analysis.

The most important factors related to friction measuring devices that influence the friction measurement are tire, position and normal force (Bachmann, 1998; Henry, 2000). In order to get stable and comparable friction coefficient values on a patch of road surface, friction measurement must be standardized in terms of these factors (Wallman & Åström, 2001). Existing standardized friction measuring devices fall in four major categories based

on the relative movement pattern between the measuring wheel and pavement: locked wheel, constant slip ratio (slip speed is normally between 10 and 20% of vehicle speed), variable slip ratio (slip speed is changing between 0 and 100% of vehicle speed) and constant slip angle (usually 20 degrees) measurement meters. The following list includes some popular friction measuring devices available in the market:

- Locked Wheel Devices: VDOT
- Constant Slip Ratio Devices: RFT, ASFT, Griptester, BV11, BV14, SFT, TWO
- Variable Slip Devices: French IMAG, RUNAR, ROAR, SALTAR
- Constant Slip Angle Devices: SCRIM, Mu-Meter, SafeDrive, Halliday RT3

Al-Qadi et al. (2002) have given a detailed review on these devices and their historical usage on measuring winter road and airport runways. Most state-of-the-art friction measuring devices are easy to install and operate, and continuously make measurements with the mounted vehicle driven at a normal speed. Also, they work well under a wide range of vehicle speeds, which makes the measurement more convenient and reliable. With the development of the GPS (Global Positioning System) technology, the location of the operating friction measuring device can be accurately tracked; therefore, each single friction reading can be easily associated with its location stamp. The CFM with high-resolution location stamps can be a powerful tool for hotspot targeting, job tailoring, resource deployment, and performance tracking. If the mapping between CFM value ranges and RSC types can be reliably established, it can further improve several important aspects of winter road maintenance works. First, maintenance personnel can straightforwardly apply their accumulated maintenance experiences based on visually distinguished RSC types. Second, decision makers will gain more freedom in transplanting relevant guidelines, standards and policies with different origins. Third, road travelers can easily understand what RSC they are facing. Finally, with this mapping as a translator, estimated RSC types through CFM can be compared, synthesized and combined with RSC measures from RWIS pavement sensors or CFM from other friction measuring devices. For these reasons, a number of studies have been conducted trying to reliably establish this mapping relationship.

2.3.2 RSC Mapping by Mean Friction Level

Most available FB maintenance standards are essentially built on the foundation of some reliable mapping relationship between RSC types and measured mean friction levels (Nordstrom, 1998; Katko, 1993; Finnish National Road Administration, 2001). One example of

Table 2.2: Friction Ranges of RSC Types in Finland Standards

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
RSC Description	bad driving conditions, wet ice, very slippery	icy, slippery	tightly packed snow, satisfactory winter conditions	rough packed ice and snow, good winter conditions	bare wet, not slippery	bare dry, not slippery

this friction to RSC category mapping is from Finland roadway maintenance standards shown in Table 2.2 (Katko, 1993; Finnish National Road Administration, 2001).

Six RSC types are defined based on mean friction level. The lowest two ranges of friction value, 0.00-0.14 and 0.15-0.19 correspond to the two worst types, namely wet ice and icy. When the friction value increases, the RSC becomes better. When the friction measurements are between 0.45-1.00, the category is bare, i.e., there is no snow or ice. It can be found that the descriptions for some types in Table 2.2 are very similar, for instance wet ice and icy. It is really difficult to visually distinguish between these two types, implying that CFM could be more accurate in RSC classification than human observation. Another example is Swedish airport runway maintenance standards (Antvik, 1997), which evaluate the quality of RSC using friction measurement, and the RSC is only described with simple words like "good" or "poor". Certain mean friction ranges are mapped to these simply described RSC types as shown in Table 2.3.

Table 2.3: Friction Ranges of RSC Types in Swedish Airport Runway Standards

RSC Type	Friction Value
good	0.4 and above
medium to good	0.36-0.39
medium	0.30-0.35
medium to poor	0.26-0.29
poor	0.25 and below

Different from Swedish and Finnish standards, a study in Japan addresses the issue of mapping mean friction levels to RSC types which are easier for human visual discrimination

(Tokunaga et al., 2008). As shown in Figure 2.3, the probability distributions of friction measurements for six RSC types are plotted together. Dry and wet types have higher friction levels while types with snow or ice have lower friction levels. Friction ranges of different RSC types may largely overlap although they are visibly different.

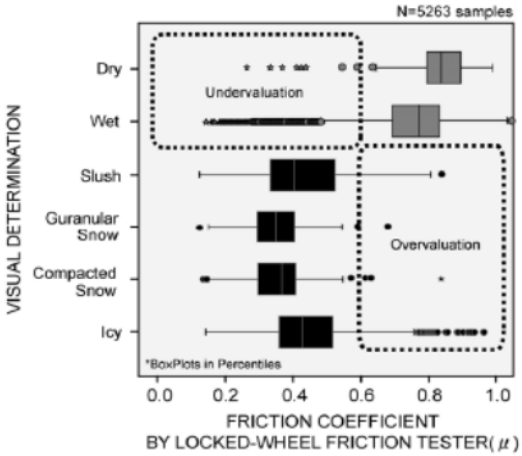


Figure 2.3: Mapping of Mean Friction Levels to RSC Types (Tokunaga et al., 2008)

A survey by Öberg & Gregersen (1991) showed similar overlapped fiction ranges for different visually discriminable RSC types as shown in Table 2.4.

Table 2.4: 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

In summary, three major problems exist in the above RSC classification schemes:

1. For some very similar RSC types, friction levels are different in different classification schemes, which could be due to two reasons. One is that the description for the same

RSC type may vary from study to study, or from country to country. The other reason is that the measuring devices used in these studies may be different, so that even in similar RSC classification schemes, very different measurements could result.

2. The mean friction levels of different RSC types may significantly overlap, which makes it difficult to reliably identify the nature of pavement contaminants solely based on friction measurements.

3. The valid friction value ranges of some RSC types are relatively wide suggesting friction measurements are usually associated with large tolerance intervals. This phenomenon can substantially compromise the reliability and validity of the whole classification scheme. Note that the second problem is directly linked to this problem.

To address the first problem, measurements of standard friction devices working under different mechanisms can be calibrated to each other. Some field studies have shown that the measurements of most friction measuring devices can be inter-calibrated in a simple linear way (Wambold et al., 1995; Vaa, 2001). Although device inter-calibration can be successfully achieved, it is still unlikely to unify so many classification schemes due to the heterogeneity in application context, language and cultural barriers, or just researchers' personal preferences or bias.

The second problem cannot be technically solved until some other explanatory parameters than the mean friction level are introduced into the classification scheme. These new parameters should be able to provide certain information about RSC that the mean friction level cannot sufficiently capture. This information, by itself or combined with the mean friction, should be able to efficiently distinguish between RSC types.

The third problem arises out of the general friction measurement mechanism, as the factors related to road surface textures, contaminants and tires all affect the measurement and cause more uncertainty when directly mapping mean friction levels to different RSC types. These factors are summarized in Table 2.5 (Kummer, 1996; Sandberg, 1997).

2.3.3 RSC Estimation Modelling Using CFM

The second and third problems identified in the last section suggest a need for further research to build more reliable CFM-based RSC estimation models using other explanatory variables in addition to the mean friction level, and it is preferable that these RSC estimation models address the uncertainty of friction measurement. Some studies have attempted to address this challenge by different approaches.

Table 2.5: Factors Influencing Road Friction Measurement

Road Surface	Contaminant	Tire
Macrotexture	Chemical structure	Tread pattern design
Microtexture	Viscosity	Rubber composition
Unevenness/Megatexture	Density	Inflation pressure
Chemistry of materials	Temperature	Rubber hardness
Temperature	Thermal conductivity	Load
Thermal conductivity	Specific heat	Sliding velocity
Specific heat	Film thickness	Temperature
		Thermal conductivity
		Specific heat

Perchanok (2002) conducted a discriminant analysis using three friction measurements, namely, peak resistance (F_p), slip speed at which the peak resistance occurs (V_{crit}), and locked wheel resistance (F_{60}), to classify RSC into categories, like bare wet, bare dry, loose snow, packed snow, slush, and so on. The analysis built a series of linear discriminant functions of F_p, V_{crit} and F_{60} which could optimally discriminate different RSC types. Although the discriminant functions were not explicitly listed in the article, the validation results reported in the article suggest that these discriminant functions using more variables than solely mean friction level could significantly improve the classification accuracy. This study also suggests that certain probability density parameters of friction measurements, which reflected the spatial distribution patterns of snow cover, could be used to improve the discrimination power of the classification model. Variance and skewness of CFM are two such parameters recommended by this study.

In another study by Perchanok (2008), CFMs were treated as time series and transformed from the time domain to the frequency domain so that some hidden properties of CFMs related to RSC types were disclosed. A special frequency-domain exploratory graph, called "periodogram", was used to show the influence of roadside terrain features on large-scale highway spatial snow cover distribution patterns. The periodograms in his study suggest that the spectral density curves of CFMs on snowy RSCs can reflect the characteristics of road side terrain features, such as periodical changes of the elevation of the adjacent terrains. Accordingly, some CFM parameters in the frequency domain can possibly improve RSC classification quality.

Although these two studies have shed light on which potential aggregate features of CFMs could be used to distinguish between different RSC types, they are exploratory in

nature and are short in developing any substantial models. Our study will probe further along the path they pointed out to systematically examine the distinguishing capabilities of those aggregate features. Another implication of these two studies is to use stochastic models to capture the uncertainty of friction measurements, so that the estimation results are more realistic and reliable.

Chapter 3 will revisit some important technical details of these two studies and present our extensive study on RSC type estimation using CFM.

2.3.4 RSC Estimation Using Other Data

Chapter 1 has reviewed some studies which show that RSC is associated with the decrease in traffic speed and volume. However, using these traffic data to deduct RSC type is not a good practice due to the following reasons:

- The changes of traffic speed, density and volume on a slippery road surface are more reasonably considered as the results of drivers' self-adjustment on driving behaviors according to RSC. Therefore the large variance of people's driving habits would cause substantial unreliability if they are used to estimate RSC.
- Driving behavior and traffic flow properties could be affected by other factors simultaneously in addition to RSC, such as weather condition, traffic controls, road geometries, and so on. With so many confounding factors, estimating RSC from driving behavior would face overwhelming challenges.
- In a common sense, it is more logical to infer driving behavior from RSC, but not vice versa. In other words, a specific driving behavior caused by slippery RSC could also be caused by some other factors. Therefore, this inference link is considered unidirectionally valid.

The above reasons can also explain why there is almost no study trying to use vehicle driving data to infer RSC. Lack of study does not mean impossibility. If some driving data are found to be strongly affected by RSC but affected much less than drivers' behavior and other factors, they are potentially useful for classifying RSC.

The driving data, which commonly include speed and longitudinal acceleration, can be recorded in real time by most vehicles in the market. These data can be read from the computer port of the vehicle, called on-board diagnostics port. The commercially available

vehicle telematics devices can export those data and wirelessly transfer them at very low cost.

Besides, a lot of cell phones, especially smart phones, have been equipped with GPS chips and 3-axis accelerometers. GPS chips can not only measure the location of the vehicle, but also can give highly accurate estimations on the Doppler speed of the vehicle (http://en.wikipedia.org/wiki/Doppler_effect). 3-axis accelerometers can measure the vehicle acceleration in longitudinal, lateral and vertical directions all at once. All the data collected by smart phones can be transferred by GSM to computation servers to estimate RSC. Actually most smart phones already have substantial computation capability to process the data and do calculations in situ.

Although we have found no studies addressing the RSC estimation problem with those data, we will try to calibrate a set of RSC classification models using vehicle speed. Those models are only to show possibilities, with no intention of giving physical interpretations or applying to the real world.

2.4 RSC Forecasting

RSC forecasting has for years caught the attention of researchers, maintenance practitioners, standard and policy makers, and roadway authorities all around the world. The highly developed weather monitoring technologies have made weather forecasting models more and more intelligent and powerful, which gives an edge in RSC forecasting. Numerous studies, mostly initiated by governments, have attempted to tackle this problem with a variety of approaches; however, most of the study results are not satisfactory enough to establish reliable models which can be efficiently and reliably used in real-world winter maintenance practice. We will review those studies in the following sections and summarize what has been gained and what is missing.

2.4.1 Factors Affecting RSC Changing

RSC during a snow event could change quickly due to many factors related to meteorological process, maintenance operations and traffic (Klein-Paste, 2008). Meteorological (or "weather" colloquially) factors such as temperature, precipitation, humidity, cloudiness, wind, and so on, are direct causes of formation and melting of snow/ice on road surface. Maintenance operations, such as plowing and salting, are aimed at reducing the amount of snow/ice on road surfaces. Plowing is to remove snow physically while salting is to melt the

snow/ice and break their bondage with the pavement surface. Abrasives such as sand, slag, or cinders are also used to provide an immediate increase in traction. Sanding is usually done at low temperatures where chemical reactions are ineffective, and snow or ice has a strong bond to the pavement. However, traffic can cause this increase to be short-lived by rapidly dispersing the applied abrasives (Ketcham et al., 1996; Minsk, 1998; Perchanok et al., 1991).

The salting chemical agent can be applied in a dry state, or pre-wetted before application. It can also be 100% dissolved in water and applied in a liquid state, which is called Direct Liquid Application (DLA). Dry salt (usually referring to $NaCl$) needs initial moisture before it goes into solution and this results in a lag between the time of application and the time it becomes effective. Also, solid salt particles could be displaced from the pavement by wind and traffic actions even before the solution process starts. Pre-wetted salt can reduce this time lag by providing the initial moisture for the salt to go into a solution and it can also help salt particles to adhere to the pavement. This has the potential to reduce salt usage and be more effective than dry salt since the amount of displaced salt is minimized and initial melting is expedited (Ketcham et al., 1996; Minsk, 1998; Transportation Association of Canada, 2008; Burtwell, 2001).

DLA is used primarily for anti-icing purposes, i.e., the chemical application is done before the start of or very early into a snowfall to ensure that snow or ice are not bonded to the pavement. Chemicals commonly used for DLA include solutions of calcium chloride, magnesium chloride, sodium chloride, calcium magnesium acetate, and potassium acetate (Ketcham et al., 1996; Minsk, 1998; Transportation Association of Canada, 2008).

As discussed in Section 1.2, pavement surface temperature is a critical variable in deciding the amount and type of chemical agent to use. The effect of chemical agents in melting snow or ice decreases as the pavement temperature drops. Table 2.6 lists the lowest effective temperature (Eutectic point) for most common salting chemicals (Ketcham et al., 1996; Minsk, 1998).

Besides application methods and temperature, some other factors, such as timing for application, precipitation type, prevailing weather conditions, pavement structure and material, and so on, also affect the efficiency of salting operations (Perchanok et al., 1991; Ketcham et al., 1996; Minsk, 1998).

The effect of traffic on RSC is three-fold. First, vehicle tires can pack the snow/ice to the road surface causing reduced traction. Secondly, traffic can accelerate the blending of de-icing chemicals with snow/ice through pressure and heat generated from vehicles, improving the de-icing effect of chemicals. Third, traffic can dilute the contaminant solution and decrease the amount of the de-icing chemical on the road surface. The chemical dilu-

Table 2.6: Eutectic Point of Chemicals

Chemical	Eutectic Point (°C)
Calcium chloride ($CaCl_2$)	-29.0
Calcium magnesium acetate (CMA)	-7.0
Magnesium chloride ($MgCl_2$)	-23.0
Potassium acetate (KAc)	-25.0
Sodium chloride (NaCl)	-7.0

tion effect of traffic has been explored and modelled in a number of field studies (Blomqvist & Gustafsson, 2004; Hunt et al., 2004; Lysbakken & Norem, 2008; Marchetti et al., 2008). The findings of these studies indicate that the traffic-induced dilution process is exponential in nature as shown in Figure 2.4. The difference of the two dilution models in the figure suggests that other factors, such as weather and pavement characteristics, affect the way traffic dilutes the road surface contaminant solution. Few studies, however, have systematically investigated these factors, and there is a shortage in both theoretical studies and empirical observations in this aspect. Hence, knowledge about the effect of traffic on salt residual is largely vague and incomplete.

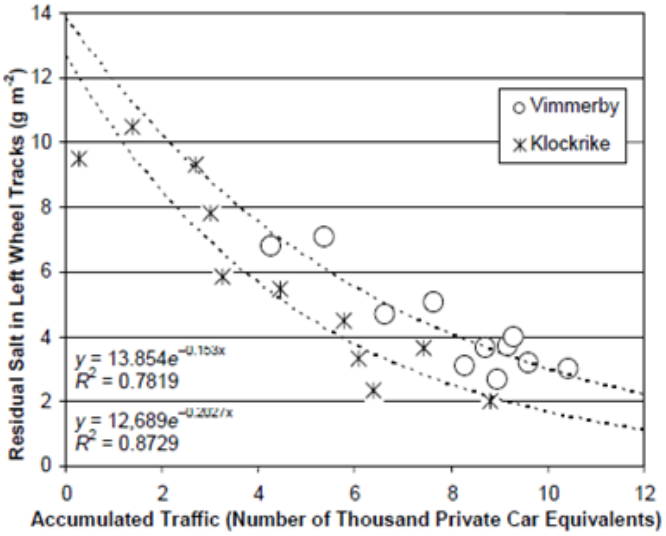


Figure 2.4: Traffic and Amount of Residual Salt at Two Locations (Blomqvist & Gustafsson, 2004)

2.4.2 RSC Forecasting with Physical Approach

As discussed previously, winter RSC is affected by weather, maintenance and traffic factors simultaneously and thus changes with high uncertainty and by a complex mechanism, making it a challenging task to calibrate reliable forecasting models. Previous studies address this RSC forecasting problem with two major approaches: physically based vs. experience based models. We called them *physical models* and *empirical models* hereafter in this thesis.

Physical models explain the whole snow/ice melting process taking place on a road surface using physical mechanisms, which usually include complex numerical equations modelling pavement energy balance/conduction and accumulation of surface contaminants. One typical physical model is the Model of the Environment and Temperature of Roads (METRo), which has been integrated into the MDSS (Maintenance Decision Support System)–a sophisticated winter maintenance decision system developed in the United States (Crevier & Delage, 2001; Pisano et al., 2004; The National Center for Atmospheric Research, 2008). METRo uses roadside weather observations from RWIS stations together with the weather forecast as input and solves the energy balance at the road surface and the heat conduction in the road material to calculate the evolution of important RSC factors, such as temperature and water/ice depth. METRo is composed of three modules: the surface energy balance model, the road material heat-conduction model and the surface water/ice accumulation model. These three models work together to develop forecasts on the pavement contaminant depth. A series of mathematical equations is applied to forecast the energy flux of the pavement material and then the surface water/ice accumulation. The surface energy balance model of METRo (Crevier & Delage, 2001) is used to estimate the net energy flux at/near the road surface and its equation is:

$$R = (1 - \alpha)S + \epsilon I - \epsilon\sigma T_s^4 - H - L_a E \pm L_f P + A \tag{2.2}$$

R: the net solar radiation flux ($joule/m^2$)
S: incoming solar radiation ($joule/m^2$)
α: albedo of pavement surface, reflectivity of sun light (%)
ϵ: emissivity (%)
I: incoming infrared radiation flux ($joule/m^2$)
ϵI: absorbed incoming infrared radiation flux ($joule/m^2$)
σ: the Stefan-Boltzmann Constant ($joule/m^2K^4$)

T_s: pavement surface temperature (K)
$\epsilon\sigma T_s^4$: emitted flux ($joule/m^2$)
H: turbulent heat flux ($joule/m^2$)
L_a: vaporization heat or sublimation heat ($joule/kg$)
E: water vapour flux (kg/m^2)
L_f: heat of fusion of water either for freezing or thawing ($joule/kg$)
P: precipitation rate (kg/m^2)
A: anthropogenic flux, positive flux caused by traffic ($joule/m^2$)

Any time units, such as second, minute or hour, can be applied to Equation 2.2, but should be consistent across all variables. It can be seen that the flux of pavement surface energy on unit area of the pavement is determined by the following factors: solar radiation (S), pavement properties (α, ϵ), pavement surface temperature (T_s), precipitation (L_f, P), air humidity and wind (L_a, E, H), and traffic (A). The net energy flux of road surface (R) estimated by Equation 2.2 is conducted downward along pavement's ground profile as a function of time given the heat conductivity and capacity of the road material. This is done in the road heat-conduction model whose equation is not listed here.

The net energy flux calculated by Equation 2.2 can then be used to model the change of amounts of water and ice on the pavement surface. The road surface condition model in the surface water/ice accumulation model consists of Equation 2.3 and 2.4.

$$\frac{dW_l}{dt} = P - E + \frac{R - G_1}{L_f} - r \tag{2.3}$$

$$\frac{dW_s}{dt} = P - E - \frac{R - G_1}{L_f} - r \tag{2.4}$$

W_l: amount of liquid water at road surface (kg/m^2)
W_s: amount of snow/ice at road surface (kg/m^2)
t: time
G_1: downward heat flux to the lower layer of the road surface contaminant ($joule/m^2$)
r: runoff (kg/m^2)

Again, second, minute or hour, are all applicable for the time unit t, but should be consistent for all variables. The change in the amount of liquid water and solid snow/ice is

affected by precipitation (L_f, P), energy balance (R,G_1) and factors affecting runoff speed (r). The values of input variables in Equation 2.2, 2.3 and 2.4 are determined by accurate measurement of several pavement characteristics and important meteorological factors. Some of those factors, such as solar radiation, incoming infrared radiation, turbulent heat, anthropogenic flux, water vapor, and so on, are either difficult or costly to measure for current RWIS installations. When using those equations to do forecasting, the values of all those input variables should ideally be forecasted instead of monitored. Forecasting values for so many variables is a very challenging task. Using the input variables forecasted in this way to further forecast RSC would face a lot of technical difficulties in practice, one of which is that, the forecast errors from so many variables will accumulate, interact and make the resulting RSC forecast highly unreliable.

Moreover, METRo needs accurate road surface heat flux and solar radiation information to give reliable surface temperature and water accumulation forecasts. Such information is, however, only available at some fixed spots over the road network where RWIS stations are located. As a result, the spot forecasts made by METRo are quite questionable when they are applied to the road sections far from the RWIS stations.

Furthermore, although the surface temperature forecast by METRo is reliable according to the testing results (Crevier & Delage, 2001), it does not estimate chemical concentration of the road surface contaminant and its effect on RSC. This limits the application of METRo into a long-term (usually 24 hours) maintenance resource planning tool. If some major maintenance operations, like plowings or saltings, are conducted, the METRo forecast would no longer be trustworthy, which stems from the limitations in its working mechanism.

In order to enhance METRo to be capable of short-term RSC forecasting and thus used as a maintenance decision making tool, the RSC forecasting module of MDSS simply designates some simple linear dilution function to the de-icing chemical so that the chemical concentration of road surface contaminant can be estimated (The National Center for Atmospheric Research, 2008). The general form of this linear dilution function is:

$$Chem = (1 - tFactor) * (AppRate + ResChem + ChemInSol) \tag{2.5}$$

$Chem$: salt on pavement surface after current salting operation
$tFactor$: fraction of salt lost from the road surface due to traffic
$AppRate$: applied salt of current salting operation
$ResChem$: solid salt residual before current salting operation
$ChemInSol$: salt in pavement surface solution before current salting operation

When Equation 2.5 is used recursively at a short time interval such as one hour, the amount of salt on the pavement surface exponentially decreases with accumulated traffic with $tFactor$ defined as hourly decreasing rate of the de-icing chemical. By integrating this chemical concentration model with METRo, MDSS forecasts on pavement surface temperature, contaminant depth, and chemical concentration over a certain time horizon. It should be noted that the current version of MDSS is deterministic in nature and does not consider the uncertainty of the RSC changing process.

Another similar physical pavement temperature and surface condition forecasting model was developed by Tokunaga et al. (2008). The model uses a similar set of water balance equations and variables as METRo to forecast pavement temperature and RSC. The only significant difference is that snow and ice depths are modelled separately; therefore, there are three water balance equations instead of two as in METRo. Obviously, this set of models has almost the same limitations as METRo, including no consideration of maintenance operations, requiring a large amount of weather forecasting information as input, not addressing prediction uncertainty, and difficulties to apply to locales far from RWIS stations, which are theoretically shared by all possible physical models.

2.4.3 RSC Forecasting with Empirical Approach

To address the issues of physical models, empirical models have been developed in some field studies. Instead of using complex first-principle style numerical equations, empirical models are usually built with data-mining approaches. The real-world information of weather, maintenance and traffic are archived and stored in database systems, and systematically analyzed for relevant patterns. The isolated patterns can as simple as lists of rules, or as complex as structural linear or non-linear statistical functions.

The data for calibrating empirical models usually come from RWIS observations without the unrealistic requirements on the input parameters as in physical models. Also, the number of input variables appearing in an empirical model is usually relatively small compared to most physical models, which makes it much more convenient to be calibrated and applied. In addition, empirical models are inherently capable of associating the forecasting result with some uncertainty measures, such as occurrence frequency or probability. This makes it convenient to capture the uncertainty nature of the RSC changing process.

Some previous studies have resulted in some very simplistic empirical models–rule tables, which are just look-up tables with a list of rules for RSC forecasting. A simple rule table was developed in Sweden using RWIS data (Norrman, 2000). The RWIS data included real-time observed weather information, such as precipitation, air temperature,

surface temperature, relative humidity and wind speed. Some example rules are listed in Table 2.7.

Table 2.7: RSC Estimation Rule-Table (Norrman, 2000)

Precipitation	T_{air}	T_{road}	RSC Estimation
yes	>0°C	<0°C	It is probably rain or sleet that freezes on to the road surface. The probability of snowfall decreases rapidly with increasing T_{air}.
yes	<0°C	<0°C	It is probably snowfall onto a frozen road surface.
yes	<0°C	>0°C	It is probably snowfall or sleet forming slush on a warm road surface.

The estimation rules in the table are consistent with basic physical principles and easy to explain and use. If the real-time observed weather information is substituted by a short-term weather forecast, this rule table can be then used for RSC forecasting. The validation results showed poor estimation accuracy, which is not surprising because the rules were built by directly borrowing findings of previous field studies and experts' opinions without any formal statistical analysis, necessary local calibrations, or considerations of complex interaction effects of weather factors (Norrman, 2000). The rules are deterministic in nature and do not consider winter maintenance and only consider traffic effect in a very limited manner. This type of rule tables can hardly be considered as models and can only give rough guesses without any uncertainty measures.

A study in Japan tried to make forecasts according to the occurrence probability in observed field data (Yamagiwa et al., 2004). RSC is classified into 13 types, which are assigned to three severity ranks as shown in Table 2.8. The RSC forecasting is given based on the occurrence frequency of each surface type under similar weather conditions. The forecast frequencies are grouped by rank, and the probability of each rank is calculated by simply summing up the frequency percentages of all RSC types in the rank. This rule-based system is calibrated for the forecasting horizon of 6-14 hours, and can only be used for maintenance operation planning. Probability distributions for the three ranks can serve this purpose very well, because the long-term snow/ice danger can be effectively identified with this rank resolution. Obviously, the rules listed in a tabular format cannot handle situations with a higher level of complexity, especially the one caused by maintenance operations. Therefore, both of the previous rule-table-based forecasting systems are incapable of making reliable forecasting on RSC under de-icing operations.

Table 2.8: RSC Types by Rank (Yamagiwa et al., 2004)

	RSC Type	Rank
1	very slippery compacted snow	1-very slippery road surface
2	very slippery ice sheet	
3	very slippery ice film	
4	ice sheet	2-slippery road surface
5	ice film	
6	powder snow on ice	
7	granular snow on ice	3-relatively easy driving road surface
8	compacted snow	
9	powder snow	
10	granular snow	
11	slush	
12	wet	
13	dry	

In order to simultaneously incorporate the effects of weather, maintenance operations and traffic, a conceptual short-term RSC forecasting model has been proposed by the Swedish National Road and Transport Research Institute (VTI in short) (Wallman, 2004; Möller, 2008). Following the similar empirical-style modelling philosophy to the Japanese study, weather, maintenance and traffic data are collected, archived and integrated with RSC data at one-hour time intervals. Cross-classification rules by important conditional variables of weather, maintenance operations and traffic are used together with current RSC to forecast RSC of the next time stamp. Figure 2.5 shows the model framework.

To incorporate the complexity and uncertainty characteristics of the RSC changing process, the VTI model forecasts in the form of multiple possible RSC types instead of just one single type. Some important input parameters for this RSC model are:

- road condition at t
- amount of residual salt at t
- weather condition from t to $t+1$, including air temperature, road surface temperature, dew point temperature, precipitation type and amount, wind speed, and so on.
- traffic volume from t to $t+1$
- maintenance operations taking place during the period from t to $t+1$

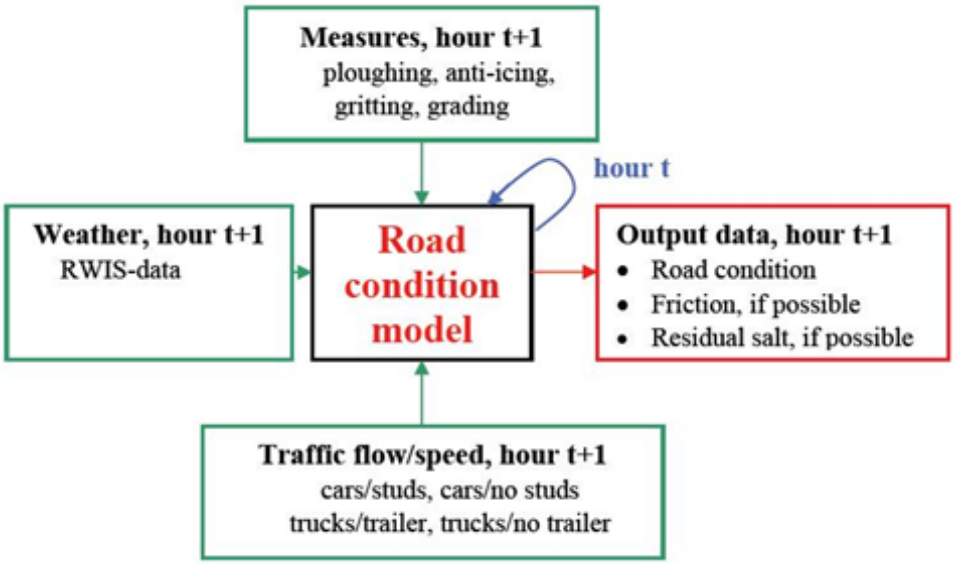


Figure 2.5: Framework of the VTI RSC Model

The forecasting output includes possible RSC types, amount of residual salt and friction level at $t + 1$. RSCs are classified into five types as follows:

- dry bare ground
- moist or wet bare ground
- hard-packed snow or thick ice
- black ice or hoar-frost
- loose snow or slush

The model is still in the data accumulation stage and under development, so the systematic evaluation on its performance has not yet been available. Although the VTI model is intent on short-term forecasting with one hour time intervals, forecasts with longer horizons can be obtained by chaining the short-term forecasts of each time step sequentially. One major limitation of this model is that the design of the model does not explicitly give probability estimation for the forecasted RSC at each time step as Yamagiwa's model has done.

In summary, the empirical models reviewed above share the following characteristics:

- They are all based on large amounts of archived data, which usually come from multiple sources and need significant amount of processing and integration work.
- The data are usually analyzed using data mining techniques by people with considerable statistical knowledge. The resulting models are stochastically rooted, capable of producing probabilistic forecasts.
- The models are not built based on first-principles and are not intended to describe the process in a physical way. Instead, they summarize the observations into rules, which could be some simple look-up tables or possibly some highly complex regression equations.

The above characteristics reflect the empirical philosophies of those studies, which reflect some important properties of the RSC changing process. As for the particular modelling techniques, the following characteristics can hardly win common consent.

First, in all available models, RSC is classified into discrete types, which are mostly defined by descriptive language. RSC is either observed and reported based on visual observation or RWIS measurement. Visual observation lacks objectivity and repeatability compared to RWIS measurement. Even for RWIS measurement, however, the categorical RSC measurement still faces the challenge related to classification resolution. As an extreme example, if the RSC is classified into only two types–bare or icy, the RSC could jump from one type to the other by very few times even within a long snow event. In this case, the RSC forecasting becomes too obtuse making it of little use. If the RSCs are further classified into more types, more RSC jumps can be identified and the forecasting becomes more sensitive thus more helpful for maintenance practice. But how many types should the RSC be classified into? Are there any better alternative options other than discretizing RSC? Is current sensing technology able to support such alternatives?

Second, all available models are built upon a single time interval Markovian assumption (Lefebvre, 2007), i.e., the RSC at the next time stamp is completely conditional on the RSC at the current time and the interventions happening between the current and the next time stamp. The interventions refer to the influence exerted by weather, maintenance and traffic. The Markovian assumption in this context is reasonable only when all critical influencing factors are measured and considered in the resulting model. If some important factors are missing, the state of the process the model is trying to capture is incomplete and the Markovian assumption is violated. Take the salting operation for example; the VTI model only considers the amount of salt residual, and overlooks its mixture state with road surface contaminant. This could make its short-term RSC forecasting highly unreliable

especially after salting operations, because the mixture state could essentially determine the melting speed of the road surface contaminant.

2.5 Summary

In this Chapter, major studies on RSC estimation and forecasting are reviewed. Some possible problems, limitations or shortcomings of them are identified and discussed.

To improve the reliability of RSC estimation, more features need to be extracted from CFM to capture the spatial distribution patterns of snow coverage. Moreover, the uncertainty associated with RSC estimation needs to be addressed.

RSC forecasting needs to address the high uncertainty and complex nature of the RSC changing process. Models based on physical theories face the challenges of obtaining measurement for a large number of factors. Empirical models face fewer challenges in data availability but the model framework and modelling techniques need to be given more consideration.

In the next two chapters, the RSC estimation and forecasting problems are defined and tackled under new analysis frameworks based on our own recognitions and understandings. In addition, the modelling results and major findings are discussed and concluded.

Chapter 3

RSC Estimation

A major research objective of this thesis study is to systematically investigate the relationships between winter RSC types and CFM, enhance existent RSC classification schemes, and calibrate RSC classification models. Besides CFMs, vehicle driving data are also within our research scope, especially the available vehicle speed data. In this chapter, CFM and speed data are separately integrated, explored, analyzed and introduced into the RSC type estimation models under a new study framework.

3.1 Problem Definition and Modelling Framework

As discussed in Chapter 2, there are several major challenges in obtaining reliable mapping from CFM to specific RSC types. The first and foremost is that the friction levels corresponding to different RSC types may overlap, implying that the direct mappings from CFM value ranges or means to RSC types, as already tried in many previous studies, may not work well in many situations. The second challenge is that many factors besides RSCs affect the CFM, and the resulting uncertainty and complexity make it difficult to infer RSCs based only on mean friction level, especially in a deterministic manner. The third challenge is the lack of lateral coverage of CFM, which adds more uncertainty to the resulting RSC estimation.

To address these three major challenges, a probabilistic modelling framework incorporating multiple aggregate measures of CFMs is proposed, as shown in Figure 3.1. The first challenge is addressed by considering additional statistical measures that can be derived from CFMs. Those statistics, particularly features extracted from probability distribution

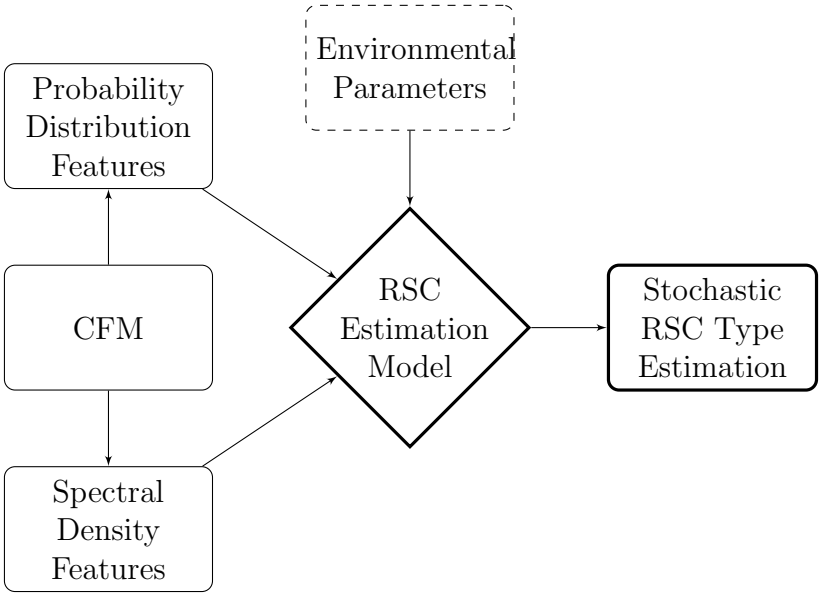


Figure 3.1: RSC Estimation Model Framework

and spectral density of CFMs, are intended to represent the spatial variation patterns of CFMs along the measuring route. Section 3.3 explains in detail why these features can enhance the RSC classification quality. To address the second and third challenges, the proposed framework has adopted a stochastic modelling approach which estimates the probability distribution of all possible RSC types. This approach is considered more appropriate to calibrate models based on data of high uncertainty or variance.

The focus of this framework is to estimate the RSC types based on CFMs and other relevant data, especially environmental parameters. The RSC types are several major different RSC states defined by visual observation and judgment. The classification features are extracted from the CFM or the environment. This is a typical classification problem, and is usually solved through the following steps:

1. Define the set of concerned types (classes, states or categories) of the studied object, which in this study is the winter RSC.

2. Explore available data and find/extract relevant classification features.

3. Calibrate and cross-validate a single or multiple structured classification model.

The framework emphasizes the importance of considering environmental parameters, like air temperature, tire surface characteristics and so on, because they directly affect

friction measurement as discussed in Chapter 2. Unfortunately, in our collected data, the only available environmental parameter is air temperature and a large proportion of the test runs did not collect it. For this reason, environmental parameters are not considered in our subsequent modelling process. Ideally, these environmental parameters should be simultaneously measured with CFMs, and directly introduced into the modelling process.

3.2 Data Collection

The CFM and speed data were collected on a section of Highway 417 in eastern Ontario, which has asphalt pavement with a total length of 40 km. Friction data were collected using a fixed-slip-ratio continuous friction measuring device called Traction Watcher One (TWO) (http://www.pon-cat.com/en/Traction-Watcher-One/). The TWO trailer was attached to the rear of the tow vehicle with its slip speed ratio fixed at 16.6%. During data collection, the tow vehicle was operated in the driving lane with the measuring tire running in the left wheel track. The sampling frequency is one reading per 10 meters, i.e., the device outputs one friction measurement for every ten meters the driving vehicle covers. The driving speed was measured at the same time and reported at the same spatial frequency.

The data were collected during the winter seasons of year 2007 and 2008. The lengths of test runs varied between 2km and 30km. The test runs were scheduled carefully to cover different stages of the snow storms with a wide variety of snow coverage and environmental conditions. In addition to friction measurement and driving speed, vehicle locations were recorded using a GPS device. A video camera was also operated to record the RSC, and the video images were used to determine the type of RSC at corresponding road segments on the testing route and served as the basis for further supervised classifier calibration. According to the recorded video images, the RSC is assigned into the following six visually discernible types:

1. Type 0: bare dry
2. Type 1: bare wet
3. Type 2: thin snow cover
4. Type 3: slushy snow cover
5. Type 4: partially snow covered
6. Type 5: mostly snow covered

Type 0, 1, 4 and 5 are easy to distinguish using only video images. Type 3 is much like a mid-type between Type 2 and 4. Therefore, the boundaries between Type 2 and 3, and Type 3 and 4 are quite ambiguous and subject to observers' personal judgment. Typical images of Type 2, 3 and 4 are shown in Figure 3.2. Although sometimes it is difficult to visually discriminate between these three types, some guidelines can be followed. For example, slushy snow means chemical treated, wet and traffic trodden dirty snow. Partially snow covered means significant amounts of dry snow distributed unevenly across the lane.

(a) Thin Snow Cover (b) Slushy Snow Cover (c) Partially Snow Covered

Figure 3.2: Video Images of RSC: Type 2, 3 and 4

Thin snow cover means slushy snow of very small depth and coverage, and usually occurs just before the BP condition is regained.

Besides the easy visual discernibility, another reason that these six types are defined is that the order of their appearance during a snow storm period corresponds to the typical RSC improvement process with maintenance operations. Usually, when the pavement is fully snow covered due to snowfall, plowing and/or salting operations are conducted, and the RSC usually changes from Type 5 to 4, 3, 2 and 1 sequentially, and finally to Type 0, which is bare dry. In other words, these six types are the major critical states the highway pavement condition experiences in a typical BP recovering process. Accurate estimation of these states is beneficial to the decision making and performance measure of maintenance operations.

3.3 Feature Extraction

To estimate RSC by friction measurement, certain "relevant" features (or statistics) need to be extracted from CFMs as classification parameters. Several previous studies have suggested that specific parameters extracted from the probability distribution and spectral density of CFMs could serve as these "relevant" features and enhance the discrimination power of the classifiers (Perchanok, 2002a, 2008). Although, these parameters are all defined in a statistical perspective, they all have solid physical interpretations. The following sections systematically explore and analyze how those parameters can enter the model to improve the classification accuracy.

3.3.1 Probability Distribution Parameters

Perchanok (2002a) suggested that some aggregate features of CFMs, e.g. variance and skewness, are correlated with RSC types, thus are potentially able to enhance RSC estimation accuracy. In his study, continuously collected point-wise friction measurements were aggregated by a fixed spatial interval to obtain variance and skewness. It was shown that when average friction measurements were extremely low or high, variance was very small. A low average friction reading coupled with small variance was a sign of a surface with fully covered snow/ice. Large variance usually corresponded to a partially snow/ice covered surface, due to the heterogeneity of tire-pavement contact caused by snow/ice contaminants. It was also shown that mean friction did not represent a continuous variation in surface characteristics. A low mean could result from a combination of many small values and a few large values while a high mean could come from many high values and a few much lower values. Therefore, low snow/ice coverage can correspond to a strong left-tailed frequency distribution and high snow/ice coverage to a less left-tailed distribution. At the middle range of mean friction, the numbers of high and low readings were balanced and the frequency distribution was not skewed. These findings suggest that the parameters reflecting the shape of a probability distribution, like skewness and variance, could contain additional information on snow/ice coverage distribution pattern in space.

Equation 3.1, 3.2 and 3.3 define the sample mean, standard deviation (representing variance) and skewness.

$$\bar{f} = \frac{\sum_{i=1}^{n} f_i}{n} \tag{3.1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (f_i - \bar{f})^2}{n-1}} \tag{3.2}$$

$$\gamma = \sqrt[3]{\frac{\sum_{i=1}^{n} (f_i - \bar{f})^3}{(n-1)s^3}} \tag{3.3}$$

f_i: i^{th} friction measurement
$\bar{f}$: mean friction
n: total number of friction measurements
s: standard deviation of friction measurements
γ: skewness of friction measurements

Although only standard deviation and skewness are mentioned here, some other probability distribution features, like kurtosis, could also have certain degrees of RSC discriminating power. It is impossible to test every probability distribution parameter; however, one should focus on those which have physical interpretations.

3.3.2 Spectral Density Parameters

In the modelling framework shown in Figure 3.1, it is proposed to incorporate spectral density parameters of CFM in the RSC estimation model. This is motivated by some positive results from initial efforts by Perchanok (2008), which showed that the spectral density patterns of CFMs can reflect the spatial distribution patterns of the pavement contaminants. The concept of spectral density stems from time-series analysis works focused on the frequency domain. As CFMs are collected on highway surfaces continuously and evenly in distance, the CFMs collected in a single test run can be treated as a time series and certain time series analysis techniques can be applied. Spectral analysis is one such technique, which can be applied to the series data in the frequency domain. The observed series of values is considered as the superposition result of a number of underlining periodic processes with complex sinusoidal forms of different amplitudes and periods. The periods of those underlying processes correspond to the series of discrete frequency values. More often than not, it is difficult to reveal important data patterns from the traditional time

domain analysis. In such cases, frequency domain analysis could be a handy and efficient alternative (Zadeh, 1953; Jenkins & Watts, 1968).

Frequency domain analysis is a major technique of time series analysis and has been used in a wide variety of science and engineering fields (Shumway & Stoffer, 2006). For our RSC estimation problem, the periodicity and the period-specific amplitudes in the data series could reflect the spatial distribution patterns of the snow coverage. As a result, the "relevant" frequency domain features could potentially significantly increase the estimation accuracy and reliability. Based on the preliminary analysis of the CFM data, we could find out what those "relevant" features could be and how they should be mathematically extracted from raw CFMs. Before starting the formal discussion about feature extraction from spectral distribution patterns, several basic concepts in frequency domain analysis are introduced first, including auto-covariance function, auto-correlation function, spectral density function, and periodograms.

Basic Concepts

The sample **auto-covariance function** of a time series TS $(x_1, x_2, \ldots, x_t, \ldots, x_n)$ can be calculated as

$$\gamma(h) = cov(x_{t+h}, x_t) \tag{3.4}$$

$$= \frac{1}{n}\sum_{t=1}^{n}(x_{t+h} - \mu_x)(x_t - \mu_x) \tag{3.5}$$

$\gamma(h)$: covariance of two x values in TS with time distance h
μ_x: mean of TS
n: number of observations in TS

$\gamma(h)$ measures to what degree the pairs of the values within TS are related to each other using a function of their distance h. The sample **auto-correlation function** (ACF in short) is derived from the auto-covariance function and can be calculated as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \tag{3.6}$$

$\gamma(0)$ is the covariance of x with distance 0, which is exactly the variance of TS. Therefore, the auto-correlation function $\rho(h)$ is the auto-covariance function scaled by the total

sample variance. The range of $\rho(h)$ is from -1 to 1. The absolute value of $\rho(h)$ represents the strength of the auto-correlation, where the sign represents the direction of the auto-correlation. If TS is "stationary" and $\rho(h)$ satisfies

$$\sum_{h=-\infty}^{\infty} |\rho(h)| < \infty \tag{3.7}$$

then it has the representation

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} f(\omega) d\omega \tag{3.8}$$

where

$$h = 0, \pm 1, \pm 2, \ldots$$

as the inverse transform of the **spectral density function**, which has the representation

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} \tag{3.9}$$

where

$$-1/2 \leqslant \omega \leqslant 1/2$$

In other words, $f(\omega)$ is the discrete Fourier transform of $\gamma(h)$ and these two measurements of TS consist of a so-called "fourier transform pair". ω is the frequency, whose range is from $-1/2$ periods per data point to $+1/2$ periods per data point. Putting $h = 0$ in Equation 3.8 yields

$$\gamma(0) = var(TS) = \int_{-1/2}^{1/2} f(\omega) d\omega, \tag{3.10}$$

i.e., the total variance is the integrated spectral density over all of the frequencies. In probability theory, the fourier pair like $\gamma(h)$ and $f(\omega)$ are considered to contain the same information, which is expressed in two different ways. Shumway & Stoffer (2006) explained this concept in this way:

"The autocovariance function expresses information in terms of lags, whereas

> the spectral density expresses the same information in terms of cycles. Some problems are easier to work with when considering lagged information and we would tend to handle those problems in the time domain. Nevertheless, other problems are easier to work with when considering periodic information and we would tend to handle those problems in the spectral domain."

The idea of extracting spectral domain features from CFMs and using them to classify RSC has the same philosophy as Shumway & Stoffer (2006) have said. In the next chapter, we will be dealing with another major research problem, RSC forecasting, which is easier to work with by considering lagged information in the time domain. Therefore, the set of analyzing and modelling methods used in the next chapter is very different from this chapter, although they are all under the same roof: "time series analysis".

A common frequency domain graphical exploratory tool is **periodograms**, the simplest form of which is just a plot showing $f(\omega)$ against ω. As $f(\omega)$ is an even function, i.e. $f(\omega) = f(-\omega)$, usually periodograms only show $f(\omega)$ where $0 \leqslant \omega \leqslant 1/2$.

Periodograms

Perchanok (2008) used periodograms of CFMs on winter snowy road surfaces to analyze the influence of roadside terrain features on large-scale highway spatial snow cover distribution patterns. The study showed that the shape of spectral density curves of CFMs can reflect the characteristics of road side terrain features, particularly periodical changes of the elevation of roadside terrains. The study is focused mostly on the spectral density (or simply "spectra") in relatively low frequency ranges, i.e., long period ranges, which is suitable to study friction variations at a scale of maintenance route length (10∼40km).

Following a similar approach, we have conducted a preliminary examination on the periodograms of Type 0, 1, 2, 4, 5 based on 34 test runs randomly selected from available CFM data. These 34 CFM series are treated as 34 time series (or spatial series), each of which has 250 continuous point-wise measurements, i.e., covering 2.5km. Fourier frequencies (ω) for all runs are discretely set to $k/250$ with $k = 0, 1, \ldots, 125$. That is, the range of ω is from 0 to 1/2, and the unit of ω is *periods/measurement* or *periods/datapoint*. Our primary examination of the periodograms suggests that the CFM series of all runs are "non-stationary" with some drastic low-frequency spectra. In mathematical terms, "stationary" means

1. the expectation of all x_t in a time series is an equal constant, and

2. the covariance of any two observations in this time series can be determined solely by their distance.

This definition of stationary is usually called "secondary stationary" in most of the literature. The terms "stationary", "stationarity" and "non-stationary" previously mentioned in this thesis are all based on this definition, i.e., secondary stationary.

The non-stationarity characteristics of our CFM data, especially the existence of long-period (or low-frequency) trends (level changes), can be interpreted by the distribution patterns of snow cover along the test route. On a bare dry surface, this level change of friction could be caused by the natural variation of the pavement surface texture. On a snowy surface, this level change could be caused by the variations in snow depth or snow coverage along the test lane. Because the major concern of this study is the characteristics that distinguish different types of road surface contaminants, the spectra in both low and high frequency ranges must be fully considered. However, the low-frequency level changes tend to obscure the appearance of spectra at the high frequency range. A basic approach to solve this problem is to apply a "high-pass" filter to the raw time series data, in other words, to retain or pass high-frequency variation and attenuate low-frequency variation, so that the long-term trend of the series is removed and the mean of the filtered series is forced to be zero. A common high-pass filtering operation is "first-order differencing" or "first differencing" in short, mathematically defined as:

$$\Delta x_t = x_t - x_{t-1} \tag{3.11}$$

Figure 3.3 shows the original and the first-differenced CFM series of a test run. The overall mean of the original CFM series is larger than zero, and there is an obvious long-term variation of the mean level. After first-differencing, the mean becomes zero and the whole series looks to be randomly fluctuating around zero, i.e., much more stationary. The effect of differencing on the shape of the periodogram curve is shown in Figure 3.4. By first differencing, the spectra in the extreme low frequency range are largely attenuated and the spectra in the relatively high frequency range are kept unchanged or strengthened moderately. In this way, the spectra information contained in both high and low frequency ranges is revealed in a balanced manner in terms of spectral magnitude.

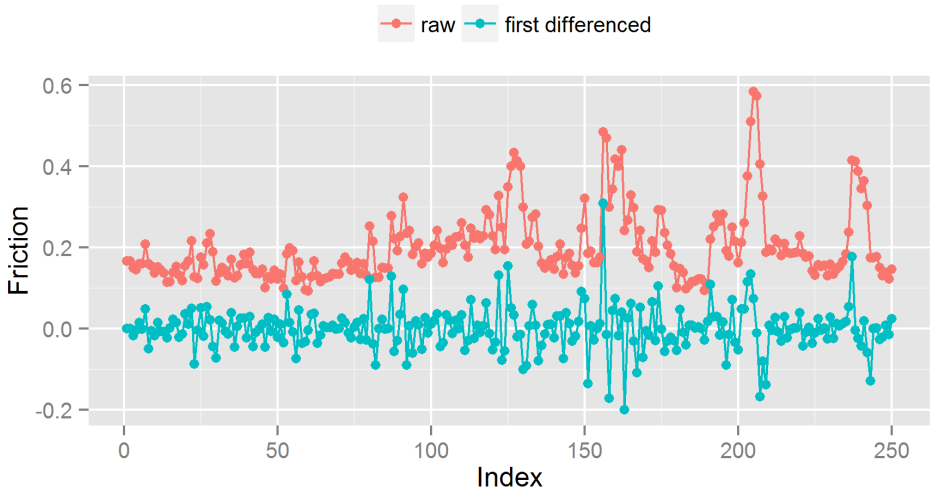


Figure 3.3: Raw and First-Differenced CFMs

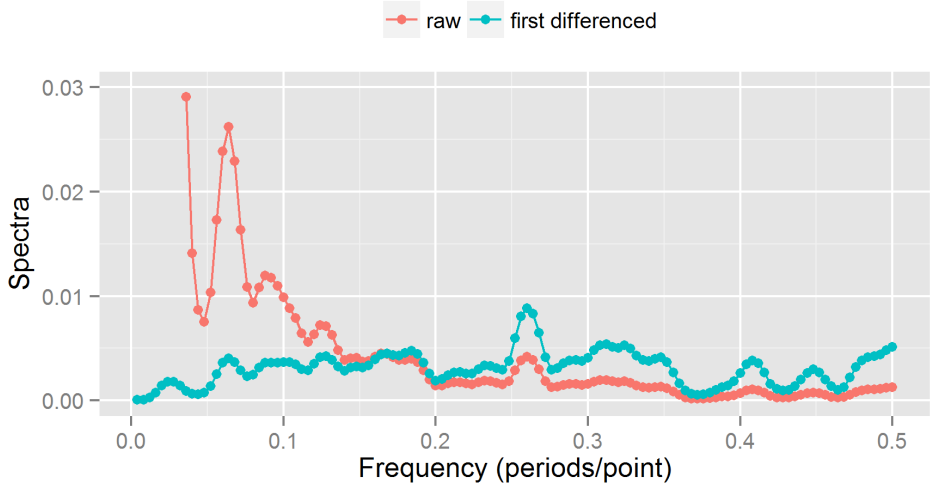
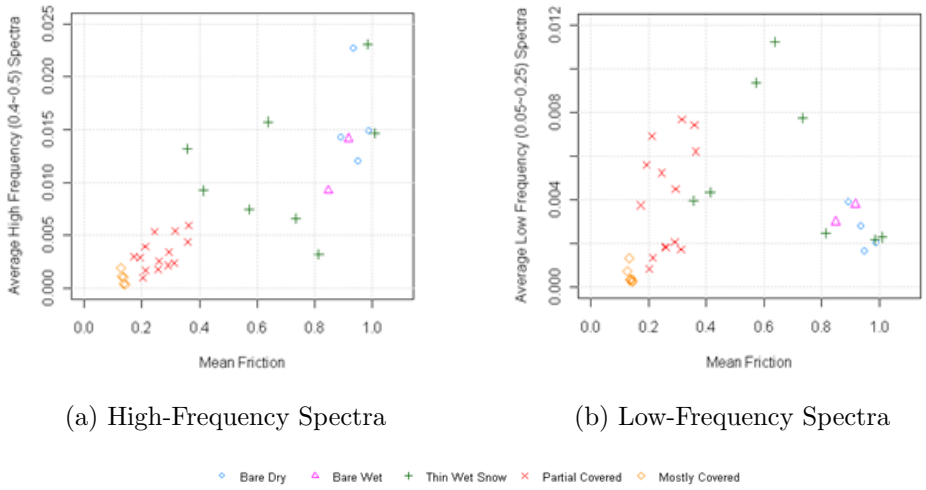


Figure 3.4: Periodograms of Raw and First-Differenced CFMs



(a) High-Frequency Spectra (b) Low-Frequency Spectra

Figure 3.5: Spectra and Mean Friction by Frequency Range

Preliminary Findings

In the preliminary exploratory analysis, after first differencing is applied to the CFMs of all 34 test runs, the mean spectra of two frequency ranges (high frequency: 0.4∼0.5 periods/point; low frequency: 0.05∼0.25) are used to illustrate the relation between the spectral patterns and the overall snow coverage levels (Feng & Fu, 2009). Mean frictions of the runs are plotted against their mean high-frequency and low-frequency spectra respectively in Figure 3.5.

Each data point in Figure 3.5 represents one single test turn. The mean friction is considered as a surrogate for the level of snow coverage; therefore, this figure presents the correlation patterns between snow coverage and spectra in high and low frequency ranges. Figure 3.5 (a) shows that the high-frequency spectra are positively correlated with mean friction in a linear manner implying that the area of the exposed bare pavement is positively related to the high-frequency spectra. Although the high-frequency spectra of the thin wet snow surface vary widely, most of them are greater than those of the partially snow covered surface and smaller than those of the bare dry surface. Based on this observation, it can be assumed that the main source of the high-frequency spectra is the bare pavement surface not covered by snow/ice.

When a significant portion of the road surface is covered by snow/ice with a low mean

friction (left part of Figure 3.5 (b)), low spectra result due to the low proportion of bare pavement. On the other hand, extremely low snow coverage leads to the low spectra of bare surfaces as shown in the right part of Figure 3.5 (b). When snow coverage is in the middle range, the low-frequency spectra could be very high due to the long period changes in friction level caused by the alternation between large patches of snow/ice and bare surfaces. Additionally, surfaces with very different degrees of snow coverage could result in similar low-frequency spectra, which implies that low-frequency spectra are not a direct indicator of snow coverage.

Table 3.1 summarizes the correlation patterns between CFM spectra and friction levels.

Table 3.1: Spectral Comparison by Frequency Range

RSC Type	Low-Frequency Spectra	High-Frequency Spectra
Bare Dry	low	high
Bare Wet	low	high
Thin Wet Snow Cover	wide range	high
Partially Snow Covered	high	middle
Mostly Snow Covered	low	low

Although not tested in any strictly controlled experimental set-up, the following propositions can be made, which are strongly supported by our preliminary exploratory analysis:

1. The major source of high-frequency spectra of CFM is bare pavement; thus the less the snow coverage is, the higher the high-frequency spectra are. Therefore, the high-frequency spectra of mostly snow covered condition are the lowest, while bare dry and wet conditions highest.

2. As the source of low-frequency power of CFM is mainly from the spatial alternation of snow and bare pavement along the longitudinal direction of the measured lane, the partially snow covered condition produces strong low-frequency spectra. Both bare conditions and mostly-snow covered conditions lack this alternation and thus produce much lower low-frequency spectra.

3. The mostly snow covered condition has the lowest spectra at all frequency ranges.

4. It is difficult to discriminate bare dry and bare wet conditions solely by spectral information.

3.4 Data Preparation

The features from the probability distribution and spectral density of CFMs are all aggregate measures, i.e., they are derived from a set of raw single CFMs and reflect certain summary characteristics of the set. For the purpose of this chapter, a CFM set should be the CFMs consecutively collected within a certain length of the roadway. This length is the spatial aggregation interval of the CFM samples, which should be determined with the following considerations:

- RSC estimation of a high spatial resolution is preferable to maintenance personnel or decision makers. As the friction data can be collected automatically, transferred to servers in real time, and presented on maps with accurate location, RSC estimations with as high as possible spatial resolutions can give the most accurate RSC information. Therefore, the spatial integration interval should be minimized conditional to other possible limitations.
- Most aggregate measures need certain minimum sample size in order to obtain reliable estimation results. To get a reliable estimation of skewness, a large sample size is needed as skewness is of a higher order moment than standard deviation. Some prior studies suggest that a sample size as large as 50 may not be enough to get the reliable skewness estimation (http://www.spcforexcel.com/are-skewness-and-kurtosis-useful-statisticsquick).
- The amount of available data poses a strong limit to the aggregation interval. For this study, only the segments of CFMs of a single RSC type are used in data aggregation. If a longer interval, e.g. 2km, is used, we have to exclude the CFMs of all test runs shorter than 2km or without any pure segments longer than 2km.

3.4.1 Sample Size

The CFMs are divided into parts of certain length, which is the aggregation interval. Each part becomes a single sample, whose aggregation features are calculated. Figure 3.6 shows how the sample size of each RSC type decreases when the length of the integration interval increases. Note that there is one friction reading for every 10 meters, so the number of friction measurements in each sample is the distance in meters on X-axle divided by 10. Table 3.2 lists the sample size by RSC type and integration interval.

From Figure 3.6, it can be seen that the sample sizes of all RSC types decrease exponentially with the increasing aggregation interval length. Table 3.2 shows that Type 0 and

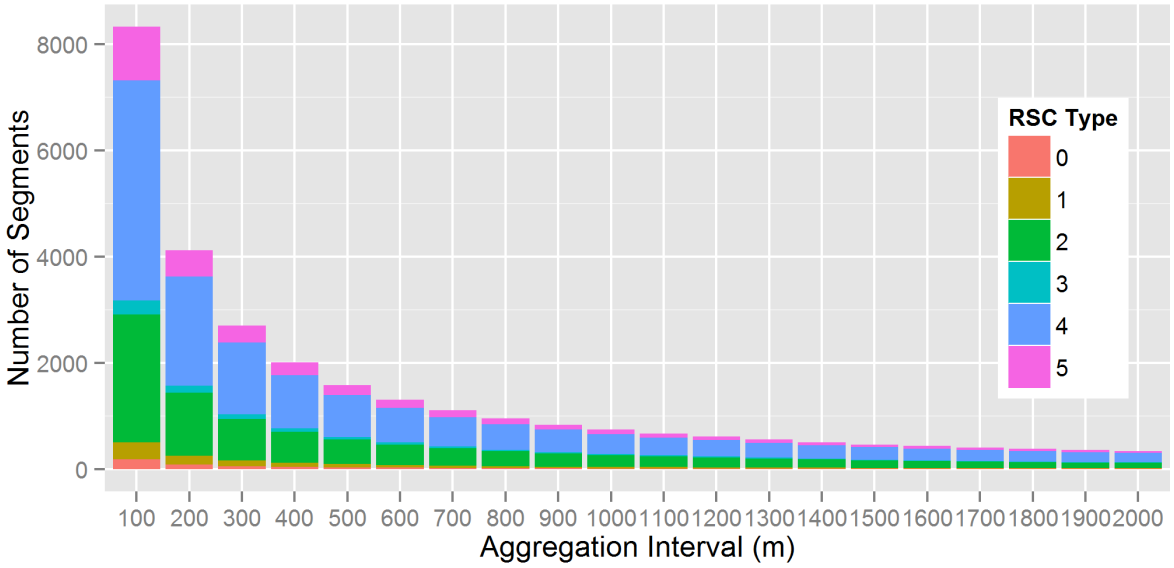


Figure 3.6: Sample Size and Aggregation Interval Length

Type 3 have the smallest sample sizes among the 6 RSC types at all aggregation intervals, and both their sample sizes drop under 20 if the aggregation interval is larger than 1000m. In our following modelling work, the sample will be divided into two sets, one of which is for model calibration and the other is for model validation. Thus the sample size less than 20 could be the lowest tolerable number of samples for our study, i.e., 1000 meters might be the longest aggregation distance limited by our collected CFMs.

Although it seems that we can select any aggregation interval length between 0 and 1000 meters, we still need to find a meaningful way to optimize the resulting models' classification performance. In our preliminary modelling works, logit models using the aggregation interval of 100m, 500m and 1000m have been calibrated and validated, and the models' performance was found acceptable (Fu et al., 2008). With a systematic approach detailed in Section 3.6.3, 700m is choosen as the major spatial aggregation interval for the later modelling work in this chapter.

Table 3.2: Sample Size and Aggregation Interval Length

Interval (m)	Type0	Type1	Type2	Type3	Type4	Type5
100	183	325	2400	264	4147	1003
200	91	161	1189	130	2051	494
300	61	105	784	85	1351	322
400	45	77	581	62	1008	239
500	36	62	460	48	794	182
600	30	50	384	40	650	154
700	25	45	325	33	555	129
800	22	35	279	29	479	114
900	19	31	248	26	420	95
1000	17	29	223	22	373	83
1100	15	27	201	20	333	76
1200	15	23	183	19	307	70
1300	13	22	169	14	282	59
1400	12	22	152	13	254	57
1500	11	17	135	13	239	52
1600	11	15	128	13	220	51
1700	10	15	119	11	206	45
1800	9	13	115	11	191	42
1900	9	13	105	11	184	37
2000	8	11	101	10	175	35

3.4.2 Data Aggregation

The CFM test segments of all test runs are cut into 700m (70 point-wise measurements) long intervals. For each interval, the sample mean, standard deviation and skewness of friction measurements are calculated using the formulae listed in Section 3.3.1. Measurements of each interval are also treated as a time series and first differenced. Then the mean spectra for low (0.00∼0.25 periods/point) and high frequency (0.25∼0.50 periods/point) are calculated using the statistical analysis software R (R Core Team, 2012) based on Equation 3.8 and 3.9. In this way, each aggregated interval of CFMs becomes a single sample, and the RSC type of each such sample is designated by visual observation as mentioned before. The basic features of one such single sample are

1. mean friction
2. standard deviation
3. skewness
4. low-frequency spectra
5. high-frequency spectra

Using exactly the same integration method, interval lengths of 300m and 1000m are also applied to the CFM data. Although these two aggregation intervals are not considered in the subsequent modelling work, they will be studied together with 700m aggregated samples in the exploratory analysis to give a more complete picture about the existent patterns of the extracted features. The sample sizes after aggregation by 300m, 700m and 1000m have already been listed in Table 3.2.

3.5 Exploratory Analysis

In this section, the aggregate features are examined by RSC types in order to find the patterns which are potentially helpful in RSC classification and model interpretations. Each aggregate feature is separately explored, and the interaction effect of two spectra features–low-frequency and high-frequency spectra, is also examined.

3.5.1 Mean Friction

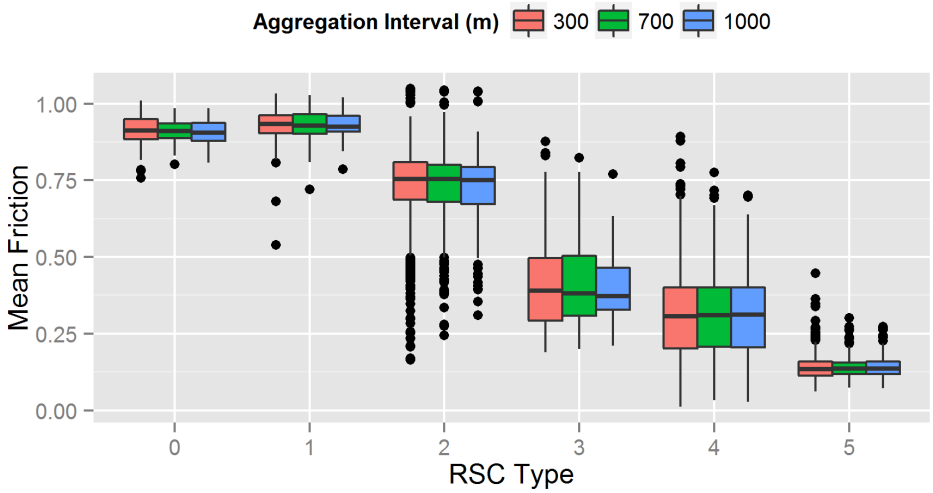


Figure 3.7: Mean Friction

Figure 3.7 shows the distribution patterns of the mean friction of the aggregate samples. For all three aggregation intervals, the average mean frictions decrease when the snow/ ice coverages increase. The ranges of mean friction of Type 0 and 1 overlap to a large extent, and both of them are within the upper 25% of the range of Type 2. Type 3 and 4 overlap a lot, and the full range of Type 5 is bracketed by Type 4. As mean friction is the most direct and accurate estimator of RSCs suggested by most previous studies, the overlapping structure of the mean friction levels of multiple RSC types could be considered as a representation of the degrees of pair-wise similarity between these types, and thus treated as the reliable basis to design any nested model structure regardless of model type.

3.5.2 Standard Deviation

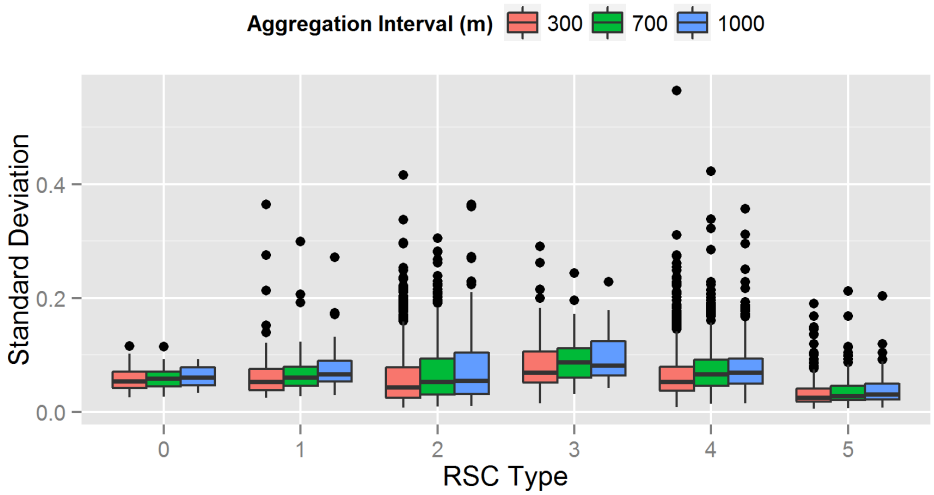


Figure 3.8: Standard Deviation of Friction

Figure 3.8 shows the distribution patterns of the standard deviation of the aggregate samples. For all three aggregation intervals, the standard deviations of Type 5 are the smallest, and of Type 3 and 4 are the largest. The average standard deviations of Type 0, 1 and 2 are very close, but the range of Type 2 encloses the full ranges of Type 0 and 1. The patterns suggest that standard deviation can be effective in distinguishing certain RSC types, e.g. Type 3 and 5, from the others.

3.5.3 Skewness

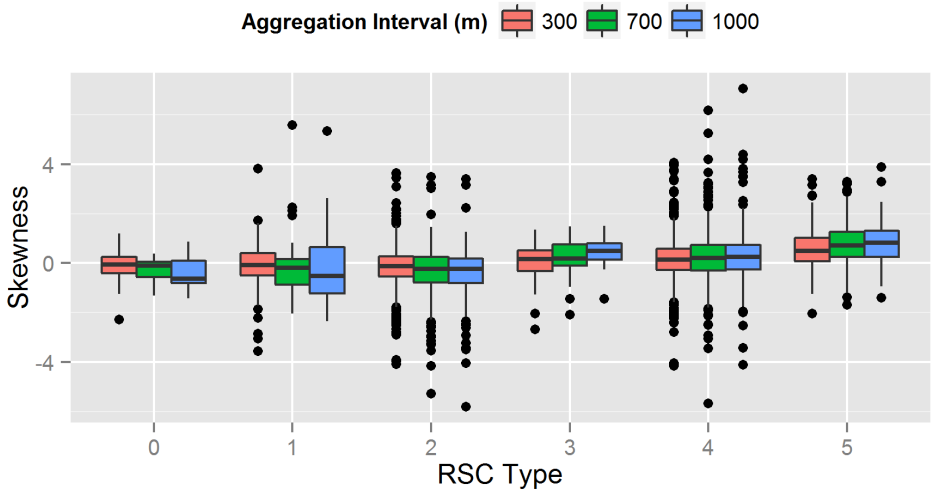


Figure 3.9: Skewness of Friction

Figure 3.9 shows the distribution patterns of the skewness of the aggregate samples. For all three aggregation intervals, the average skewness of Type 0, 1, and 2 are negative, and the other three types are positive. Negative skewness corresponds to a heavy left-tailed distribution pattern, i.e., a lot of large friction values with a small proportion of small friction values. This pattern means the snow coverage is small and the surface is mostly bare. Vice versa, positive skewness implies that the snow coverage is relatively large, which corresponds to Type 3, 4 and 5. This observation implies that the sign of skewness could be an efficient discriminator of the degree of snow/coverage.

3.5.4 Low-Frequency Spectra

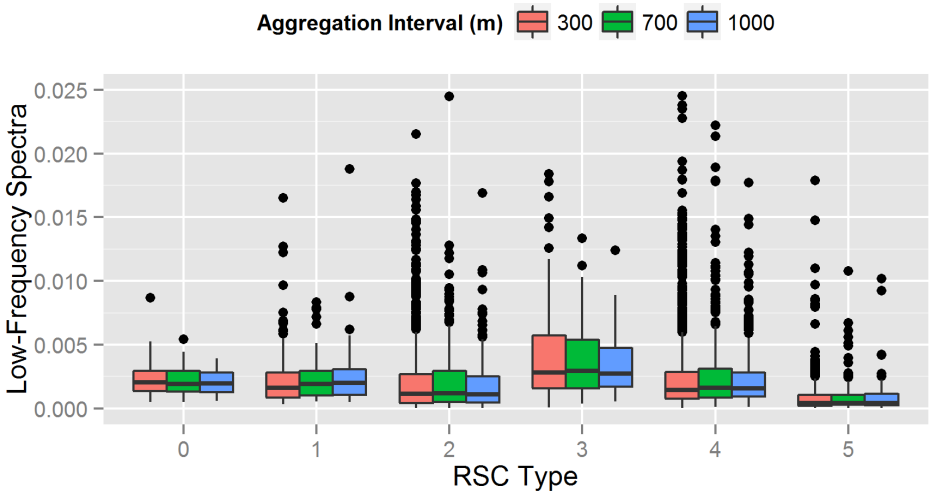


Figure 3.10: Low-Frequency Spectra of Friction

Figure 3.10 shows the distribution patterns of low-frequency spectra. For all three aggregation intervals, Type 3 is the highest and Type 5 is the lowest, with other types in the middle range. It appears that low-frequency spectra can be used to reliably identify Type 3 and 5. The other four types have heavily overlapped ranges and very similar means, suggesting that skewness is most likely not an effective measure for discriminating them from each other.

3.5.5 High-Frequency Spectra

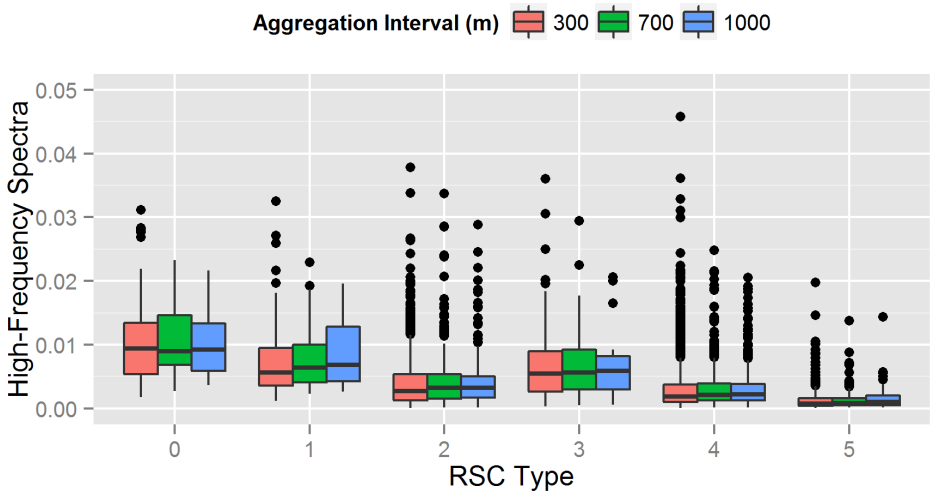


Figure 3.11: High-Frequency Spectra of Friction

Figure 3.11 shows the distribution patterns of high-frequency spectra. For all three aggregation intervals, the high-frequency spectra generally decrease with increasing snow coverage. Type 0 and 1 have the highest high-frequency spectra, while Type 5 the lowest. Type 3 is on average higher than Type 2, which contradicts our presumption made in Section 3.3.2 about the correlation pattern between high-frequency spectra and snow coverage. In general, the patterns in Figure 3.11 suggest that high-frequency spectra could be another powerful estimator of snow coverage.

3.5.6 Proportion of High-Frequency Spectra

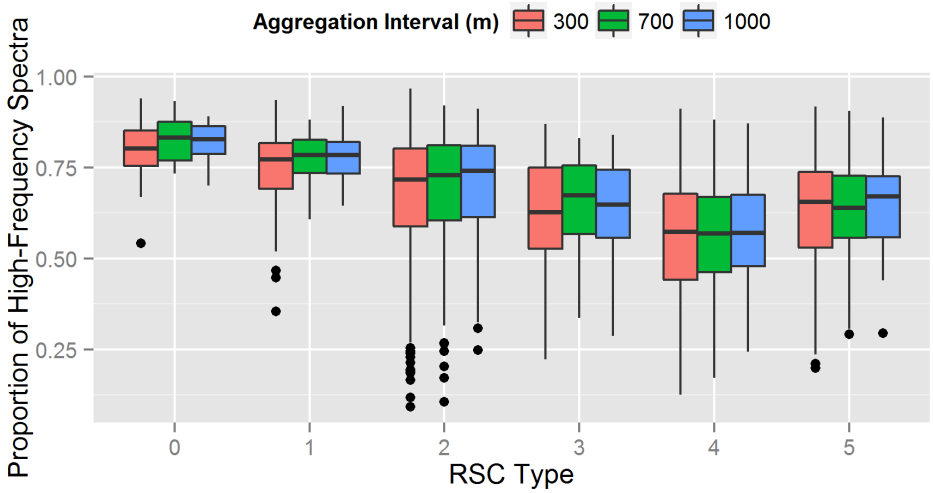


Figure 3.12: Proportion of High-Frequency Spectra of Friction

In addition to the main effects of individual features, the interaction effects of some features could be used to distinguish some important patterns in high-dimensional feature space for enhanced model performance. Adding an interaction term, which is usually obtained by multiplying relevant explanatory variables, to the model is a common technique to examine the corresponding interaction effect. This interaction term can be obtained in a manner other than multiplication, e.g. by division. In this study, we are interested in studying the interaction effect between the spectra of two frequency ranges; therefore, the interaction term between them is obtained by dividing the high-frequency spectra by the sum of high-frequency and low-frequency spectra. The reason for including the interaction effect in this way is that the division operation has an easier interpretation. Equation 3.10 indicates that the total variance of the series is the integrated spectral density over all of the frequencies. The interaction term obtained in the above way is the proportion of high-frequency spectra in the total spectra, i.e., the proportion of the total variance caused by high-frequency sub processes.

Figure 3.12 shows the distribution patterns of high-frequency spectra proportion. For all three aggregation intervals, the high-frequency spectra proportion generally decreases with increasing snow coverage, which is similar to high-frequency spectra distribution. Type 4 is the lowest instead of Type 5 being the lowest in Figure 3.11. This could be caused by

the smaller variance of CFMs on Type 5 surfaces. Such an interaction term can provide some additional information and help in identifying certain RSC types, for example Type 4 in this case.

In a similar way, all relevant and meaningful interaction effects could be explored and tested. In this study, we did not proceed with this process for the following reasons.

- Main effects are usually much more important than interaction effects. If there are some limitations in sample size or data quality, main effects should be studied with priority.
- Interaction effects could sometimes cause unexpected difficulties in model interpretation.
- Interaction effects may seriously blur the main effects of some important explanatory variables, an undesirable result for empirical studies.

In fact, another consideration in deciding the number of features and interactions to include in a model is the "curse of dimensionality", which is a term referring to the properties of high-dimensional feature spaces often appearing counterintuitive because human experience with the physical world is in a low-dimensional space, such as a space with two or three dimensions. Kantardzic (2002) has clearly explained this phenomenon conceptually, numerically and graphically. To avoid the "curse of dimensionality", the number of the tested explanatory variables should be as small as possible, preferably fewer than four. Although the numbers of tested explanatory variables in this chapter and the next chapter both exceed four, we should always be aware of this issue and focus on the most important variables. Therefore, in the model calibration process, only the parameters discussed in this explanatory analysis are considered. Other features and interaction terms will not enter the resulting models, notwithstanding some of them might be good estimators of RSC.

3.6 Methodology

Using the features explored in the last section as explanatory variables to classify the RSCs into the most possible RSC type or types is a typical supervised classification problem and can be addressed by a number of traditional modeling approaches. In order to model the uncertainties inherent in friction measurements, the modelling method must be capable of estimating RSCs in a probabilistic way, i.e., the probabilities associated with individual types. Several statistical methods can be used to deal with this stochastic estimation problem, and they generally belong to two schools: white-box classifiers and black-box classifiers.

A white-box classifier is a classification model whose internal structure and the effects of explanatory variables are transparent to the users and can be easily interpreted and corroborated with the modelled physical process. A white-box classifier usually consists of one or a set of structured numerical models each of which is a linear combination of relevant explanatory variables. Linear discriminant functions (FISHER, 1936; Gnanadesikan, 1997), logistic regression models (McCullagh & Nelder, 1989), and classification trees (Breiman, 1984) are three common white-box classifiers.

A black-box classifier is usually a data structure calibrated and adapted to the dataset according to some highly complex optimization criteria. This data structure could be a set of critical data points outlining the class boundaries, a function dynamically calculating the class similarity of neighboring sample points, or multiple layers of weighing, combination and transformation of explanatory variables. The data structure is usually high-dimensional and very complex thus difficult to interpret. The most popular black-box classifiers are nearest neighbors (Cover & Hart, 1967; Ripley, 1996), neural networks (Ripley, 1996; Venables & Ripley, 1999) and support vector machines (SVMs) (Vapnik, 1999; Kecman, 2001).

3.6.1 Logistic Regression

As the major objective of this part of the thesis study is to understand how the aggregate measures of CFMs can augment the capability of the resulting RSC estimation models, white-box classifiers usually have a decided advantage over black-box classifiers as they can clearly reveal the effect of each individual feature and allow both the modeller and model users to comprehend the possible causal relationships between features and the estimated variable. Table 3.3 compares the main characteristics of the mentioned white-box classifiers.

Table 3.3: White-Box Classifier Comparison

Classifier	Complexity	Interpretation	Statistic Assumption
Linear Discriminant Analysis (LDA)	simple	easy	strong
Logistic Regression (LR)	simple	easy	mild
Classification Trees	complex	difficult	mild

Fisher's discriminant function (DF) is a single linear combination of the explanatory variables. The coefficients of the DF are chosen so that the DF values are differentiated between the two classes as much as possible in the training data. In particular, those coefficients of the DF are estimated by maximizing the ratio of between-class variation to within-class variation. The basic DF cannot give any probabilistic estimation, as the only important assumption for LDA is that the distribution of observations within each class has an identical variance-covariance matrix of explanatory variables. In order to estimate probabilities, certain distribution has to be assumed within each class, and Gaussian distribution is mostly used. The assumption of Gaussian distribution with identical within-class variance-covariance is usually too strong for high dimensional complex classification problems, although the discriminant functions generated by linear discriminant analysis are simple and easy to interpret.

Another widely used white-box classifier is the classification tree. The basic idea of the classification tree is to partition the space of explanatory variables into successively smaller hyper-rectangles in order to make the sample more and more pure in terms of the response variable's class within the newly created hyper-rectangles. This recursive structure can be represented as a tree and the successive hyper-rectangles are then represented by the nodes of the tree. Breiman (1984) has shown that multiple explanatory variables can be linearly combined at each split to model interaction effects, but the resulting tree tends to be more difficult to explain and more data-sensitive. The criteria for finding optimal splittings can be Shannon entropy (Shannon, 1951), Gini coefficient (Gini, 1971) or any customized numerical measurement of class dispersion. So classification trees are very flexible in incorporating different statistical assumptions, and thus can be applied to a wide variety of classification problems. Additionally, although its structure tends to be more complex and difficult to interpret when interaction terms are considered, the whole composition of a classification tree is completely transparent.

Logistic regression (LR) is a special form of generalized linear models and is an extension of the regression-motivation for Fisher's LDA. As LR is chosen as the major modelling method for the RSC type estimation, a detailed discussion is given as follows.

The general form of an LR model is

$$\ln \frac{p(Y = C_k)}{1 - p(Y = C_k)} = \eta(x) \quad \forall C_k \in C \tag{3.12}$$

Y: the categorical response variable Y
C: the set of classes, which includes all possible discrete classes of Y
C_k: Class k in C
$p(Y = C_k)$: the probability of Y being in Class k
X: the explanatory variable vector of d features, i.e., $X = (x_1, \ldots, x_d)^T$

$\eta(x)$ is a function describing the dependence of the odd ratio $\frac{p(Y=C_k)}{1-p(Y=C_k)}$ on X. $\eta(x)$ is commonly defined as a linear function as

$$\eta(x) = \beta_0 + \beta_0 x_1 + \ldots + \beta_d x_d \tag{3.13}$$

where β_0,β_1,...,β_d are model coefficients to be estimated. Therefore the probability of Y belonging to any specific class can be estimated by the following simple transformation of Equation 3.13:

$$p(Y = C_k) = \frac{e^{\eta(X)}}{1 + e^{\eta(X)}} \tag{3.14}$$

Equation 3.14 shows that the estimated probabilities can never go beyond the range of $[0, 1]$, which is a major appeal of logistic regression versus the ordinary linear regression model. The coefficients for $\eta(x)$ in Equation 3.13 are usually estimated using the maximum likelihood method, which is implemented in most statistical software packages.

In order to estimate coefficients for the ordinary binary case, (i.e., there are only two possible classes), Y is assumed as a Bernoulli random variable, which is a special case of a binomial random variable with one trial. Where there are more than two classes, based on their assumption and corresponding strategies in handling inter-class structures, LR models could be classified into three types in terms of models structures, namely, multinomial logit (MNL) models, ordered MNL models and multi-Level LR models.

Ordinary MNL Models

If all observed Ys in the training dataset are statistically independent, and all classes (more than two) are unordered and non-overlapping, the categorical variable Y is multinomially

distributed, such that the binary LR model in Equation 3.14 is extended to a multinomial logistic (MNL) model as follows:

$$p(Y = C_k) = \frac{e^{\eta_k(X)}}{\sum_{i=0}^{c-1} e^{\eta_i(X)}} \quad k = 0, \ldots, c-1 \tag{3.15}$$

With c classes, $c-1$ linear predictor functions $\eta_k(x)$ are calibrated for $k = 1, \ldots c-1$ and $\eta_0(x)$ are arbitrarily set to 1 in order to estimate the unknown parameters in $\eta_k(x)$ for $k = 0, \ldots, c-1$ using the maximum likelihood method.

Ordered MNL Models

If C_k have some ordered relationship in terms of certain feature, which appears in $\eta_k(X)$, this relationship can be utilized by calibrating ordered MNL models to enhance the analytical power of the resulting model set. For instance, the studied set of RSC types in our study can be ordered by their slipperiness level as Type 0,1,2,3,4,5 as suggested by Figure 3.7. If the mean friction level or other correlated explanatory variables is entering the model calibration process, a set of MNL models can be calibrated and interpreted as follows.

Let $p(C_k)$ denote the probability of belonging to Type k for $k = 0, \ldots, 5$ given the explanatory variable vector X. β_j are held the same across $\eta_k(X)$ for $j = 1, , d$, where d is the number of explanatory variables. The intercepts of all $\eta_k(X)$, (i.e., all β_0), are allowed to vary. In this way, five binary LR models are calibrated sequentially for $p(0)$, $p(01)$, $p(012)$, $p(0123)$ and $p(01234)$. The types can accumulate in the direction consistent with their assumed order; thus the probability of a single type can be recovered in the accumulation direction.

The set of MNL models organized in this way and calibrated assuming equal $\beta_j's$ and varying $\beta_0's$ across models are called proportional-odds models as the odds ratio of $p(Y \leq C_k)$ for two different observations X_1 and X_2 is

$$e^{-\beta^T(X_1 - X_2)} \tag{3.16}$$

where β is the vector of β_j for $j = 1, \ldots, d$.

This odds ratio of X_1 and X_2 is independent of class id so that no matter how the classes are directionally divided and the model of which class is tested against, the effect of X can be analyzed in a more consistent manner. Therefore, after the binary models

are calibrated, neighboring classes can be flexibly amalgamated and the models are not required to be recalibrated. This property makes ordered MNL models more appealing than ordinary MNL models in terms of easiness of interpretation and application.

Multi-Level LR Models

The structure of the ordinary MNL model is based on the assumption that the classes are completely independent and mutually exclusive. This assumption may be violated in some applications where some of the classes are more similar to each other and thus share more common characteristics than others. As a result, a more proper classification scheme is a multi-level structure. Take the RSC type classification scheme of this study for example; the mean friction level ranges of RSC types may overlap, and some pairs of RSC types may overlap more than others, i.e., the similarity degrees between discrete classes may be different. Figure 3.7 shows the box-plot of mean friction of all six types. Ranges of the mean friction of Type 0 and 1 overlap seriously, while both of them are within the range of Type 2. Type 3 and 4 overlap a lot, while the full range of Type 5 is bracketed by Type 4. As the mean friction value is the most significant estimator of RSC as shown in most previous studies, the overlapping structure of the mean friction distribution could be treated as the reliable basis to design a multi-level structure representing the degree of similarity between RSC types.

In our preliminary work, a classfication tree was calibrated to provide quantitative evidence to support this intuitive finding. The mean friction, skewness, standard deviation, low-frequency spectra and high-frequency spectra were extracted from 500m aggregated sample and used as the exploratory variables of this classification tree. The response variable was the RSC type valued 0 to 5. The calibrated classification tree is shown in Figure 3.13.

Mean friction (F) and low-frequency spectra (LowFreq) appeared in the tree to classify samples at different levels of the tree structure. This tree could not efficiently distinguish between Type 0 and 1, and between Type 3 and 4, which suggests that Type 0 and 1 are very similar and so are Type 3 and 4. Based on this observation, a multi-level modelling structure is designed in the way shown in Figure 3.14.

At each level of this structure, a single binary LR model is calibrated. A new observation can be classified by sequentially applying the LR model at each level from top to bottom. In total, five binary LR models are required to assign the RSC into one of six types. Split 1 model resides at the top level of the structure and estimates $p(0, 1, 2)$ and $p(3, 4, 5)$. Split 2 model estimates $p(01|012)$ and $p(2|012)$. Similarly, Split 3 model estimates $p(34|345)$ and

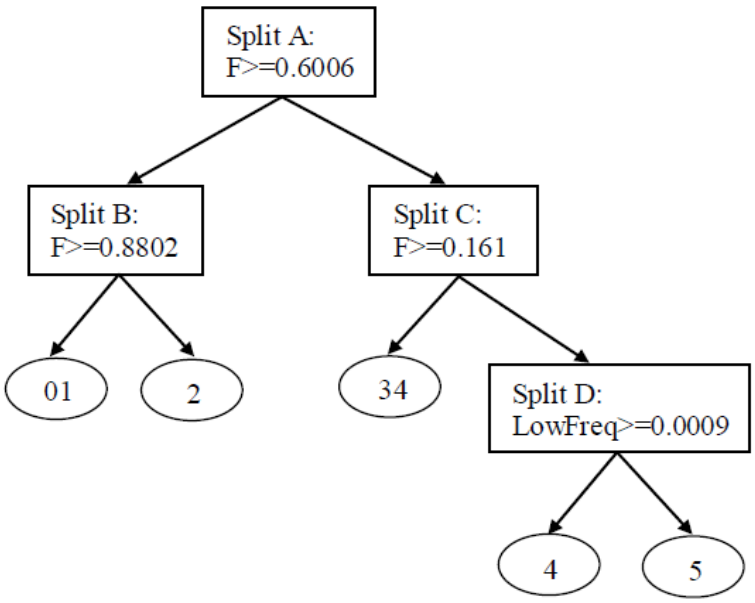


Figure 3.13: RSC Type Classification Tree

$p(5|345)$. Split 4 model and Split 5 model are at the lowest level of the tree for capturing binary probability divisions between Type 0 and 1, and between Type 3 and 4.

The unconditional probability for each individual type can be derived by sequentially multiplying the corresponding estimated probabilities in different levels of the tree. Therefore, the probability estimations for the six studied types are calculated as

$$p(0) = p(0,1,2)p(0,1|0,1,2)p(0|0,1)$$
$$p(1) = p(0,1,2)p(0,1|0,1,2)p(1|0,1)$$
$$p(2) = p(0,1,2)p(2|0,1,2)$$
$$p(3) = p(3,4,5)p(3,4|3,4,5)p(3|3,4)$$
$$p(4) = p(3,4,5)p(3,4|3,4,5)p(3|3,4)$$
$$p(5) = p(3,4,5)p(5|3,4,5)$$

The multi-level LR models, which are usually called nested LR models or structured LR models, can deal with classes with various similarity levels. In certain circumstances, this nested model structure is a requisite to circumvent possible fallacies in classfier design so that the estimated probability distribution of classes is more reasonable and reliable.

For example, the RSC estimate using a binary CFM-based LR model is {BareWet: 0.5, PartiallySnowCovered: 0.5}. If the bare dry type is added to the possible RSC type set and a CFM-based MNL model is applied, according to the Independence of Irrelevant

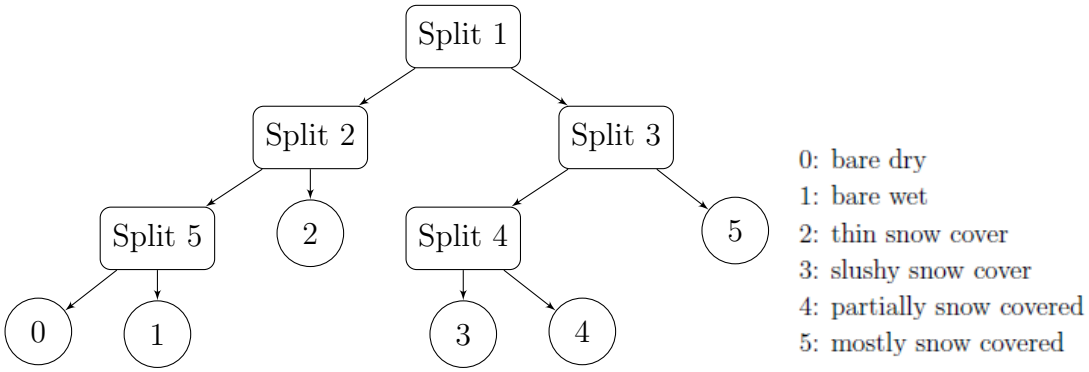


Figure 3.14: RSC Estimation Model Structure

Alternatives (IIA) property of the logit model, the probabilities of bare wet and partially snow covered types are still equal. We assume CFMs of bare dry and bare wet types are very similar, which is supported by Figure 3.7, the resulting probability distribution would be close to {BareDry: 0.33, BareWet: 0.33, PartiallySnowCovered: 0.33}. The probability of the partially snow covered type decreases significantly from 0.5 to 0.33. If we keep adding more types for consideration whose CFMs are similar to bare wet, we would expect the estimated probability of the partially snow covered type will keep decreasing; however, the true RSC could be 50% partially snow covered, and the other 50% possibility is shared by all other types, i.e., the probability of being partially snow covered should not be significantly affected by just considering more possible RSC types. In such situations, MNL models can not successfully capture this distribution pattern and nested structures must be implemented.

Besides this forte at handling complex class structures, multi-level LR models can easily incorporate interaction effects by adding interaction terms to the concerned single LR models. Moreover, the statistical assumption behind LR models is a binary distribution, which is very mild thus friendly to data anomalies. Additionally, the LR models are in the linear form and all model coefficients straightforwardly reflect the effect of corresponding features. Finally, LR models directly produce probabilistic estimations which fit our study objective very well. Due to these merits, we use logistic regression as the major modelling approach for RSC estimation and the modelling process will follow the multi-level structure specified in Figure 3.14. Other alternatives of the structure can also be designed depending on how the class similarity is defined.

3.6.2 Support Vector Machines

Although black-box classifiers are usaully difficult to interpret, they can capture some highly complex interaction effects which even the most advanced white-box classifiers cannot. Usually such interaction effects are only visible within the high-demensional feature space, and since the class boundaries in this space are irregular, they cannot be represented as interpretable numerical functions. Therefore, more often than not, the performance of a black-box classifier is better than a white-box classifier using the same set of classification features. So even black-box classifiers are not considered as the major modelling tool for our RSC estimation problem, such classifiers can be calibrated in parallel with the LR models in order to examine the maximal classification accuracy that can be obtained by using the selected set of explanatory variables. Although such classifiers are almost impossible to interpret, they set some visible upper limits for the model performance so that we can assess the calibrated LR models with more confidence. For this purpose, we choose support vector machines as such auxiliary estimation tools; i.e., for each split, a support vector machine is trained with the same set of variables and data as the corresponding LR model.

The support vector machine (SVM) is a modern classification method, which can find the class boundaries in a transformed feature space in a very efficient way. Its algorithm depends on so called "kernel functions" to find highly irregular boundaries in high-dimension feature space. Due to its special mechanism, a small sample size does not pose a serious problem to the successful calibration of an SVM. Although the calibration is fast for SVM, the final form of the model is fairly complex and the degree of complexity depends on the number of support vectors. Support vectors are the samples close to the class boundaries in the feature space and the number of support vectors depends on the geometrical complexity of the corresponding boundaries. The mathematical properties and training method of SVMs are detailed by Vapnik (1999) and Kecman (2001).

3.6.3 Selection of Aggregation Interval

One of the most important goals of interval aggregation is to find reliable estimations of aggregate features within the interval, thus using them to calibrate models with the highest classification accuracy about the interval's RSC type. Although there is more than one way to achieve this goal, for our study, we solidify this idea with the following steps:

1. The continuous segments, over which CFMs were collected are divided by different interval lengths from 100 meters to 2000 by 100-meter increments, i.e., the intervals of 100, 200, 300, ..., 2000 meters are respectively tested.
2. The aggregate features are calculated for each divided interval of CFMs.
3. The aggregate features are used to calibrate classification models for each aggregation interval length.
4. Choose the aggregation interval, which results in a classification model with the highest classification accuracy. The classification accuracy is measured using "hit rate", which is the proportion of correct estimates of all the estimates made by the model.

The aggregate features include standard deviation, skewness, mean high-frequency (0.25∼0.5 periods/point) spectra, and mean low-frequency (0.00∼0.25 periods/point) spectra. An additional term, mean high-frequency spectra divided by the total spectra, is added as the interaction term of high-frequency spectra and low-frequency spectra. Physically speaking, it is just the proportion of high-frequency spectra in the total variance of the spatial series. Additionally, the mean friction does not enter the model at this stage for the following reasons:

- As previous studies suggested, mean friction is a strong discriminator of RSC types. If it enters the model, sometimes its effect may overwhelmingly mask the discriminating power of other features.
- A main purpose of this analysis on the aggregation interval length selection is to choose an appropriate interval so that the aggregate features can substantially show their discriminating powers.

The RSC types are first divided into two classes–Type (0,1,2) and Type (3,4,5), and a binary logit model and a support vector machine (SVM) with the Gaussian radial basis

function as the kernel function are respectively calibrated with all available aggregated samples for each aggregation length. In other words, 20 logit models and 20 SVMs are calibrated for 20 aggregation interval lengths. Logit models and SVMs will be discussed in detail in the methodology part of this chapter. These models are used to estimate RSC types in terms of the two classes. The hit rates of all these models are shown in Figure 3.15. Another classification scheme is also tested, i.e., RSC types are divided into two classes: Type (0,1,2,3) and Type (4,5), and the hit rates of the new 40 models are shown in Figure 3.16.

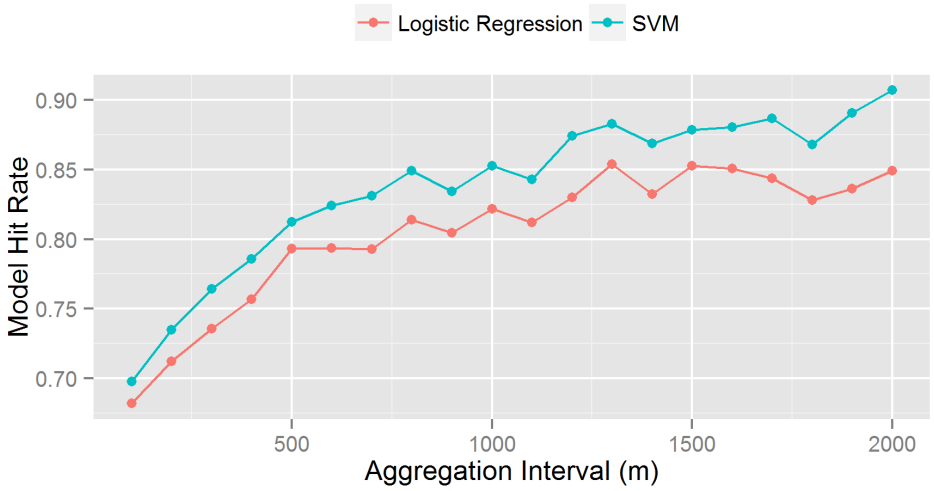


Figure 3.15: Model Hit Rates for Different Aggregation Interval Lengths
Type (0,1,2) vs. Type (3,4,5)

For both classification schemes, the hit rates of both logit models and SVMs increase with the aggregation interval length. When the aggregation interval length is shorter than 500 meters, the hit rates increase faster, i.e., increasing the aggregation interval length can significantly improve model hit rates. When the aggregation interval length is longer than 500 meters, this improving effect still exists but is not that significant, as suggested by the curves becoming more and more flat in both figures. The most significant changes of the slopes of all these four curves coincidentally occurs within the area where the aggregation interval length is between 500 and 1000 meters. Another observation from the figures is that the hit rates of SVMs are higher than logit models at all aggregate interval lengths with the biggest difference being less than 10% and most less than 5%. The following are some interpretations of these observed patterns.

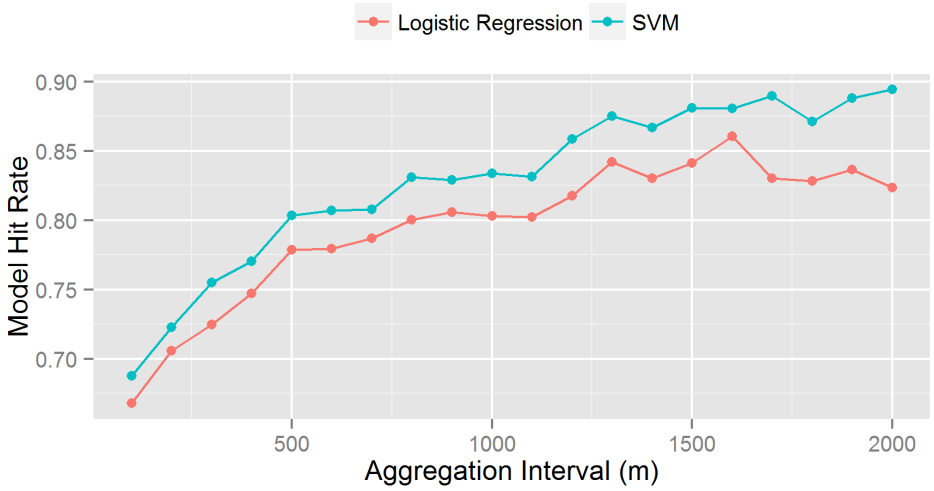


Figure 3.16: Model Hit Rates for Different Aggregation Interval Lengths Type (0,1,2,3) vs. Type (4,5)

- With the aggregation interval becoming longer, more point-wise CFMs are used to calculate aggregate features, which makes those feature estimations more reliable, thus reflecting the RSC characteristics of the interval in a more accurate and stable way. Consequently, the performance of the resulting models using these features keeps improving with the increase of aggregation interval. Therefore, there is an upward trend for all four curves.

- For both classification schemes and both classifiers, the improvement of the model performance becomes more and more marginal after the aggregation interval reaches 500 meters. This model performance transition zone exists in the aggregation interval length between 500 to 1000 meters.

- SVMs are better classifiers than logit models for all tested aggregation interval lengths, suggesting the RSC classification features have some high dimensional interaction effects, which are difficult to capture using traditional classifiers; however, the overall performance of SVMs is not significantly better than logit models as their hit rates are not substantially different, suggesting that those high dimensional interaction effects are not playing an important role.

From the perspective of model performance, a longer aggregation interval length is preferable; however, as mentioned previously, a longer aggregation interval means a fewer aggregated sample size; and our available dataset does not allow the interval length to be larger than 1000 meters. Furthermore, from the operational perspective of maintenance personnel and decision makers, a shorter aggregation interval is a better choice. Therefore, in our study, 700 meters is chosen as the primary aggregation interval length for extracting aggregate features and calibrating models. Again, this interval length is determined based on a carefully selected trade-off point of sample availability, model performance and the easiness of model application.

Another implication by Figure 3.15 and 3.16 is that although traditional linear classification models, like logit models, are not as efficient as modern classifiers, such as SVMs, in capturing high dimensional feature patterns, their performance is comparably acceptable, at least for the purpose of this study. Moreover, one major objective of this study is to give explicit interpretations for the resulting models so that they can be applied in the real world with more confidence.

3.7 Model Calibration and Validation

For each split in Figure 3.14, a binary LR model and an SVM with the Gaussian radial basis kernel function are separately calibrated using the following explanatory variables:

- Mean Friction ($mean\_fric$)
- Standard Deviation (sd_fric)
- Skewness($skew\_fric$)
- Low-Frequency Spectra ($lowspec\_fric$)
- High-Frequency Spectra ($highspec\_fric$)
- Proportion of High-Frequency Spectra ($highspec\_prop\_fric$)

The outliers are filtered out at first. Outliers are the observations with the value of any major feature larger than the respective 97.5% quantile. The values of all variables are linearly scaled to the range [0,1], with 0 as the scaled minimum and 1 as the scaled maximum. The purpose of this scaling is to show the effects of explanatory variables in a comparable way within the same model. Unless specified otherwise, all model calibration and validation will be conducted using the scaled data. Scaling variables in this way for linear regression models leaves the t and F tests and R^2 unchanged and coefficients rescaled by a constant (Faraway, 2002).

As in this study, only conditional probability is modelled at each split, so only samples of relevant RSC types are used for model calibration and validation. For example, at Split 1, samples of all RSC types are involved, so the samples of Type 0, 1, and 2 are treated as Class 0, and the other three types are treated as Class 1. At Split 5, the relevant RSC types are limited to 0 and 1; thus the samples of Type 0 are treated as Class 0, and Type 1 as Class 1. Therefore, the closer the split is to the root of the structure, the more RSC types are involved.

Another issue in model calibration and validation is that some samples should be held out for cross-validation so that only a part of total relevant samples are used for model calibration. The cross-validation results obtained by applying the calibrated model to the hold-out data can represent the performance of the model when it is applied to new data. As the available sample sizes of some RSC types (e.g. Type 0, 1 and 3) are very small, when specific splits involve any of these types, the allocation of calibration and validation data should comply with the following rules of thumb:

- No matter how many RSC types are relevant to the split, the samples are first assigned the class id of 0 or 1 according to the splitting direction. After that, "balanced numbers" of samples in Class 0 and 1 are randomly selected for model calibration from all relevant samples. The term "balanced" means the numbers should not differ too much, as extremely unbalanced numbers of samples in two classes could seriously bias the calibration results. If the sample size is large enough, 50%-50% is the most preferable proportion.

- When the sample size is very small, the priority is to leave enough data for model calibration and at the same time to avoid extreme unbalance. In other words, we have to meet the minimum sample size requirements for model calibration first (usually 5 each class) and try to get balanced numbers of sample for each class. Then the rest of the samples can be used for model validation.

At each split, the calibration and validation datasets are carefully chosen according to the above rules. The classification accuracy of the calibrated LR model and SVM at each split is evaluated using hit rates at that split, which include the hit rates of both calibration and validation datasets. For each sample, the LR model estimates the probabilities of belonging to both classes; therefore, the class with an estimated probability greater than 0.5 is deemed as the sample's estimated class, which is used to calculate the hit rate of the LR model.

All the calibrated LR models are calibrated with the "glm" function in R and the R outputs are listed in the Appendix of this thesis. The SVMs are trained with the "svm" function in the R package "e1071" with the Gaussian radial basis kernel function; but R outputs are not explicitly listed because the model structures are too complex and we do not attempt to give any interpretations on them. Again, the only thing concerning SVMs is their hit rates compared to those of LR models.

The calibrated LR models based on CFM data for all five splits are summarized in Table 3.4. The coefficients of significant and marginally significant explanatory variables are listed in the table. The numbers in parentheses are the standard deviations of respective estimated coefficients. Several major diagnostic statistics for model fitness are also shown in the table including AIC, BIC, log likelihood and deviance.

Table 3.4: CFM-Based LR Models for RSC Estimation

	Split 1	Split 2	Split 3	Split 4	Split 5
$(Intercept)$	25.73^{***}	90.06^{*}	1.48	1.36^{**}	$-43.00^{\cdot}$
	(4.13)	(42.84)	(0.98)	(0.52)	(26.12)
$mean\_fric$	-21.94^{***}	-97.49^{*}	-19.71^{***}		$41.20^{\cdot}$
	(2.70)	(47.26)	(2.98)		(21.26)
$highspec\_prop\_fric$	-9.33^{**}		3.25^{*}		$-20.76^{\cdot}$
	(3.22)		(1.41)		(11.01)
$skew\_fric$	-12.50^{***}				$35.53^{\cdot}$
	(3.75)				(19.03)
sd_fric	-7.21^{***}				36.02^{*}
	(2.06)				(17.12)
$highspec\_fric$	10.29^{**}	-13.66^{*}		-9.82^{***}	
	(3.27)	(6.02)		(2.91)	
$lowspec\_fric$	-6.59^{*}			6.66^{*}	$-24.48^{\cdot}$
	(2.79)			(2.95)	(12.59)
AIC	128.23	21.61	153.12	68.95	36.60
BIC	157.40	29.40	163.08	75.52	44.59
Log Likelihood	-57.12	-7.81	-73.56	-31.47	-12.30
Deviance	114.23	15.61	147.12	62.95	24.60
Num. obs.	477	99	204	66	28

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

3.7.1 Split 1: Type 0, 1, 2 vs. Type 3, 4, 5

To calibrate Split 1 LR model needs the samples from all RSC types, which are divided into two classes: Class 1 {0,1,2} and Class 2 {3,4,5}. The calibration process in R is step-wise with AIC as the criterion of parameter entrance/exit.

AIC is the short form of the Akaike Information Criterion, which is a measure of the relative fitness of a statistical model (Akaike, 1974). The AIC of a model describes the tradeoff between accuracy and complexity of the model; thus it provides a means for model selection. AIC can be generally calculated as

$$AIC = 2k - 2\ln(L) \tag{3.17}$$

where k is the number of parameters in the model and L is the maximum likelihood for the model with those variables. According to Equation 3.17, a smaller AIC indicates a better balance between model complexity and model fitness. An important goal of using AIC to control the model complexity is to avoid overfitting, i.e., to make the model fit the assumed physical process instead of the dataset used for model calibration.

There are several other major model fitness measures based on similar ideas as AIC, among which BIC is widely used (Schwarz, 1978). BIC reprents Bayesian Information Criterion and is calculated as

$$BIC = -2\ln(L) + k\ln(n) \tag{3.18}$$

where n is the sample size. The BIC penalizes the number of parameters less strongly than does the AIC.

Using different fitness measures in the process of step-wise model calibration may generate very different models, although the performance of these models could be very close. Therefore, a model stepwisely calibrated using such fitness measures is not the only valid one. This is a very important perception for statistical modelling, which underscores a modeler's involvement in parameter selection. For most modelling works in this section, we will adjust the model structure in all the step-wise calibration processes with the AIC, and the resulting models are reviewed and intepreted in terms of parameters' physical meaning. If there is no major conflict in this process, the model is considered valid.

In the calibrated model for Split 1, besides the intercept, only $mean\_fric$ is left at the end of the step-wise calibration process. As indicated before, this does not mean that the effects of other variables are insignificant. It just suggests that with the specified fitness measure–AIC in this case, the model takes this form. The coeffecient of $mean\_fric$ is

negative suggesting that the larger the mean friction level is, the more likely the RSC type belongs to Class 0 (Type 0, 1, 2). The sign of this coefficient is consistent with the fact that higher friction measurements are associated with better RSCs.

The hit rates of the calibrated LR model and the correspoinding SVM are listed in Table **??**. The hit rates of the LR model are very high for both calibration and validation datasets, and the corresponding confusion matrices are reasonably symmetrical suggesting the estimation is reliable with little bias. Compared to the SVM, this LR model has very similar hit rates, which from another perspective confirms its good classification performance.

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	223	11	Est	0	80	17	Est	0	224	12	Est	0	78	21
	1	11	232		1	6	384		1	10	231		1	8	380
95%				95%				95%				94%			

In the LR model, all aggregate variables are significant, suggesting that all of them have contributed to the high hit rates. Taking a closer look at the coefficients, it can be found that

- Higher $mean\_fric$, $highspec\_prop\_fric$, $skew\_fric$ and sd_fric correspond to lower probability of being Class 1, which is consistent with the physical meanings presumed in previous discussion and/or the dicovered patterns in the exploratory analysis.
- Higher $highspec\_fric$ corresponds to higher probability of being Class 1, which conflicts with the explored pattern of $highspec\_fric$ in Section 3.5.5. As the calibration dataset is a mixture of the samples from all six RSC types with unequal proportions, the effect of $highspec\_fric$ exhibited in the model could be biased by this inequality of sample proportions and thus the sign of the coefficient is not as expected.
- Higher $lowspec\_fric$ corresponds to lower probability of being Class 1, which is difficult to interpret in terms of physical meaning. This is likely to be caused by the unbalanced sample sizes in certain RSC types.
- Compared to $mean\_fric$, the effect magnitudes of other variables are smaller but not negligible, which suggests that besides the mean friction, the other aggregate variables can significantly improve the RSC classification accuracy.

Although the signs of coefficients of some significant variables are questionable, the overall performance of this LR model for Split 1 is satisfactory. Its hit rates are almost the same as the SVM, suggesting that the model structure has efficiently exploited the classification power of the explanatory variables.

3.7.2 Split 2: Type 0, 1 vs. Type 2

The hit rates of the Split 2 model and the corresponding SVM are listed in Table **??**. Similar to the Split 1 model, the hit rates of the LR model are very high for both calibration and validation datasets, and the confusion matrices are reasonably symmetrical suggesting the estimation is reliable.

LR									SVM								
Calibration					Validation				Calibration					Validation			
		True				True				True				True			
		0	1			0	1			0	1			0	1		
Est	0	28	0		Est	0	8	9		Est	0	29	0		Est	0	9	8	
	1	2	69			1	2	202			1	1	69			1	1	203	
98%					95%				99%					96%			

In this model, only $mean\_fric$ and $highspec\_fric$ are significant, and their coefficients are both negative, i.e., higher $mean\_fric$ and $highspec\_fric$ correspond to higher probability of being Class 0. This fact is consistent with their presumed physical meaning, and also reflects the feature patterns as explored in Section 3.5. The coefficient magnitude of $mean\_fric$ is overwhelmingly larger than that of $highspec\_fric$ suggesting $mean\_fric$ is very important for this split model and the contribution by $highspec\_fric$ could be relatively marginal. This model is satisfactory in terms of of both model structure, coefficient interpretation and classification performance.

3.7.3 Split 3: Type 3, 4 vs. Type 5

The hit rates of the Split 3 model and the correspoinding SVM are listed in Table **??**. Compared to the Split 1 and 2 models, the hit rates of this model are a bit lower but acceptable. The confusion matrix of the calibration dataset is symmetrical, but that of the validation dataset is extremely asymmetrical. For the validation dataset, a much larger proportion of Class 0 samples are incorrectly classified into Class 1, but most of Class 1 samples are correctly classified, which implies some possible bias associated with the model. This phenomenon is substantially alleviated in the SVM, although the overall hit rates of the LR model and the SVM are very close.

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	88	12	Est	0	331	3	Est	0	91	11	Est	0	336	4
	1	22	82		1	83	23		1	19	83		1	78	22
83%				80%				85%				81%			

In this model, only $mean\_fric$ and $highspec\_prop\_fric$ are significant. As expected, higher $mean\_fric$ corresponds to higher probability of being Class 0. Higher $highspec\_prop\_fric$ corresponds to higher probability of being Class 1, which is consistent with its distribution patterns as explored in Section 3.5.6.

Similar to the Split 2 model, the coefficient magnitude of $mean\_fric$ in this model is overwhelmingly large, suggesting $mean\_fric$ is very important for this split model and the contribution by other variables could be marginal. This model is also acceptable in terms of validity and classification performance.

3.7.4 Split 4: Type 3 vs. Type 4

The hit rates of the Split 4 model and the corresponding SVM are listed in Table **??**. Its hit rates are close to those of the Split 3 model, but lower than those of the Split 1 and 2 models. The confusion matrices of the calibration and validation datasets are both asymmetrical, and the same phenomenon occurs with the SVM. This suggests at this split, the boundaries of Type 3 and 4 are fuzzier in the feature space and the estimation based on current features could face a large degree of uncertainty. In other words, these two RSC types are similar and it is more difficult to distinguish one from the other.

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	9	3	Est	0	0	46	Est	0	10	1	Est	0	1	46
	1	12	42		1	3	409		1	11	44		1	2	409
77%				89%				82%				90%			

In this model, $mean\_fric$ does not appear significant. The step-wise calibration process in R has chosen $highspec\_fric$ and $lowspec\_fric$ as the only significant variables. Higher $highspec\_fric$ corresponds to higher probability of being Class 0, which is consistent with the presumed physical interpretation that larger high-frequency spectra are associated with better RSCs. Higher $lowspec\_fric$ corresponds to higher probability of being Class 1, which is obviously in contradiction to the distribution pattern shown in Figure 3.10. This is difficult to interpret and may be attributed to some interaction effects not captured by the model. This model is acceptable in terms of classification performance, but some discretion should be given in both interpretation and application.

3.7.5 Split 5: Type 0 vs. Type 1

The sample sizes for model calibration and validation of Split 5 are both very small, so the model validity and estimation performance could be highly questionable; however, we can still get some interesting hints from the modelling results. For example, the hit rates of both the LR model and SVM shown in Table **??** beat the uninformative ratio of 50% for the calibration dataset, which suggests that both classifiers have shown certain capability in RSC estimation.

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	10	4	Est	0	3	5	Est	0	13	5	Est	0	2	8
	1	4	10		1	0	4		1	1	9		1	1	1
71%				58%				79%				25%			

The only significant variable in the model is sd_fric with a positive coefficient; i.e., larger standard deviation is associated with Class 1, which corresponds to RSC Type 1 (bare wet). This is consistent with our expectations based on its physical interpretation and distribution patterns across RSC types. Until more CFM data are collected on Type 0 and Type 1 road surfaces, it is difficult to give a more conclusive assessment on this model. It is easy to see that the hit rates and the structure of the current model both imply there is a large room for improvement.

3.7.6 Propagation Error

Besides the validation of each binary split model, validating the performance of the whole model structure is also a great concern, especially for the decision support in maintenance practice. Propagation error is evaluated for this purpose.

When a sample is input into the structured models, starting from the root of the tree shown in Figure 3.14, the binary model at each split is recursively applied to assign the sample to the appropriate branch until it reaches a certain leaf of the tree, which is its estimated RSC type. If the estimated type is not the true type, it is a case of propagation error. In this way, all the samples are propagated along the tree through several binary assignments, and the propogation errors for the calibrated logit models are summarized in

the confusion matrix shown in Table 3.5. The row names of this confusion matrix are the true RSC types and the columns names are the estimated types. Two columns are added to represent the total numbers and estimation accuracy (represented by hit rates) for all RSC types. Besides, two rows are added to represent the total numbers and estimation confidence (represented by specificity rates). The rate of specificity for each type is the number of cases correctly classified as this type divided by the total number of cases classified as this type.

Table 3.5: Propagation Errors of LR Models

	0	1	2	3	4	5	Total	Accuracy (%)
0	11	4	2	0	0	0	17	64.7
1	8	13	2	0	0	0	23	56.5
2	5	4	254	4	13	0	280	90.7
3	1	0	3	6	14	0	24	25.0
4	0	0	24	38	333	105	500	66.6
5	0	0	0	1	14	105	120	87.5
Total	25	21	285	49	374	210	964	
Specificity (%)	44.0	61.9	89.1	12.2	89.0	50.0		

The accuracy is relatively high for true types 0, 2, 4, and 5, and low for the other two types. As for the specificity, estimated types 1, 2 and 4 are relatively high.

From the perspective of practical maintenance decision, different types of propagation errors may have various consequences. For example, if the models are used to determine BP status, a reliable classification between bare (0 and 1) and non-bare (2, 3, 4 and 5) types is more important than other classifications. When the true condition is bare but estimated to be non-bare, it is a false alarm. On the other hand, if the true condition is non-bare but wrongly estimated to be bare, it is a miss. Observed from Table 3.5, it can be found that the false alarm rate is $4/40 = 10\%$ and the miss rate is $10/924 = 1.08\%$. The extremely low miss rate means the models can identify non-bare slippery road surface in a highly efficient manner. The 10% false alarm rate appears to be a little high, but the estimated types of those false alarms are all Type 2, which is very similar to the bare condition so that the cost of those false alarms is very low. These observations about the propagation errors suggest that this structure of models can be a powerful tool for BP status monitoring.

Another observation about the propogation errors is that the misclassifications mostly occur at the RSC types of similar severity. For example, 17 out of 18 misclassifications

of Type 3 cases are estimated as Type 2 or 4; 143 out of 167 misclassifications of Type 4 cases are estimated as Type 3 or 5; and 14 out of 15 misclassifications of Type 5 cases are estimated as Type 4. All those misclassifications are associated with low cost, which implies the performance of the structured models is reliable and they can be applied with confidence.

The propogation errors for the trained SVMs are summarized in Table 3.6. It is very similar to Table 3.5 in terms of accuracy, specificity, misclassification rates, false alarm rates and miss rates. This suggests that SVMs do not show significant advantage over logit models when organized and applied under the proposed multi-level structure.

Table 3.6: Propagation Errors of SVMs

	0	1	2	3	4	5	Total	Accuracy (%)
0	14	2	1	0	0	0	17	82.4
1	12	10	1	0	0	0	23	47.1
2	7	1	254	2	16	0	280	90.7
3	1	0	3	7	12	1	24	29.2
4	0	0	29	36	339	96	500	67.8
5	0	0	0	0	15	105	120	87.5
Total	34	13	288	45	382	202	964	
Specificity (%)	41.2	76.9	88.2	15.6	87.4	52.0		

3.8 RSC Estimation by Vehicle Speed

As mentioned at the beginning of this chapter, speed data are used separately to classify RSC types. Although the features in the speed data are totally different from CFMs, the definition of the study problem, the model framework and the modelling steps are very similar except that the basic classification data are the aggregate features of vehicle speed instead of CFMs.

The speed data available to this study were collected by the GPS device mounted on the friction meter operation vehicle. As discussed in Section 2.3.4, the ideal features for RSC estimation should not be closely related to the driving behavior. Obviously, speed is not such a feature, so the models calibrated in this section should be taken cautiously. The purpose of calibrating this set of models is to show a possible framework for such type of studies.

3.8.1 Aggregation Interval Length

The vehicle speed is studied under a very similar frame work as in Figure 3.1. The only difference is that all the aggregate measures are extracted from vehicle speed instead of CFMs. The aggregate measures from speed data are exactly the same, i.e.,

1. mean speed (of original series)
2. standard deviation (of original series)
3. skewness (of original series)
4. low-frequency spectra (of first differenced series)
5. high-frequency spectra (of first differenced series)
6. proportion of high-frequency spectra (of first differenced series)

Like what has been done in Section 3.6.3, the hit rates of preliminary LR models and SVMs calibrated with speed data are plotted against corresponding aggregation intervals in Figure 3.17 and 3.18.

For both classification schemes, the hit rates of logit models and SVMs increase with the aggregation interval length. The hit rates of SVMs are all higher than LR models as expected. When the aggregation interval length is longer than 500 meters, the hit rates of LR models level off, while the hit rates of SVMs keep increasing but at a lower rate. This pattern suggests an optimal aggregation interval between 500 to 1000 meters, which is very similar to what has been suggested by Figures 3.17 and 3.18.

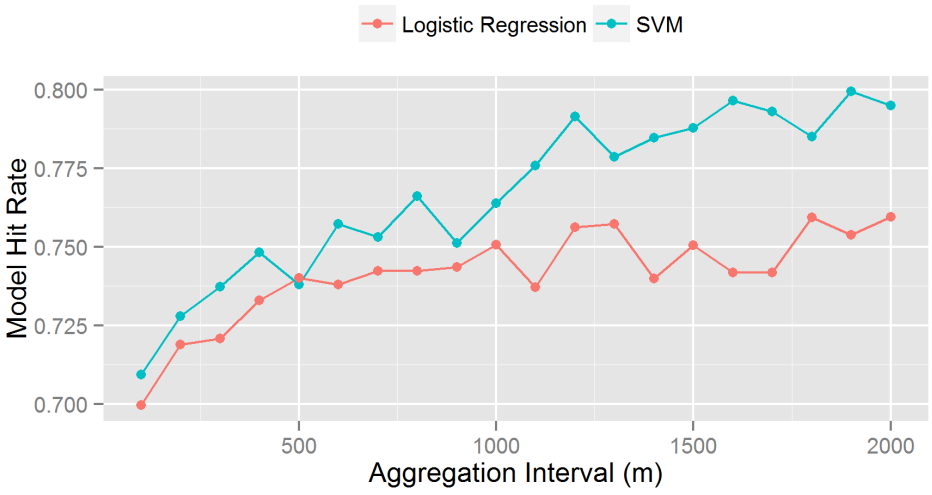


Figure 3.17: Model Hit Rates for Different Aggregation Interval Lengths (Speed Data) Type (0,1,2) vs. Type (3,4,5)

Compared to speed-based models with the same aggregation interval, CFM-based models have significantly higher hit rates, suggesting friction measurements can give more reliable RSC estimations than speed data. However, the reasonable hit rates of speed-based models suggest that vehicle speed may be used to explain some variations in road surface conditions.

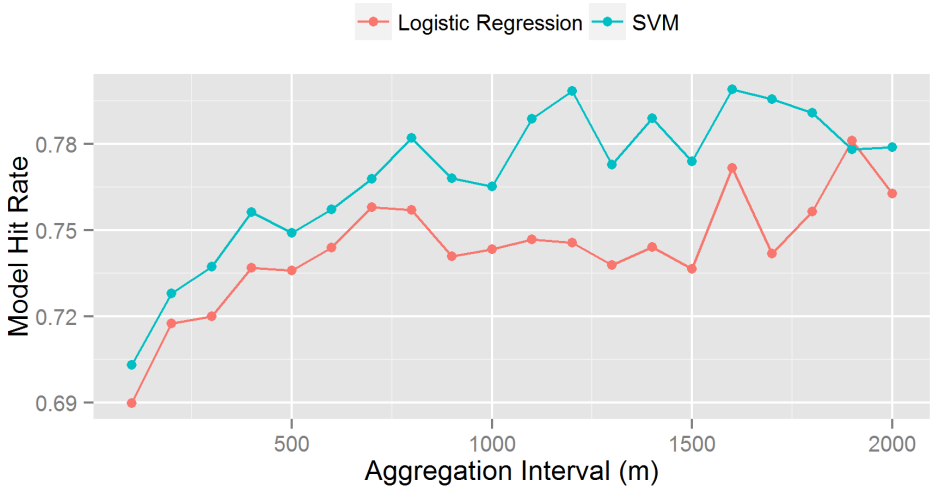


Figure 3.18: Model Hit Rates for Different Aggregation Interval Lengths (Speed Data) Type (0,1,2,3) vs. Type (4,5)

3.8.2 Exploratory Analysis

Similar to what has been done in Section 3.5, the aggregate features are examined by RSC types. Using exactly the same integration method, the interval lengths of 300m, 700m and 1000m are applied to the speed data.

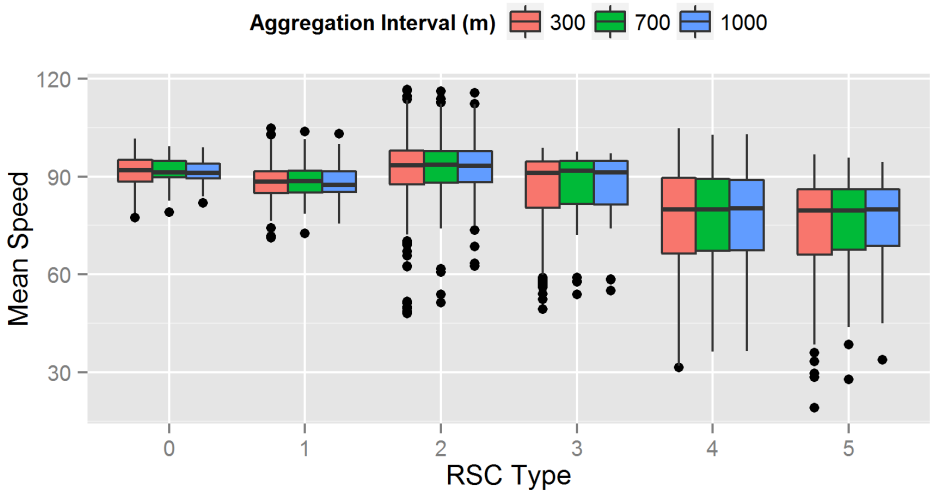


Figure 3.19: Mean Speed

The distribution of each extracted feature is plotted against the RSC types as shown in Figure 3.19 to 3.24. It can be found that

- Mean speed generally decreases when the RSC becomes worse but this could be due to the drivers' self-adjustment of the speed and is only an indirect reflection of RSC. The average levels of the mean speed of Type 0 to 3 are very similar, while those of Type 4 and 5 are also similar with a slightly lower average.
- Standard deviation, skewness and low-frequency spectra may not be good estimators of RSC, as their value ranges overlap a lot across all six RSC types.
- The high-frequency spectra of Type 4 and 5 are very small, while the other four types have similar levels of high-frequency spectra.
- The proportion of high-frequency spectra of Type 4 and 5 is similar and lower than the other four types and the other four types significantly overlap.

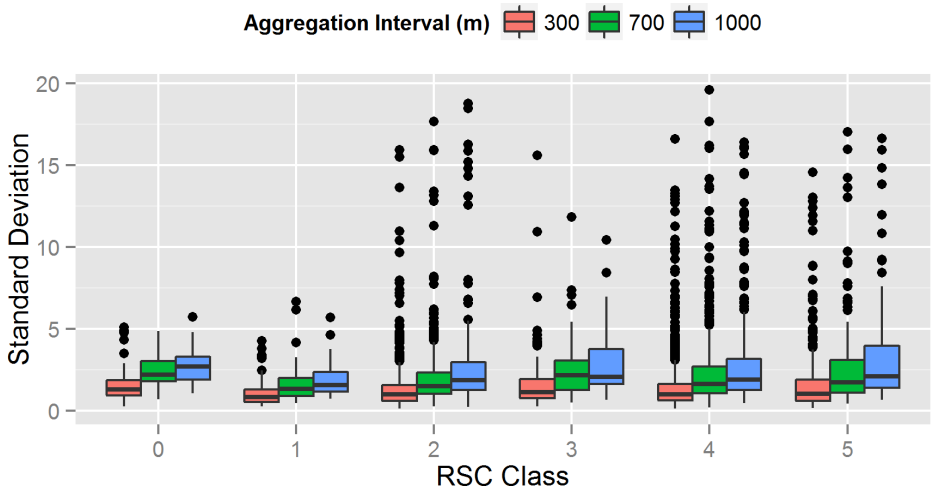


Figure 3.20: Standard Deviation of Speed

These observations suggest that, if mean speed is not used in model calibration, only the classification between Types (0,1,2,3) and Types (4,5) can be reliably made using the current set of aggregate features.

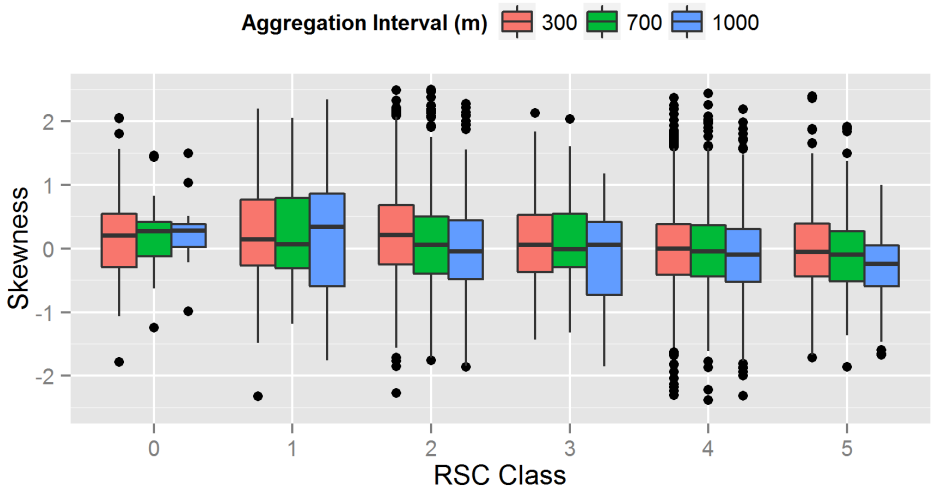


Figure 3.21: Skewness of Speed

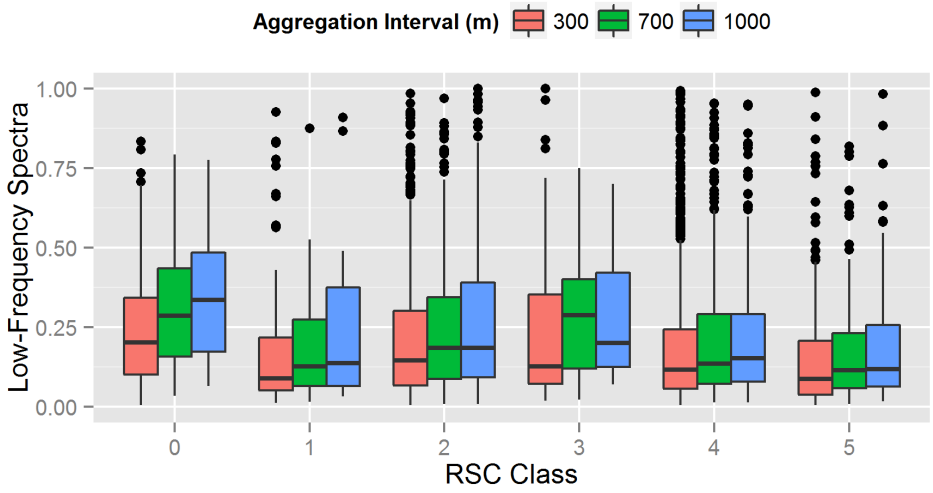


Figure 3.22: Low-Frequency Spectra of Speed

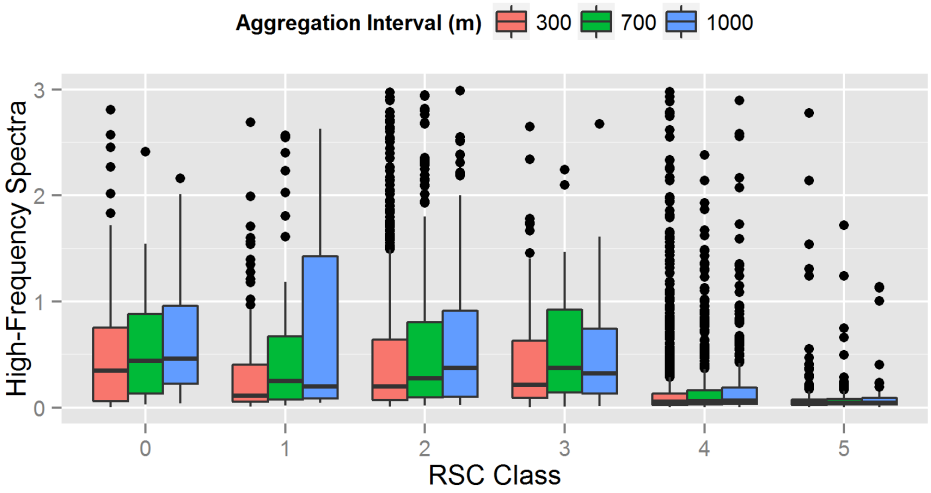


Figure 3.23: High-Frequency Spectra of Speed

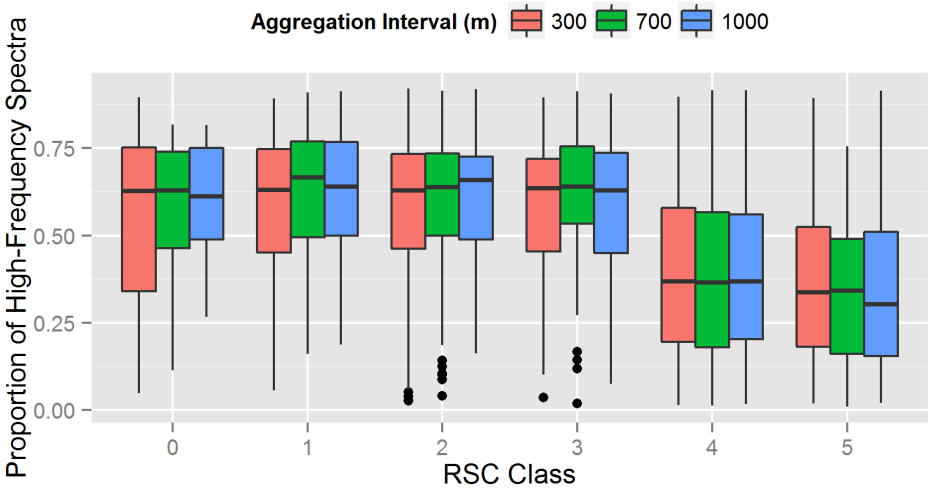


Figure 3.24: Proportion of High-Frequency Spectra of Speed

3.8.3 Model Calibration and Validation

The same nested structure as shown in Figure 3.14 is used for RSC estimation. For each split, a binary LR and SVM are calibrated using the following variables:

- Standard Deviation (sd_speed)
- Skewness($skew\_fric$)
- Low-Frequency Spectra ($lowspec\_speed$)
- High-Frequency Spectra ($highspec\_speed$)
- Proportion of High-Frequency Spectra ($highspec\_prop\_speed$)

Each aggregate measure of speed in the list corresponds to the same aggregate measure of CFMs but the mean speed is not used as we assume it is significantly affected by non-RSC factors, such as speed limits, traffic controls, roadway geometries, traffic flow characteristics and driving behaviors. The other variables could be correlated more closely with RSC, and thus are tested in our modelling process. Again, this decision is based on our presumption with no solid support from any empirical studies.

In fact, the model structure based on CFMs may not fit the speed data and should be adjusted according to the patterns revealed in the exploratory analysis on speed data. As in this study, the speed-based modelling is only for the purpose of demonstration. The same model structure and sample allocation scheme are used so that it is more convenient to compare the resulting models to the CFM-based models at individual splits.

The calibrated LR models based on speed data are summarized in Table 3.7. The hit rates of the LR models and SVMs for the calibration and validation datasets for all the splits are listed in Table 3.8 to Table 3.12. From the calibrated LR models, it can be found that

- No explanatory variables are significant for Split 2 and 5, which suggests that the tested aggregate features can not be used for RSC classification for these two splits.
- The hit rates of the LR models for Split 1, 3 and 4 are from 60% to 80%, which are acceptable for the testing purpose. The Split 4 model has the highest hit rates and the confusion matrices are symmetrical. This is expected because at this split, there are some good discriminators, like $highspec\_prop\_speed$, as shown in the exploratory analysis.

Table 3.7: Speed-Based LR Models for RSC Estimation

	Split 1	Split 2	Split 3	Split 4	Split 5
(*Intercept*)	4.14^{***}	0.83^{***}	-0.15	11.27^{**}	0.00
	(0.89)	(0.22)	(0.24)	(4.09)	(0.38)
skew_speed	$-2.47^{\cdot}$			$-10.15^{\cdot}$	
	(1.48)			(5.92)	
lowspec_speed	-3.49^{***}		$-4.36^{\cdot}$		
	(1.01)		(2.43)		
highspec_prop_speed	-4.15^{***}			-5.50^{***}	
	(0.48)			(1.61)	
sd_speed			2.70^{*}	-7.43^{**}	
			(1.22)	(2.84)	
highspec_speed			-5.80^{*}		
			(2.85)		
AIC	544.97	123.46	273.18	67.42	40.82
BIC	561.64	126.05	286.46	76.18	42.15
Log Likelihood	-268.48	-60.73	-132.59	-29.71	-19.41
Deviance	536.97	121.46	265.18	59.42	38.82
Num. obs.	477	99	204	66	28

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

- At all splits, the hit rates of speed-based models are equal to or lower than those of the CFM-based models except Split 4. It again concurs with the findings of most previous studies that friction level is a very reliable measurement of RSC.

On the whole, the RSC estimation based on speed data is substantially worse than CFM-based estimation, but at Split 4, the speed-based model has performed very well, suggesting vehicle driving data could be very helpful in RSC estimation if their correlations with RSC types are well understood and they enter the model after appropriate aggregation/preprocessing.

Table 3.8: Hit Rates of Split 1

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	163	72	Est	0	67	117	Est	0	164	63	Est	0	68	109
	1	71	171		1	19	284		1	70	180		1	18	292
70%				72%				72%				74%			

Table 3.9: Hit Rates of Split 2

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	0	0	Est	0	0	1	Est	0	0	0	Est	0	0	0
	1	30	69		1	10	211		1	30	69		1	10	211
70%				95%				70				95%			

Table 3.10: Hit Rates of Split 3

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	80	49	Est	0	268	17	Est	0	83	32	Est	0	269	17
	1	30	45		1	146	9		1	27	62		1	145	9
61%				63%				71%				63%			

Table 3.11: Hit Rates of Split 4

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	12	5	Est	0	1	83	Est	0	11	4	Est	0	0	54
	1	9	40		1	2	372		1	10	41		1	2	401
79%				81%				79%				88%			

Table 3.12: Hit Rates of Split 5

LR								SVM							
Calibration				Validation				Calibration				Validation			
		True				True				True				True	
		0	1			0	1			0	1			0	1
Est	0	0	0	Est	0	0	0	Est	0	12	1	Est	0	2	4
	1	14	14		1	3	9		1	2	13		1	1	5
50%				75%				89%				58%			

3.9 Summary

In this chapter, the relationship between CFM and RSC is systematically studied. The aggregate measures of CFMs are explored from different perspectives, such as aggregation length, probability distribution, patterns in the frequency domain, and so on. According to the exploration results and certain presumptions suggested by preliminary studies, a multi-level model structure is designed in order to accurately estimate RSC types using CFM data. Several aggregate measures of CFMs are carefully chosen and used as the explanatory variables in calibrating and validating a set of binary LR models. In the mean while, an SVM is trained corresponding to each LR model to benchmark its performance. The modelling results are acceptable and the aggregate measures have shown significant effects in most LR models.

Vehicle driving data, especially speed, are also tested under the similar model structure and methodology. The overall performance of speed-based models are not as good as CFM-based models, but there is some success in classifying certain RSC types.

Chapter 4

RSC Forecasting

As discussed in Chapter 2, the approaches used for RSC forecasting expand over a wide spectrum, from rule-based to data-driven, from purely physical to statistically rooted, and from short-term to long-term perspectives; however, each of them has some of the following limitations:

- the effects of de-icing operations such as salting or plowing are not considered,
- the set of input parameters is very large,
- the uncertainty in the snow/ice melting process is not systematically addressed.

These limitations could seriously affect the model's feasibility for supporting real-world maintenance decision making and performance measure. The RSC model framework of VTI (Möller, 2008; Wallman, 2004), as shown in Figure 4.1, is an attempt to overcome the above limitations.

The core of this model is to use the RSC at t and the forecast of relevant explanatory variables during the period from t to $t+1$ to forecast the RSC at $t+1$. Although the VTI model is the most comprehensively defined short-term RSC forecasting model framework up until now, several aspects of it need some adjustment or improvement. Specifically,

1. The RSC is defined as a set of discrete categories instead of numerical measures. The categorical RSC definition could easily introduce subjective bias from both modellers and model users, impairing the validity and performance of the model.

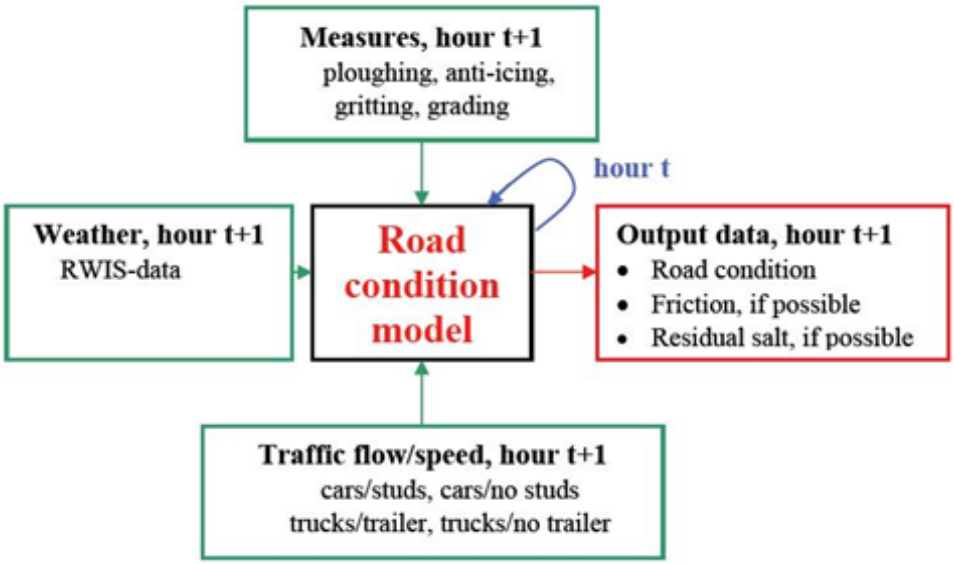


Figure 4.1: VTI RSC Model

2. The nature of the RSC changing process is essentially continuous instead of discrete in time, i.e. the depths of snow/ice contaminants usually change gradually under the influence of environmental factors. The forecast in the form of multi-class probability distributions can address this continuity nature indirectly, but it is subordinate to the forecast directly aiming at some numerical measure, like contaminant depths.

3. The model framework implicitly assumes that the RSC changing in time is a simple Markovian process. The future RSC is only conditional on the current RSC and related factors. The RSC at hour $t+1$ can be successfully predicted from hour t without any residual effect from the previous hours ($t-1$, $t-2$, ...). It is a very strong assumption having overlooked an important factor in the ice/snow melting process, namely the lasting melting effect of salt. In other words, the applied salt could take a much longer time than one hour to show its effect as melting takes time. Therefore, the knowledge of the amount of salt residual at t might not help a lot in predicting its melting effect at $t+1$.

To address the limitations of the VTI model, a new RSC forecasting model framework is proposed as described in Section 4.1. The data collection and aggregation process and the modelling methodology are respectively described in Section 4.2, 4.3 and 4.6. The model calibration and validation results are presented in Section 4.7 and 4.8.

4.1 Problem Definition and Modelling Framework

Due to different objectives of RSC modelling, the problem can be defined in a variety of constructs in terms of short or long term forecasting, stochastic or deterministic forecasting, discrete RSC classes or numerical RSC measure forecasting, considering maintenance operations or not, point-wise (study site is only a spot on the roadway) or macroscale (study site is one or multiple road sections) forecasting. The scope of this chapter is to make short-term forecasting at a fixed spot on the roadway surface. The factors incorporated in the framework include weather, traffic and maintenance operations, which is similar to the conceptual model proposed by Klein-Paste (2008). This scope is defined in consideration of the objectives of the overall thesis study, which is to address the estimation and forecasting accuracy and model applicability in real-world maintenance decision making and performance measure. In addition, this scope is also defined based on the data available for the study, which will be discussed in details in Section 4.2, 4.3 and 4.6.

The framework makes a Markovian assumption similar to that of the VTI model, but with the following major changes (Figure 4.2).

1. The RSC is defined as a set of numerical measures, which are surface temperature, friction level and contaminant layer depths. RSC types are not considered in both model calibration and application.

2. Applied salt is treated as an exogenous variable instead of an endogenous variable. It is considered as an external factor affecting the RSC changing just like weather variables, instead of as a part of the RSC measurement.

3. The RSC is predicted using its previous state and the historical observations of the relevant factors. Only one time interval of traffic, weather and plowing information is entered into the model, but a longer history of salting operations is considered.

As shown in Figure 4.2, more than one time interval of previous salting operations enter the model as it is assumed that applied salt may take more than one time interval to clearly show its melting effect. The one time interval from t to $t+1$ could be any constant length besides one hour, e.g. 20 minutes–the measuring time interval by most RWIS stations.

The model framework also assumes a Markovian process, of which a snapshot of the system state spans more than one study time interval ($t-p$ to t) in order to incorporate historical salting operations. The RSC is defined mainly by three measures:

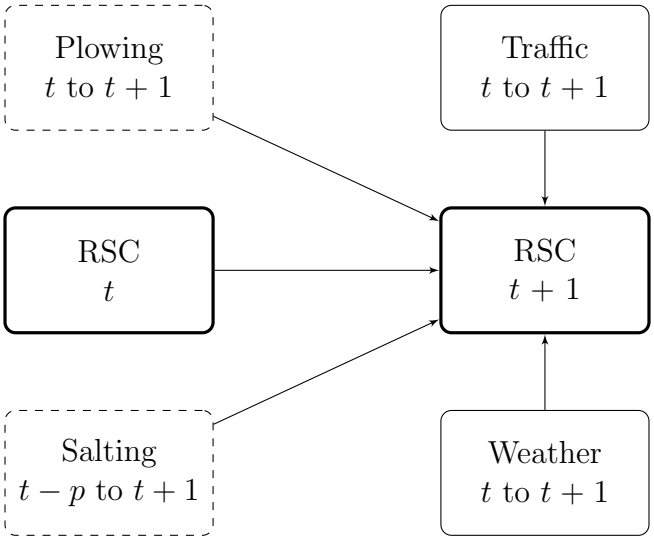


Figure 4.2: RSC Forecasting Model Framework

1. pavement contaminant depth
2. road surface temperature
3. road surface friction level

Detailed descriptions about each component in Figure 4.2 are given in Section 4.2 and 4.3.

4.2 Data Collection

The data required by the model framework in Figure 4.2 consist of four major parts:

1. numerical measures of RSC
2. weather information
3. traffic volume
4. plowing and salting operation records

A high quality model requires that all four parts of the data be collected at or very close to the same spot on the roadway. After a diligent process of searching and requesting, especially with our research partner, Ministry of Transportation Ontario (MTO), it is found that no study site in Ontario has the complete dataset for the scope of this study. The reason and the solution to this problem are discussed as follows.

4.2.1 RSC Measures

RSC measures are the foundation of the model framework. Most installed RWIS pavement surface sensors in Ontario do not measure the contaminant depth. At the data collection stage of this study, only one RWIS station on Highway 417 could give this measure. This RWIS system is produced by Vaisala and the surface sensor set is different from other traditional ones in both measuring mechanism and monitored factors. This surface sensing system consists of two new remote optical sensors (collectively called Vaisala Spectro/Cyclo sensors): namely, the Vaisala Remote Road Surface State Sensor (DSC111) and the Vaisala Remote Road Surface Temperature Sensor (DST111), as shown in Figure 4.3.

DSC111 is an active near-infrared band (-1 μm) remote sensor, which sends infrared light beams to the road surface and detects the backscattered signals at selected wavelengths. Based on observed differences in light absorption, it can differentiate pavement contaminant layers and thus surface states such as dry, moist, wet, icy, snowy/frosty or slushy. The sensor also provides a measure called grip level which is intended to represent the level of friction of the road surface. This grip level is estimated using an empirical model based on the surface state and contaminant depth detected by the sensor. DSC111 reports the following data items:

- pavement states: dry, wet (thin water layer), slushy (thick water layer, no ice or snow), snow or frost (white ice), ice (black ice)

(a) DSC111 (Spectro Sensor) (b) DST111 (Cyclo Sensor)

Figure 4.3: Image of Vaisala Sensor Set

- pavement contaminant depths in equivalent liquid water amount (in mm)
- estimated surface friction level (0.01-1.00)

DST111 is a temperature sensor based on infrared technology. It measures the difference of long wave infrared radiations between the sensor instrument itself and the road surface. This difference can be calibrated to a known temperature difference and thus used to estimate the pavement temperature. According to its product specifications, DST111 is accurate up to 0.3 °C in typical icing conditions. DST111 reports the following data:

- pavement surface temperature in °C
- air temperature in °C
- dew point temperature in °C
- relative humidity in percentage

The detailed specifications for DSC111 and DST111 are attached in Appendix A. The most recent development of both sensors can be found on the Vaisala website (http://www.vaisala.com). Pilli-Sihvola et al. (2006) and Feng & Fu (2008) have evaluated both sensors in their field studies and confirmed that their performance was fairly acceptable.

This Vaisala sensor set is located beside the eastbound lanes of Highway 417 near Casselman, Ontario, as shown in Figure 4.4. The monitored spot is right beside the sensor

set on the roadway and at the downstream direction of the Nation River Bridge. The sensors are installed on a pole at the roadside and are vertically 8.3m over the roadway level. The monitored spot covers a 20cm-diametered area near the right wheel track of the lane. The installed sensor and the monitored area are shown in Figure 4.5.

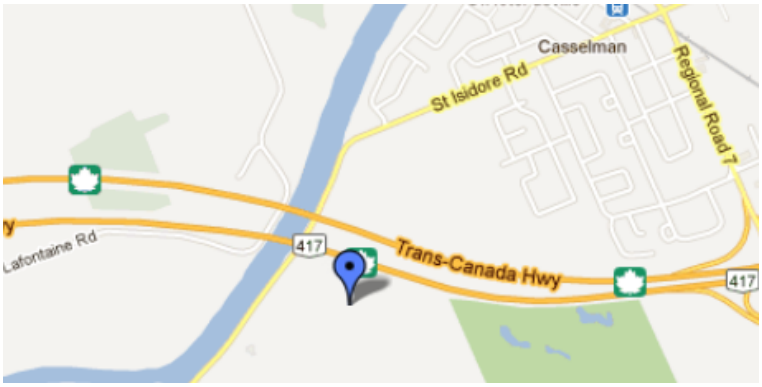


Figure 4.4: Location of Vaisala Sensor Set

(a) Sensor Installation (b) Monitored Area

Figure 4.5: Vaisala Sensor Installation and Monitored Area

The pavement contaminant depth reported by DSC111 consists of three components, which are

1. water layer depth
2. ice layer depth
3. snow layer depth

DSC111 also reports the estimated friction levels. As DST111 reports road surface temperature, this sensor set gives all the RSC measures needed in this study. So the data of other factors indicated in the model framework are collected based on this study site and the study period conforms to the available RSC dataset from this Vaisala site.

4.2.2 Weather Data

There is a traditional RWIS station located right beside the Vaisala station, which is shown in Figure 4.5. It reports data on a number of weather variables, such as air temperature, relative humidity, windchill and so on. In this study, precipitation rate, a very important variable, is available from this RWIS station but only over a limited time period.

4.2.3 Maintenance Data

As the study is concerned with short-term RSC changing considering the effects of maintenance operations, the spatial and temporal resolution of the maintenance data should be as high as possible. However, the available maintenance operation database of MTO, called MMIS (Maintenance Management Information System) (Ministry of Transportation Ontario, 2003), does not record the exact location and time of each plowing or salting operation. Therefore, the available AVL (Automated Vehicle Location system) data of maintenance vehicles are used (http://www.mto.gov.on.ca/english/transtek/m02-03/02-03fs.shtml).

The AVL system tracks the maintenance operations conducted by MTO and contractor vehicles. For location tracking, the AVL system reports the time and location of the monitored vehicles with a high frequency (mostly from several seconds to 30 seconds). For the salting operations, it tracks the application rate of both dry salt and pre-wetting salt solution. There are three maintenance vehicles doing plowing or salting at the study site. Their information is listed in Table 4.1, which is copied from MTO maintenance equipment description sheets.

Table 4.1: Maintenance Vehicles at Testing Site

Name	Equipment	Task
Plow 78-7	Plow Truck 14' Reversible	Plowing
Combo 78-8	Combo 7.6 m3 Dual Spinner	Salting
Combo 78-9	Combo 7.6 m3 Dual Spinner	Salting ramps

Plow 78-7 only conducted plowing and Combo 78-8 salting. Combo 78-9 conducted salting as well but its main task was to apply salt on ramps and bridges. Although the study site is not on any ramp or bridge, it is very close to Nation River Bridge, and Combo 78-9 could also be dispatched to salt the section of the route containing the study site. Therefore, the data of Combo 78-9 are also considered for the study.

4.2.4 Traffic Data

After checking available RSC, weather and maintenance data, it was found that the time period for which all those data are available is only one winter season–November, 2007 to March, 2008. Therefore, this time period is designated as the study period. After searching

the traffic volume data for this period near the study site, it was found that the data of only one traffic counter in Ottawa covered this period. Although on Highway 417, this traffic counter is 66.4 km from the study site, and the city of Ottawa is in between. The difference of the traffic volume between this location and the study site could be very large. There is another traffic counter at Dunvegan, which is only 17.4 km from the study site, and there is no big town in between. This traffic counter only has the data up to April, 2006, but the traffic volume after that can be estimated using Ottawa's data if some model can be built using the historical data of both traffic counters. The locations of the traffic counters at Ottawa and Dunvegan, and the study site are shown in Figure 4.6. Detailed procedure for this model calibration and traffic estimation are given in Section 4.3.

Figure 4.6: Traffic Counters Close to Study Site

4.2.5 Summary

After the availability of the data is examined, the study site and the study period is determined as:

- Study Site: Vaisala monitored spot on Highway 417 (45°18'11.81"N, 75°6'2.52"W)
- Study Period: Winter season 2007-2008 (Nov.2007-Mar.2008)

Data sources and respective collection procedures are summarized in Table 4.2.

Table 4.2: Summary of Data Collection

Data	Source	Collection Procedure
RSC	Vaisala sensor set	downloaded from Vaisala website
Weather	RWIS station	obtained from MTO
Maintenance Record	AVL system	downloaded from AVL data archive website
Traffic Volume	Traffic Counters	obtained from MTO

4.3 Data Integration

4.3.1 RSC and Weather Data

Most Vaisala and RWIS data are spaced 20 minutes apart, so 20 minutes is determined as the study time interval, i.e., one time interval in the model framework is 20 minutes. All Vaisala and RWIS data are integrated according to this time interval, and a long time series of RSC and weather datasets is generated for the study period. The data fields include the following variables:

- air temperature
- precipitation
- pavement surface temperature
- water depth
- ice depth
- snow depth
- friction level

There are other variables available, like relative humidity and windchill. As their effect on RSC changing is expected to be relatively small compared to air temperature and precipitation, they are not considered in further modelling analysis.

4.3.2 Maintenance Data

The timestamps for the maintenance vehicles passing the study site are calculated by some interpolation algorithm. As the AVL data have a very high temporal resolution, the error of these timestamps is mostly under 30 seconds, which is trivial compared to the 20 minute study time interval. If the vehicle is applying salt or salt solution, the application rate is extracted. Afterward, the salt application rate of salting operations at the study site is inserted into the RSC/weather time series dataset at the location where its timestamp corresponds to the row timestamp.

For the plower, the AVL data do not have any record indicating active plowing operations. Therefore, a simple rule is applied to estimate its behavior: i.e., when the plower is passing the study site, if the sum of the water, ice and snow layers is more than 0.4mm, and the timestamp is more than 20 minutes later than the last salting operation, the plower is

doing plowing. Otherwise, the plower is just passing without plowing. The 0.4mm threshold value is chosen after manually examining the change of the contaminant depths before and after the plower operations.

4.3.3 Traffic Data

As the traffic data at Dunvegan does not cover the study period, an ordinary regression model is calibrated to relate traffic at this site for the study period to those at a counting station near Ottawa. The data used for model calibration are the hourly traffic volume at the two traffic counters from 2004-03-28 to 2006-04-30. The R output of the model calibration is given in Appendix B. The model can be written as

$$\ln Y = 1.68 + 0.17 \times \ln X + 0.04 \times (\ln X)^2 \tag{4.1}$$

i.e.

$$Y = e^{1.68+0.17\times\ln X+0.04\times(\ln X)^2} \tag{4.2}$$

where X is the hourly traffic at Ottawa and Y is the hourly traffic at Dunvegan.

The model has shown that the traffic volumes at these two locations are highly correlated and the model itself explains more than 86% of total data variation. This suggests that the estimation made by this model is acceptable. The calibration data and the calibrated function curve are plotted in Figure 4.7. When the Ottawa traffic volume is larger than 3000, the variance of the estimation increases significantly. Except this estimation, there is no better source of traffic information at the study site for the study period, so it was used for the modelling work at the current stage. Ideally, observed traffic data should be used instead of the estimated data.

The hourly traffic volume at Dunvegan is estimated using the traffic data at Ottawa with this model, and this estimation is used to represent the traffic volume level at the study site. Obviously, it could be much different from the true traffic volume at the study site. However, we can reasonably assume that the correlation of the traffic volume between the study site and Dunvegan is high, and the variation at the study site can be surrogated to a large degree by the variation at Dunvegan. It is therefore expected that using traffic volume at Dunvegan is similar to using traffic volume at the study site for RSC modelling, especially when the RSC forecast model is essentially linear. The only difference in the RSC forecast model will be the scale of the coefficient of the traffic term.

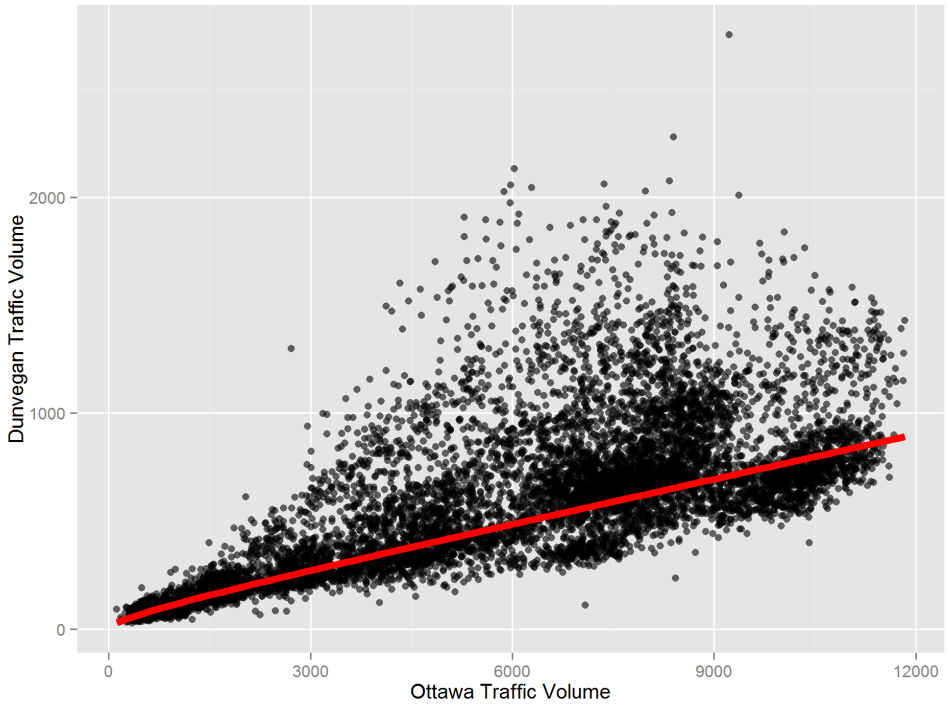


Figure 4.7: Traffic Volume from Traffic Counters

4.4 Data Summaries

The RSC, weather, maintenance and traffic data are integrated into a single data table for further analysis. As mentioned before, each row of the table has a timestamp and the time interval between the rows is an even 20 minutes. The columns of the table correspond to the endogenous and exogenous variables, where the endogenous variables are five RSC measures (surface temperature, friction level, water depth, ice depth, snow depth) and the exogenous variables are air temperature, precipitation, plowing operation, application rate of dry salt and hourly traffic volume. The application rate of the salt solution is not considered due to the following facts:

- In the study dataset, salt solution is mostly applied together with dry salt, i.e., salt is applied in a pre-wetted manner. This is one of the limitations of the study dataset, as the effects of pre-wetting vs non pre-wetting salting operations can not be compared.
- The salt solution is usually of 2% to 5% concentration for the pre-wetting purpose. Hence, the amount of salt in the solution is insignificant to that of the dry salt applied at the same time.

Therefore, this study will only examine the de-icing effect instead of anti-icing or pre-wetting effect. The concepts of de-icing, anti-icing and pre-wetting are discussed in Chapter 2.

4.4.1 RSC Variables

In total, there are 8277 data rows, which correspond to 8277 20-minute time intervals during the study period from November 20, 2007 to March 31, 2008. During a large proportion of the study period, there is no contaminant on the pavement surface. In order to show a clearer picture of the information about the RSC measures, those data rows are excluded from this analysis. The summary statistics of five RSC variables are shown in Table 4.3.

"n" in the table is the number of data rows considered by the summary. Without special notes, this meaning does not change in all the summary tables in this chapter. The value range of each RSC variable in this subset of the full data is fairly large, which gives more confidence to the reliability and validity of the modelling results. The histograms for the RSC variables in Figure 4.8 and 4.9 show their dispersions more intuitively.

Table 4.3: Summary of RSC measure Variables

	n	mean	sd	median	min	max	range
Surface Temperature (°C)	2514	-4.60	4.93	-4.20	-18.21	10.60	28.81
Friction Level	2514	0.34	0.27	0.21	0.09	0.82	0.73
Water Layer (mm)	2514	0.49	0.82	0.08	0.00	6.52	6.52
Ice Layer (mm)	2514	0.13	0.18	0.05	0.00	1.32	1.32
Snow Layer (mm)	2514	0.93	1.43	0.00	0.00	5.30	5.30

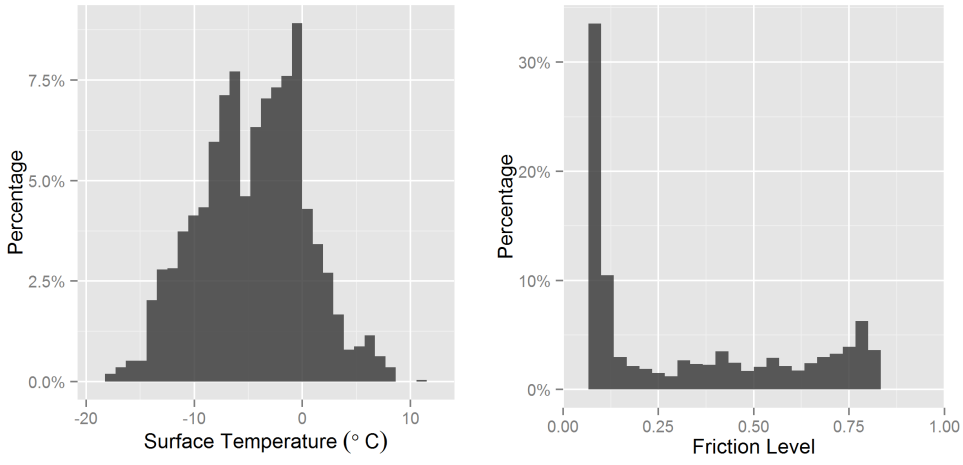


Figure 4.8: RSC Variable Distributions: Surface Temperature and Friction Level

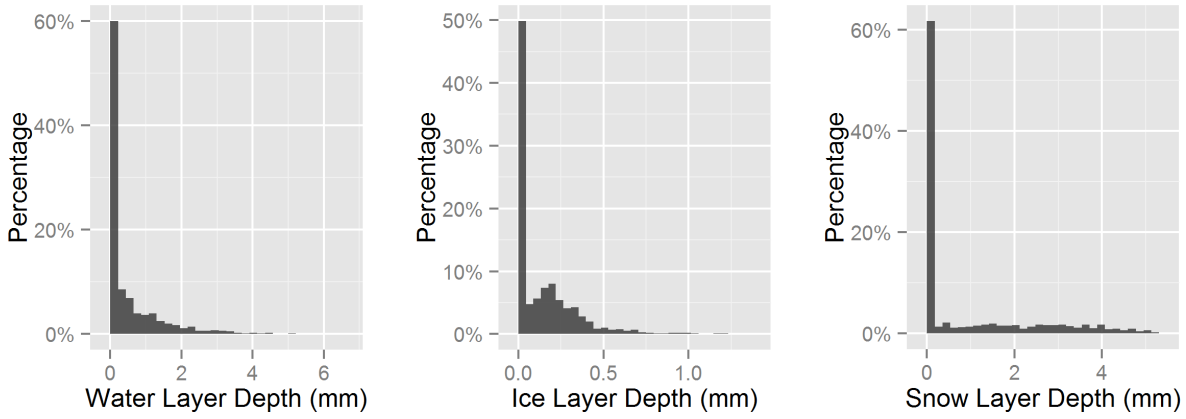


Figure 4.9: RSC Variable Distributions: Water, Ice and Snow Layers

4.4.2 Weather Variables

Only two weather variables are considered in the modelling work: air temperature and precipitation. The reasons for this decision are summarized as follows.

- These two variables are the most important weather factors affecting RSC changes, which are suggested by most previous studies.
- Some seemingly important variables, like relative humidity, are highly correlated with precipitation. When more than one highly correlated variable enters a regression-style model, the phenomenon of "multicolinearity" occurs, which makes the coefficient estimation highly unreliable (Faraway, 2002).
- Other variables, like wind and windchill, may have an effect on RSC, but this effect is likely to be much smaller in magnitude than air temperature and precipitation. Even if their effects end up significant in the resulting model, it could be very difficult to explain their causal effects on RSC changing.

The statistical summaries of air temperature and precipitation are shown in Table 4.4. Again, the data rows with zero precipitation are excluded from this summary in order to show a clearer view of the distributions.

Table 4.4: Summary of Weather Variables

	n	mean	sd	median	min	max	range
Air Temperature (°C)	1241	-3.9	5.2	-3.6	-22.4	11.0	33.4
Precipitation (mm/20 minutes)	1241	1.8	1.8	1.0	0.0	13.0	13.0

Additionally, the distributions of both variables are shown in Figure 4.10. Although the data are only for one winter season, the air temperature covers a fairly large range (from -22.36 to +11 °C) and so does the precipitation.

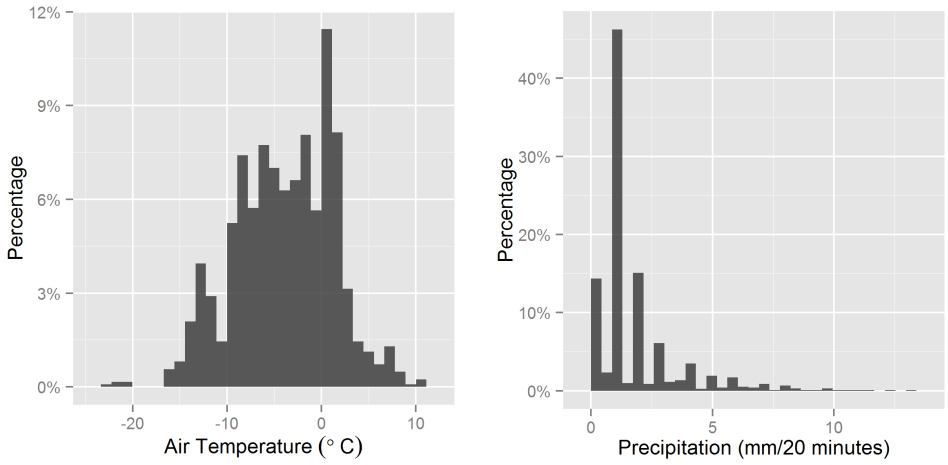


Figure 4.10: Weather Variables Distributions

4.4.3 Maintenance Operations

During the study period, there were 200 plowing operations and 305 salting operations at the study site.

As indicated in Section 4.3.2, the plowing operations are identified by a rule of thumb. According to a close review of the timestamps when the plowers pass the study site and the changes of the pavement contaminant depths before and after those timestamps, it is then determined whether or not there were plowing operations.

An interesting fact is that 143 out of 200 plowing operations are just seconds or minutes before salting operations, which reflects a common salting practice in Ontario. A plower is operating right in front and removes pavement contaminant before the salting vehicle behind applies the salt to the road surface. Another interesting fact is that most plowing operations are three hours later than previous salting operations, which is also a common practice to avoid the waste of salt residual on the roadway.

These two facts largely affect how the plowing and salting operations will appear in the RSC model and how the model is explained and applied. For instance, if a plowing operation and a salting operation take place during the same 20-minute interval, it is very difficult to isolate their effects on RSC changing for this time interval.

Since a salting operation takes place after a plowing operation, and there are no plowings in the following three hours in most cases, the effect of this salting operation can be isolated and studied without any confounding effect with the plowing operation for nine continuous 20-minute intervals (3 hours) except for the first time interval when the plowing takes place and the time intervals after another plowing operation takes place.

In other words, with the maintenance data of this pattern, it is likely that the calibrated model may not clearly disclose the instantaneous effect of plowing, but it should be able to show the effect of salting more clearly except for the time interval when the salting operation takes place.

Table 4.5: Summary of Salting Rates

	n	mean	sd	median	min	max	range
Dry Salt Rate (kg/km)	305	128.65	27.14	130.00	6.00	260.00	254.00
Solution Rate (l/tonne of salt)	305	34.02	16.57	39.00	0.00	82.00	82.00

The application rates of dry salt and salt solution are summarized in Table 4.5 and their distributions are shown in Figure 4.11. For both dry salt and salt solution, the

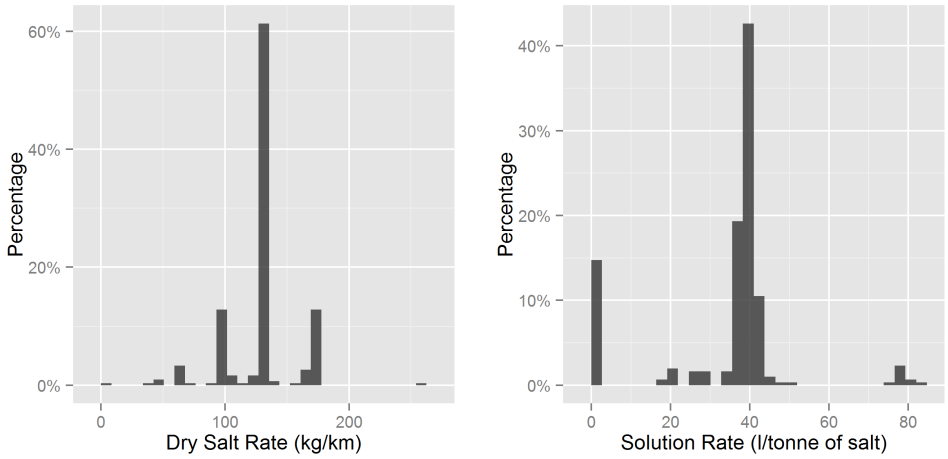


Figure 4.11: Salting Rate Distributions

application rates are concentrated in certain discontinuous ranges, which implies that the application rates are usually preset before the salting runs start. This observation has been confirmed by a review of the maintenance data. Also, this fact makes the spatial-temporal interpolation of the salt application at the study site much easier.

4.4.4 Traffic Volume

The estimated hourly traffic volume data at Dunvegan is summarized in Table 4.6, and its distribution is shown in Figure 4.12.

Table 4.6: Summary of Traffic Volume

	n	mean	sd	median	min	max	range
Traffic(vehs/lane/hour)	305	249	151	262	41	561	520

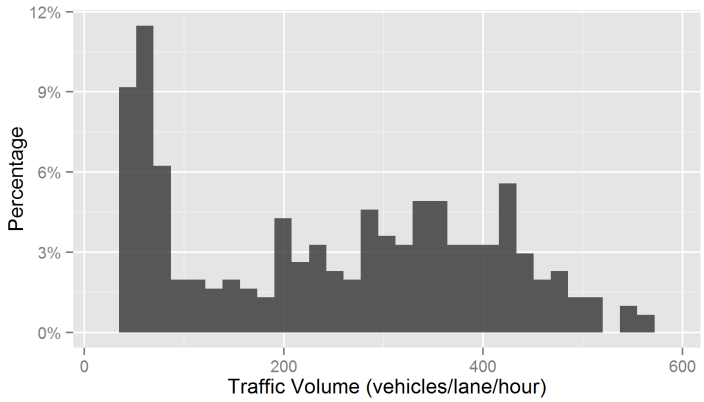


Figure 4.12: Hourly Traffic Volume Distribution

The traffic data are estimated by the traffic counts at a distant location, and they only approximately represent the variations of local traffic. Therefore, we are not reading too much into its distribution if there are no obviously anomalous patterns.

4.5 Exploratory Analysis

It is the most challenging part of the exploratory analysis to uncover how the RSC is related to other factors in an intuitive way using real world data. The reason is that many factors could jointly cause RSC changing: that is, the observed change could result from the aggregate effect of those factors. The simultaneous plowing and salting operation discussed in Section 4.4.3 is a typical example. To deal with this challenge, the complexity of the RSC changing process is explored from different perspectives. The findings together give a more complete picture, providing clues on the appropriate modelling methodology and techniques.

4.5.1 Surface Temperature and Air Temperature

Surface temperature could be the first and foremost factor affecting RSC changing, as it directly determines the state transformation of pavement contaminants. According to common physics knowledge, the variation of the surface temperature follows that of the air temperature, and shows very similar periodical patterns to the air temperature. The lengths of cycles are usually daily and yearly, i.e., a short cycle of 24 hours and a long cycle of 1 year.

Figure 4.13 shows the surface temperature as a time series during one week (2007-11-20 $\sim$ 2007-11-27). Although one week is too short to uncover any yearly periodical pattern, the daily periodical pattern is clearly shown. Other than the daily periodical pattern, the surface temperature is highly correlated with the air temperature.

An advanced approach to exploring a periodical time series is the non-parametric classical decomposition method (Chambers & Hastie, 1992), which assumes the time series is generated by the following random process:

$$Y_t = T_t + S_t + I_t \tag{4.3}$$

where $S_{t+d} = S_t$, $\sum_{t=1}^{d} S_t = 0$, $E(I_t) = 0$

S is the seasonal component, T is the trend component, I is white noise, t is the time index and d is the length of a cycle. The classical decomposition on the 20-minute intervaled surface temperature is conducted using R and the results are shown in Figure 4.14.

The raw time series data are plotted in the top panel of Figure 4.14. The third panel from the top is the discovered trend after the smoothing operation. The seasonal component is shown in the second panel from the top. There is an obvious 24-hour (72 points)

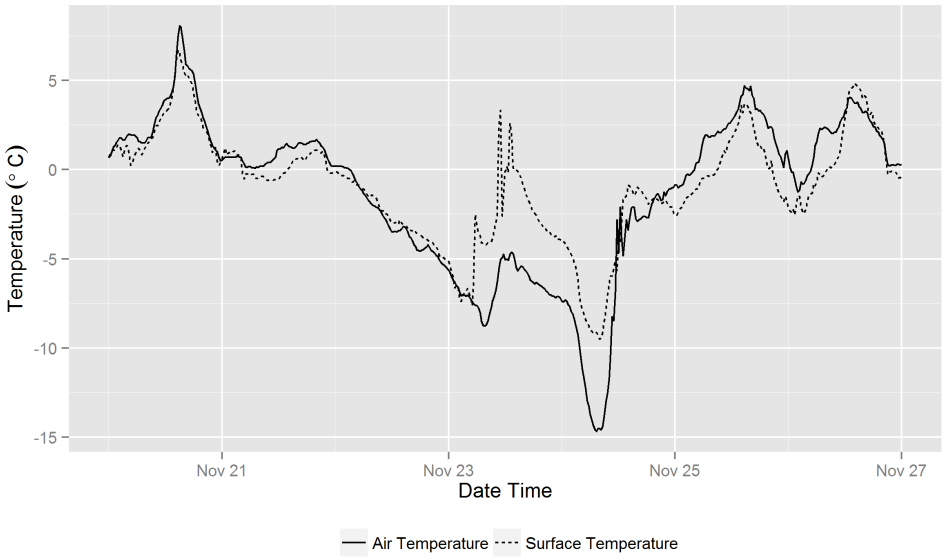


Figure 4.13: Surface Temperature of One Week

periodical variation with the peaks at early afternoon, and the bottoms at night time. The bottom panel shows the remainders (or residuals) which cannot be explained by the decomposition model. Although most remainders are within 2 standard deviations to zero, they are auto-correlated, i.e., the temporally neighbouring remainders are more similar.

The classical decomposition of the surface temperature time series suggests that an appropriate model should be able to efficiently capture the temporal correlations. If the model can also use the information of other covariates, as suggested by Figure 4.13, both its theoretical basis and forecasting performance will be further improved. For this purpose, a popular time series analysis technique, ARIMA, is applied as discussed in Section 4.6.

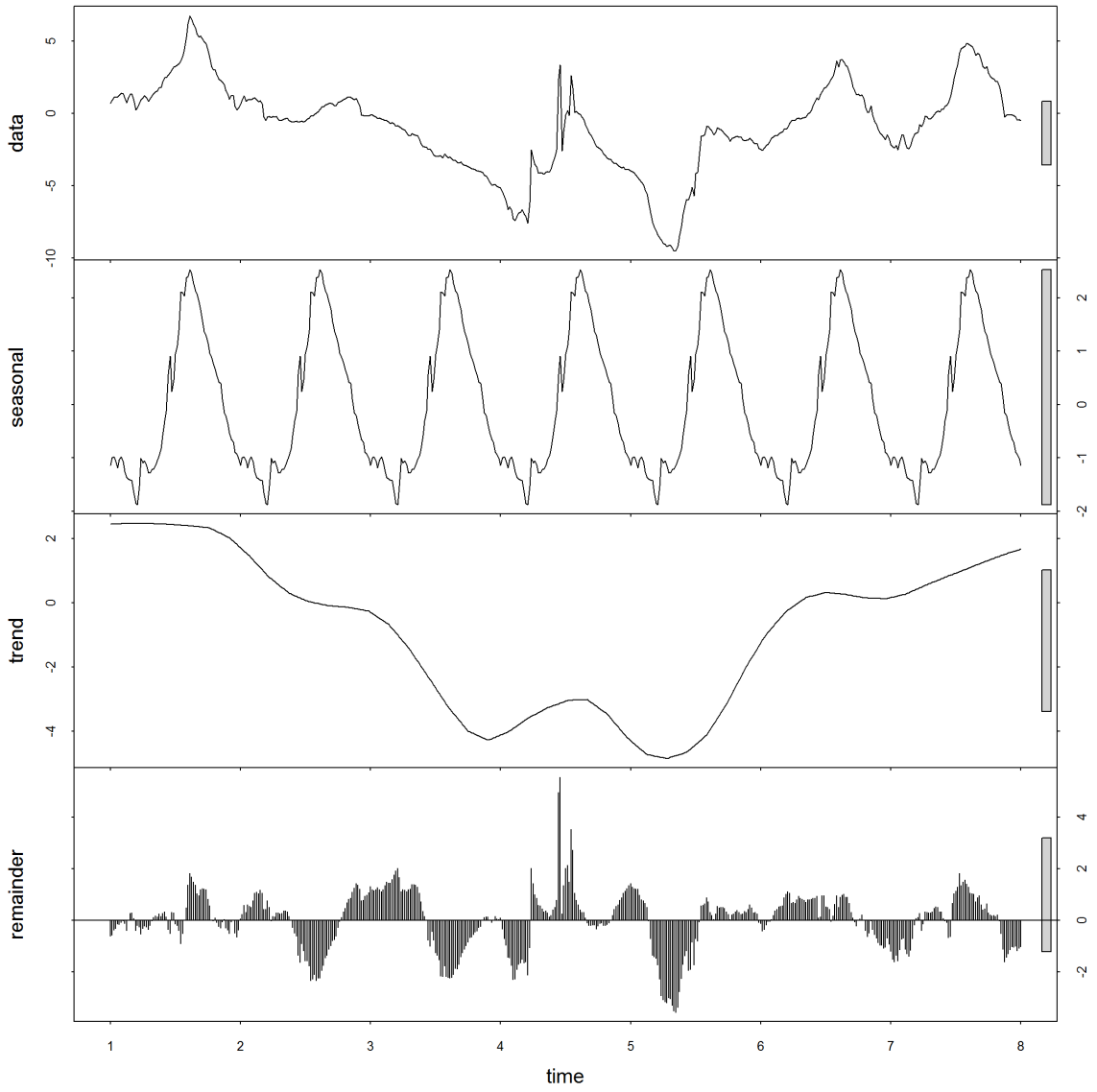


Figure 4.14: Classical Decomposition of Surface Temperature Time Series

4.5.2 Friction Level and Contaminant Depths

The friction level reported by the Vaisala sensor is not the real measurement; instead, it is estimated using current pavement contaminant depths. The estimation is made by an empirical model, which was calibrated in Vaisala's field studies with a certain type of decelerometer. The estimated friction level against the depths of three contaminant layers is plotted in Figure 4.15.

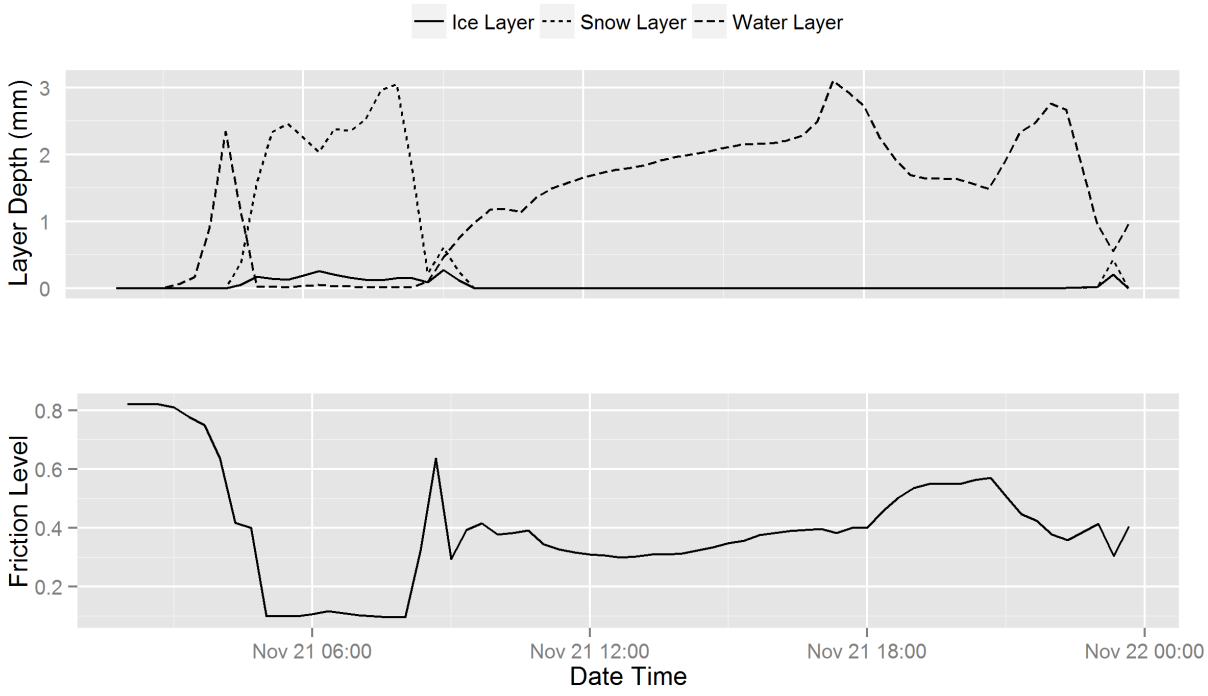


Figure 4.15: Layer Depths and Friction Level

The friction level is negatively correlated with ice and snow layer depths; however, the relationship to water layer depth is not clear. As the friction level is deterministically estimated by contaminant depths with the Vaisala proprietary friction model, it is not the real friction measurement. However, calibrating a forecasting model for friction levels can be used to demonstrate how real friction measurements, if available, can be modelled under the whole study framework.

4.5.3 Water, Ice and Snow Layers

The three layers of contaminant corresponding to water, ice and snow, are the center piece of the RSC model framework. Based on physical knowledge, the RSC changing process is essentially a freezing/melting process between water and snow/ice. These three types of contaminant interchange from one to another under different environmental conditions. It can be naturally assumed that they are interrelated temporally, i.e., the current depth of one type of layer is correlated to the depths of another two types of layers in the previous time interval. In addition, similar to surface temperature, the depth of each single type of layer should also show auto-correlation, i.e., each is correlated with its own previous depth. When multiple time series are correlated in time, cross-covariance and cross-correlation functions are used to examine the leading and lagging effects between time series. The sample cross-covariance function of two time series TS_1 $(x_1, x_2, \ldots, x_t, \ldots, x_n)$ and TS_2 $(y_1, y_2, \ldots, y_t, \ldots, y_n)$ is calculated as

$$\gamma_{xy}(h) = cov(x_{t+h}, y_t) \tag{4.4}$$

$$= \frac{1}{n}\sum_{t=1}^{n}(x_{t+h} - \mu_x)(y_t - \mu_y) \tag{4.5}$$

and the sample cross-correlation function (CCF in short) is then derived as

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}} \tag{4.6}$$

$$= \frac{\gamma_{xy}(h)}{\sigma_x \sigma_y} \tag{4.7}$$

The sample cross-covariance function and the CCF of two time series are the counterparts of the sample auto-covariance function and the ACF of a single time series. The meaning is very similar, except that they are used to explain relationships of lagged or leading values between two time series. The above definitions can be found in most of the time series literature (Shumway & Stoffer, 2006; Box & Jenkins, 1990; Mickens, 1991).

The 3×3 ACF/CCF plot matrix of water, ice and snow layer depths is shown in Figure 4.16. The three plots on diagonal lines are the respective ACFs of the water layer, the ice layer and the snow layer. They all show strong auto-correlation with very long lags (all greater than 30). Other plots show pairwise CCFs. The title of each CCF plot shows the relevant two layers and the relative lagging direction of the CCF.

Take the CCF plot at location row 1 and column 2 ([1,2] in short) for example. Its title

is "WL & IL", which means the CCF plot is for the water layer and the ice layer, and the value at lag k in the plot represents the CCF between WL_{t+k} and IL_t. As for this plot, most values before lag $k = 20$ are significantly positive, suggesting the depth of the ice layer at time t is positively correlated with that of the water layer at most time intervals in $t, t+1, \ldots, t+20$. This CCF pattern corresponds with the process of ice being melted into water.

The CCF plot at [2,1] corresponds to the CCF of the reverse process. It shows that the depth of ice is not correlated with previous water depth, which suggests that the water is usually promptly gone from the road surface either by runoff or maintenance operations and does not refreeze. The CCF plots at [2,3] and [3,2] show that the depth of snow is correlated with previous ice depth and vice versa, suggesting that snow and ice are inter-transformable.

The CCF plots at [1,3] and [3,1] represent the interchanging process between water and snow, but the significant lags in them are hard to explain with physical knowledge. Statistical analysis often faces this problem which could be caused by measurement errors, confounding effects, small sample sizes or some unseen reasons.

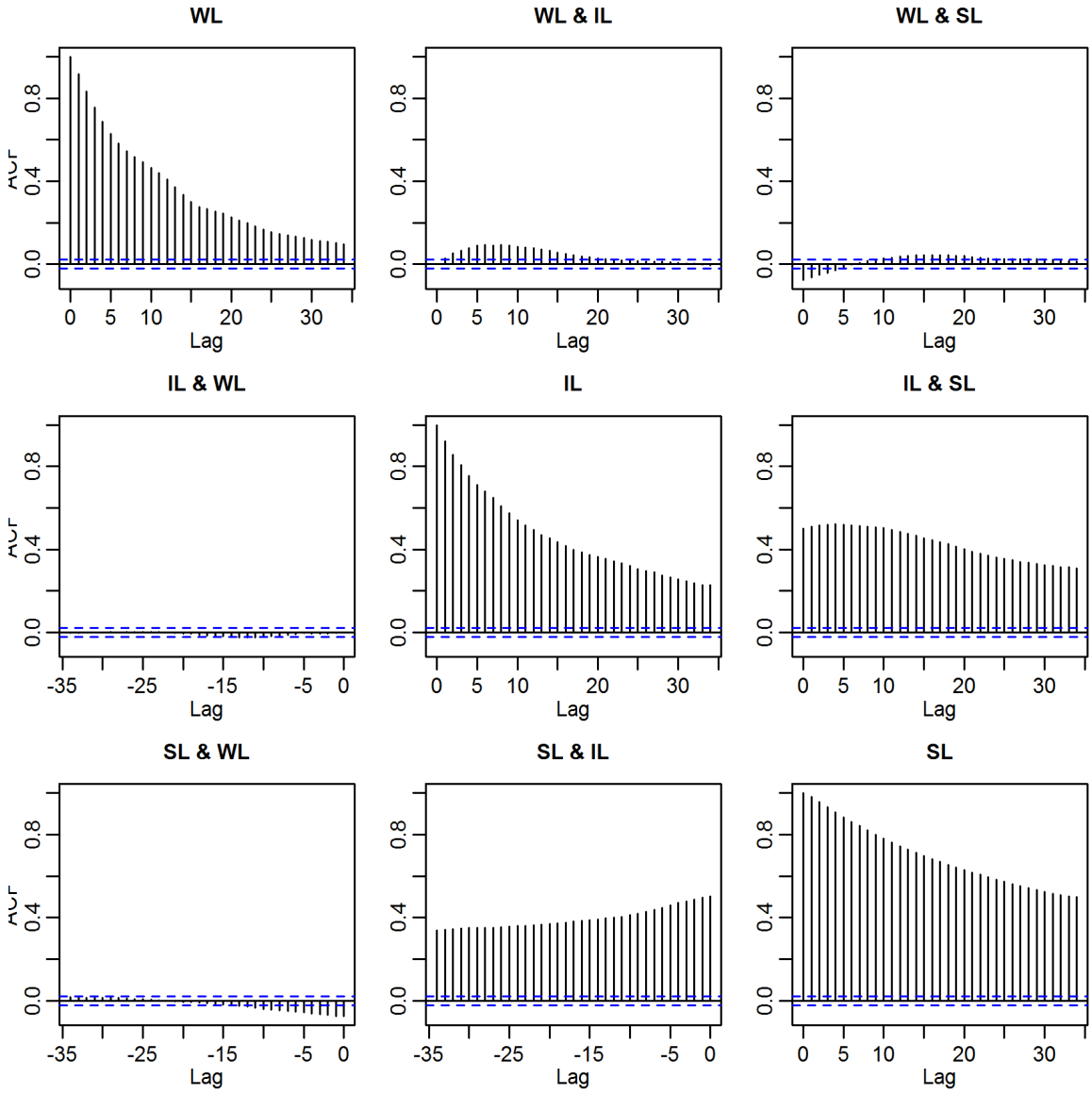


Figure 4.16: Contaminant Layer ACFs and CCFs

4.5.4 Plowing, Salting and Other Factors

To explore how the maintenance operations, like plowing and salting, affect the RSC changing, a snow event from 2007-2-14 to 2007-2-15 is plotted in Figure 4.17. The mean air temperature during the snow event is -2.65°C.

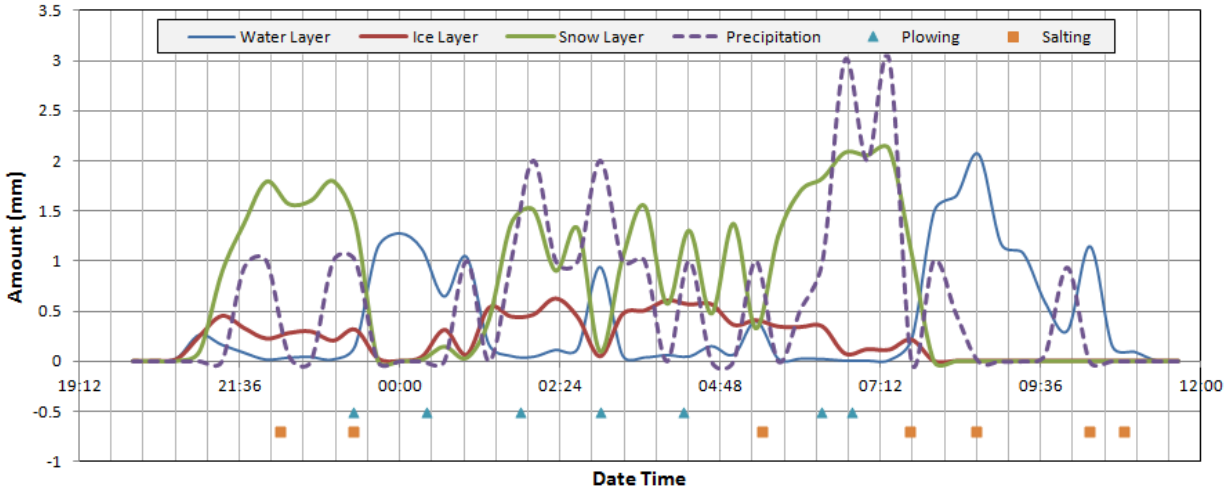


Figure 4.17: Exploratory Analysis of Winter Maintenance Operations

The snow fall is positively related to the snow layer depth. Its correlation with the ice layer depth is much weaker.

Plowing operations are not associated with significant snow depth reductions, which is unexpected. As most plowing operations are conducted during precipitation, their effect could be masked by the snowfall before the next snow depth measurement, especially when the precipitation rate is high.

The salting operations are mostly correlated with high water layer depth, which is consistent with the melting effect of salt, but they are not associated with immediate snow/ice layer depth decrease. Salting 1, 2, 3, 4 are all done when there are snow and/or ice on the pavement. The ice layer decreases significantly 40 minutes after Salting 2, but it can also be related to Salting 1 one hour earlier. The ice layer has a big drop one hour after Salting 3 is done. For Salting 4, this snow/ice decrease happens in half an hour, but it could also be related to Salting 3, which is done two hours earlier. Again, precipitation and other factors like traffic make the situation more complicated and blur the whole picture.

The time with low water depth corresponds to the time with high snow/ice depth; however, examined from another angle, high snow/ice depth is usually followed by high

water depth at a later time. This time offset is about 2-3 hours for this snow event. During the period from 12-14 20:00 to 12-15 00:00, there is a significant amount of both snow and ice, which could cause the large water depth from 12-14 23:00 to 12-15 02:00. Similarly, the source of the water from 12-15 07:00 to 12-15 10:00 is likely to be the snow/ice on the pavement during the period from 12-15 05:00 to 12-15 08:00. This phenomenon implies the salt residual on pavement is likely to take a longer time than one hour to show its melting effect, which violates the one-hour intervaled Markovian assumption of the VTI model. Our model framework in Figure 4.2 will account for this delayed salting effect.

4.5.5 Summary

In this section, a data set containing RSC measures, weather information, maintenance operation records and traffic volumes is systematically explored, and some of the major findings are as follows.

1. Surface temperature is not only auto-correlated, but also correlated with air temperature.
2. Friction level is negatively correlated to snow/ice layer depths, and its relationship to water layer depth is obscure.
3. The depths of three types of pavement contaminant are auto-correlated and cross-correlated. The cross-correlations are mostly consistent with basic physical knowledge.
4. The major source of the snow and ice on the pavement is the snow precipitation.
5. The effect of plowing is not significantly shown at the corresponding time interval, which could be caused by confounding factors.
6. Salting operations are the main factor causing snow/ice melting.
7. It could take longer than an hour for salt to take effect.

4.6 Methodology

After a systematic integration and exploration process, the following RSC measures are determined for further modelling:

1. road surface temperature (ST)
2. road surface friction (FR)
3. water layer depth (WL)
4. ice layer depth (IL)
5. snow layer depth (SL)

ST is auto-correlated and also correlated with air temperature and possible other factors, like traffic and maintenance operations. FR is also auto-correlated and affected by WL, IL and SL. The forecasting model of both ST and FR can be calibrated using the regression style modelling technique in time series analysis, called the integrated autoregressive moving average models. Additionally, some external explanatory variables can enter the model, which would extend the ARIMA model into an ARMAX model (X comes from the word eXternal).

WL, IL and SL are both auto-correlated and cross-correlated, and they are also affected by other factors like air temperature and precipitation. Their forecasting models can be calibrated in a simultaneous manner using the multivariate modelling technique called multivariate ARMAX, or shortly VARMAX (vector ARMAX). The detailed discussions on the theoretical background and application scenarios of ARMA and ARMAX models are given in various time series analysis literature. Shumway & Stoffer (2006) have given them a systematic review and present some good examples on how to use the corresponding packages in R. The rest of this section includes a brief introduction of ARIMA and ARMAX models.

4.6.1 Autoregressive Models

An integrated autoregressive moving average (ARIMA) model is a combination of an autoregressive (AR) model and a moving average (MA) model, which can deal with both stationary and non-stationary time series. An ARIMA model consists of three components: AR, MA and differencing to a certain order.

The AR model is based on the assumption that the current value of the series, x_t, can be explained as a function of p past values, x_{t-1},x_{t-2},. . . ,x_{t-p}, where p determines the

number of steps into the past needed to forecast the current value. This model is called an AR model of order p, abbreviated AR(p). Its formal mathematical expression is

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \omega_t \tag{4.8}$$

where x_t is a stationary time series, ϕ values are constants. ω_t is usually assumed to be a Gaussian white noise series with mean zero, but this assumption is not necessary for most real-world models. If the mean, μ, of x_t is not zero, replace x_t by $x_t - \mu$, i.e.,

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \ldots + \phi_p(x_{t-p} - \mu) + \omega_t \tag{4.9}$$

or

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \omega_t \tag{4.10}$$

where

$$\alpha = \mu(1 - \phi_1 - \ldots - \phi_p) \tag{4.11}$$

Equation 4.9 can be written as

$$(1 - \phi_1 B - -\phi_2 B^2 - \ldots - \phi_p B^p) x_t = \omega_t \tag{4.12}$$

or in a more concise form:

$$\phi(B) x_t = \omega_t \tag{4.13}$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \tag{4.14}$$

B in Equation 4.12 is the backshift operator, defined as

$$B x_t = x_{t-1} \tag{4.15}$$

and can be extended to powers

$$B^2 x_t = B(B x_t) = B x_{t-1} = x_{t-2}, \tag{4.16}$$

thus

$$B^k x_t = x_{t-k} \tag{4.17}$$

4.6.2 Moving Average Models

A moving average (MA) model assumes that x_t can be expressed as a linear combination of the white noise ω_t. The formal mathematical definition of an MA model of order q (MA(q)) is

$$x_t = \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \ldots + \theta_q \omega_{t-q} \tag{4.18}$$

or

$$x_t = \theta(B)\omega_t \tag{4.19}$$

where

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q \tag{4.20}$$

4.6.3 Differencing

When there are non-stationary patterns, especially trends and non-zero mean, in the series x_t, the technique of differencing can be applied to make x_t stationary. The concept of stationary and non-stationary time series has already been introduced in Chapter 3. The first differencing method for high-pass filtering in Chapter 3 can be expressed using B as

$$\Delta x_t = x_t - x_{t-1} = x_t - B x_{t-1} = (1 - B)x_t \tag{4.21}$$

A more general form of the differencing operator can be defined as

$$\Delta^d = (1 - B)^d \tag{4.22}$$

It is the operator of *differences of order d*, where $(1 - B)^d$ can be expanded algebraically to evaluate for higher integer values of d.

4.6.4 Integrated Autoregressive and Moving Average Models

A time series can be assumed as a combination of an AR(p) process and a MA(q) process expressed as

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \ldots + +\theta_q \omega_{t-q} \tag{4.23}$$

or

$$\phi(B) x_t = \theta(B) \omega_t \tag{4.24}$$

which is an ARMA(p,q) model. If a d order detrending is added, the equation turns into

$$\phi(B)(1-B)^d x_t = \theta(B) \omega_t \tag{4.25}$$

which is an ARIMA(p,d,q) model. A further extension of ARIMA models to incorporate periodical behaviors of the series is called the Seasonal Multiplicative ARIMA (SARIMA) model. Shumway & Stoffer (2006) have given detailed discussions about both ARIMA and SARIMA models.

4.6.5 Univariate ARMAX Models

The ARIMA model with all three components (AR, MA, differencing) is a general form, and a model with one or two components is sufficient to explain most real-world time series processes. For example, the surface temperature series (ST) has trends and non-zero mean and is 24-hour periodical, which suggests an ARIMA process with both differencing operations and seasonal components. If the air temperature series (AT) enters the model as an explanatory variable, the model however can be potentially simplified into an AR(1) form, e.g.,

$$ST_t = \alpha ST_{t-1} + \beta AT_t + \omega_t \tag{4.26}$$

This simplification is likely to be successful because the AT series is highly correlated with ST, and it can explain the trends, the mean and most periodical behaviors of ST. The explanatory variables added into an ARIMA model like AT, are usually called exogenous

(or external) variables in the time series analysis field. The model with exogenous variables is call an ARMAX model. A general definition of the ARMAX model is

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \ldots + \theta_q \omega_{t-q} + \Gamma U_t \tag{4.27}$$

where x_t is the element of a univariate time series, U_t is the vector of exogenous variables, and Γ is the constant coefficient vector.

4.6.6 Multivariate ARMAX Models

The univariate ARMAX model can be straightforwardly extended into the multivariate ARMAX (called "VARMAX" in the rest of this thesis) model in the following way:

$$X_t = \Phi_{1t-1} + \ldots + \Phi_p X_{t-p} + \Omega_t + \Theta_1 \Omega_{t-1} + \ldots + \Theta_q \Omega_{t-q} + \Gamma U_t \tag{4.28}$$

X_t is a vector of multiple endogenous variables at time t. The definitions of U_t and Γ are the same as Equation 4.27. Φ, Θ and Ω are the vectorial counterparts of ϕ, θ and ω in Equation 4.27.

Endogenous variables are the dependent variables, which are multiple time series and could be auto-correlated and/or cross-correlated. Usually the dependent variables with bi-directional causal relationships are grouped as a set of endogenous variables and modelled together using Equation 4.28. The bi-directional causal relationship means one endogenous variable can cause other endogenous variables to change, and the change of other endogenous variables can also cause itself to change. Consistent with this idea, water layer depth (WL), ice layer depth (IL) and snow layer depth (SL) are included in the same set of endogenous variables, as water, ice and snow can interchange with each other.

Surface temperature (ST) can cause changes in contaminant layers, but not in reverse direction. Therefore ST is not included in this endogenous variable set, and it is modelled separately using univariate ARMAX. Friction level (FR) is modelled separately as well because its changing is unidirectionally caused by contaminant layers.

4.7 Model Calibration

As discussed in Section 4.6.6, *ST* and *FR* are both modelled using the univariate ARMAX method. *WL*, *IL* and *SL* are treated as an endogenous variable set and modelled simultaneously.

The values of all variables are linearly scaled to the range [0,1] as explained in Section 3.7. The minima and maxima of all relevant variables are listed in Table 4.7.

Table 4.7: Ranges of Modelling Variables

Name	Meaning	Unit	Minimum	Maximum
ST	surface temperature	°C	-19.4	17.7
Plow	plowing	times	0	1
AT	air temperature	°C	-25.0	11.0
Precip	precipitation	mm	0.0	13.0
Traffic	hourly traffic	vehs/lane/hour	26	565
Salting_LN	salt application rate at time lag *N*	kg/km	0	260.0
WL	water layer depth	mm	0.0	6.5
IL	ice layer depth	mm	0.0	1.3
SL	snow layer depth	mm	0.0	5.3
FR	friction level	NA	0.09	0.82

4.7.1 Surface Temperature Model

ST_t can be predicted using a linear function of $ST_{t-1}, ST_{t-2}, \ldots, ST_{t-p}$ and the following exogenous variables :

- AT_t: average air temperature during $t-1$ to t
- $Precip_t$: average precipitation during $t-1$ to t
- $Traffic_t$: hourly traffic rate during $t-1$ to t
- $Plow_t$: plowing operation during $t-1$ to t
- $Salting\_L0 \sim Salting\_L9$: salting rates during $t-1$ to t ($L0$), $t-2$ to $t-1$ ($L1$), $\ldots$, $t-10$ to $t-9$ ($L9$). If there are no salting operation within any time interval, the value of zero is used. These ten variables represent salting rates during the last three hours of time t.
- WL_{t-1}: water layer depth at $t-1$
- IL_{t-1}: ice layer depth at $t-1$
- SL_{t-1}: snow layer depth at $t-1$

The exogenous variables for model calibration are chosen according to the model framework in Section 4.2. The most important factors affecting RSC from weather, traffic and maintenance operations are included in the model.

The model takes the form of ARMAX(p, d, q) with both $d = 0$ and $q = 0$, i.e., no differencing operation is conducted and no MA component is considered. Although the surface temperature series has non-stationary and cyclical patterns, these patterns can be explained by air temperature as suggested in Section 4.6.5; thus no differencing or seasonal component is explicitly added to the model. Including the MA component will increase difficulties for model interpretation and application. When the AR component and the exogenous variables can sufficiently explain the response variable, the MA component should be left out.

The data from 2007-11-20 00:00:00 to 2008-02-09 10:00:00 (81 days) are used for model calibration. The rest of the dataset from 2008-02-09 10:00:00 to 2008-03-31 23:42:00 (51 days) are held out for cross-validation.

The procedure to determine the optimal p can be found in most time series books. A general rule-of-thumb is to study the pattern of ACF and PACF (partial auto-correlation function) graphs and try the suggested p values. Detailed procedures can be found in Shumway & Stoffer (2006).

The ACF and PACF graphs are respectively shown in Figure 4.18 and Figure 4.19 up to 84 lags. The ACFs at all lags are significant with a periodical pattern. In the PACF graph, except the first 8 lags, all other lags are insignificant or marginally significant and there are some significant lags around 72 lags corresponding to the daily cycle. The patterns of ACF and PACF graphs suggest an ARMAX$(8, 0, 0)$ model, i.e., $p = 8$.

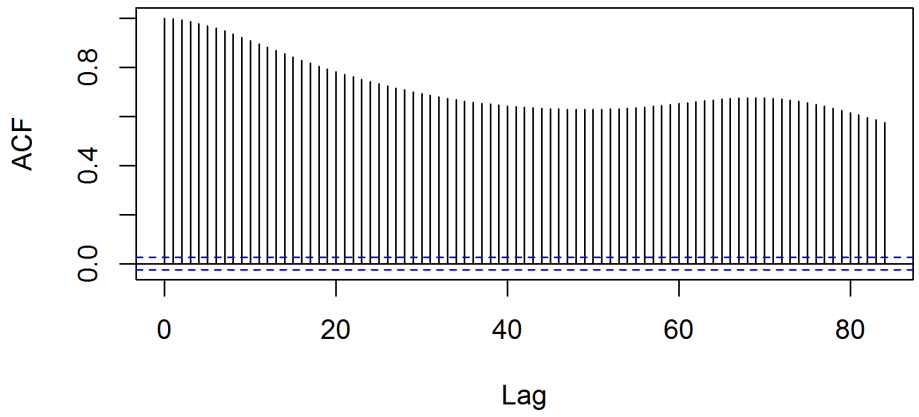


Figure 4.18: ACF of Surface Temperature

An ARMAX$(8, 0, 0)$ model is calibrated with the R package "forecast" with the maximum likelihood approach. Hyndman & Khandakar (2008) have given a detailed discussion about this package. The coefficients, AIC, BIC and log-likelihood of the model are shown in Table 4.8. The numbers in parentheses are the standard deviations of respective estimated coefficients.

The ACF and PACF plots of model residuals are shown in Figure 4.20 and Figure 4.21. Except for some marginally significant values at lags around the daily period, almost all ACFs are insignificant. A similar pattern is found in the corresponding PACF plot. It suggests that $p = 8$ is a highly acceptable trade-off point of model quality and model complexity. For the coefficients of the explanatory variables, the following can be found.

- The AR terms at lag 1, 2, 3 and 8 are significant based on the t-values of their coefficients . Including AR terms at lag numbers greater than 1 in the model actually breaks the strict Markovian assumption of the model framework. The autocorrelation and periodical patterns of ST are due to some factors not included in the

Table 4.8: ARMAX Model for Surface Temperature

Intercept	AR1	AR2	AR3	AR4
0.2327	1.1389	-0.1210	0.0549	-0.0255
(0.0138)	(0.0133)	(0.0198)	(0.0199)	(0.0199)
AR5	AR6	AR7	AR8	
-0.0274	0.0301	0.0152	-0.0756	
(0.0199)	(0.0199)	(0.0198)	(0.0131)	
WL_{t-1}	IL_{t-1}	SL_{t-1}		
-0.0089	-0.0081	-0.0112		
(0.0038)	(0.0035)	(0.0036)		
AT_t	$Precip_t$	$Traffic_t$	$Plow_t$	
0.2944	-0.0016	0.0083	0.0000	
(0.0130)	0.0026	(0.0035)	0.0001	
Salting_L0	Salting_L1	Salting_L2	Salting_L3	Salting_L4
0.0012	0.0005	-0.0020	-0.0067	-0.0046
(0.0013)	(0.0019)	(0.0023)	(0.0025)	(0.0026)
Salting_L5	Salting_L6	Salting_L7	Salting_L8	Salting_L9
-0.0016	-0.0009	0.0006	-0.0009	-0.0004
(0.0026)	(0.0025)	(0.0023)	(0.0019)	(0.0013)
AIC		-38048.26		
BIC		-37868.00		
Log Likelihood		19051.13		

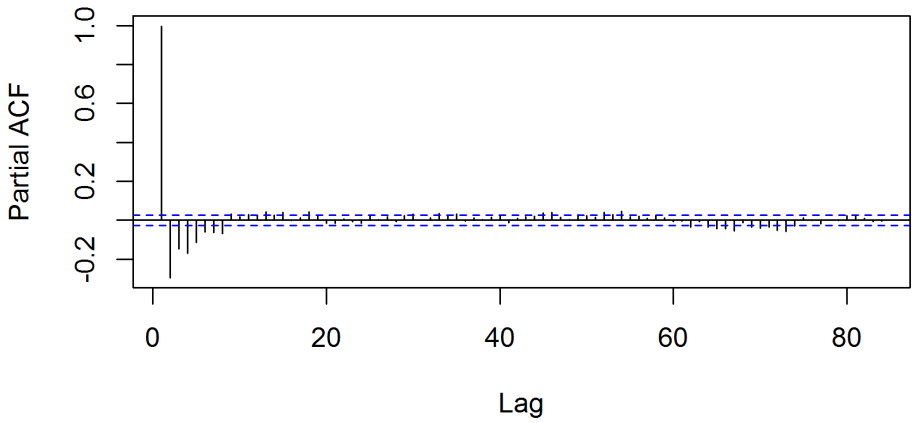


Figure 4.19: PACF of Surface Temperature

model, such as the energy influx balance of the pavement material and the daily solar radiation periodical variation. It is reasonable to loosen the Markovian assumption to add more AR terms to indirectly address those factors. Although the AR terms at lag 4, 5, 6, 7 are insignificant, when the term at lag 8 is significant, those terms are usually kept in the model to make the model more stable and the forecast more reliable.

- *AT* is significant and the magnitude of its coefficient is the second only to the AR term at lag 1. It is consistent with the observation that the air temperature and surface temperature are highly correlated, and including such an exogenous explanatory variable can significantly improve the forecasting performance of the model.

- *Traffic* is significant and its coefficient is positive, which suggests more traffic is correlated with higher surface temperature. This phenomenon can be explained as vehicles heating the road surface by emission or by wheels compacting the snow. But it can also be interpreted as a noncausal correlation, e.g., there is much more traffic during the day than at night, and on average, the daytime temperature is higher than night time.

- The coefficients of *WL*, *IL*, *SL* are all negative, suggesting that pavement contaminant is correlated with lower surface temperature. Although *Precip* is not significant, its

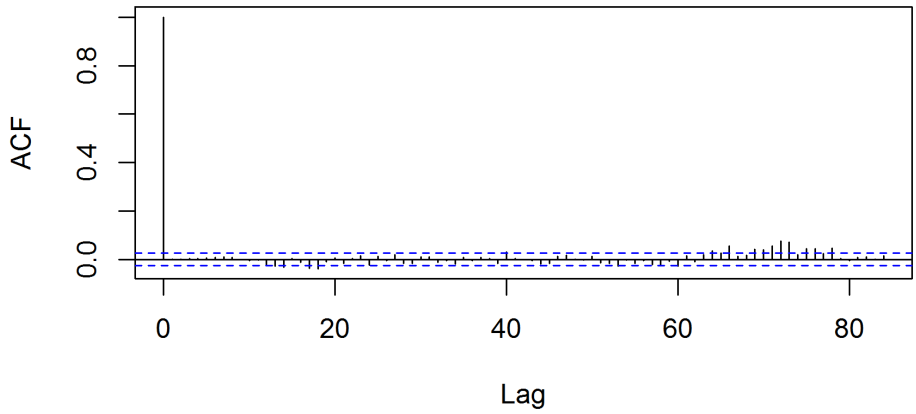


Figure 4.20: ACF of Surface Temperature Model Residuals

effect could be mediated by pavement contaminant terms.

- Most maintenance operation terms are insignificant and of very small magnitude, suggesting that plowing or salting operations do not significantly influence road surface temperature.

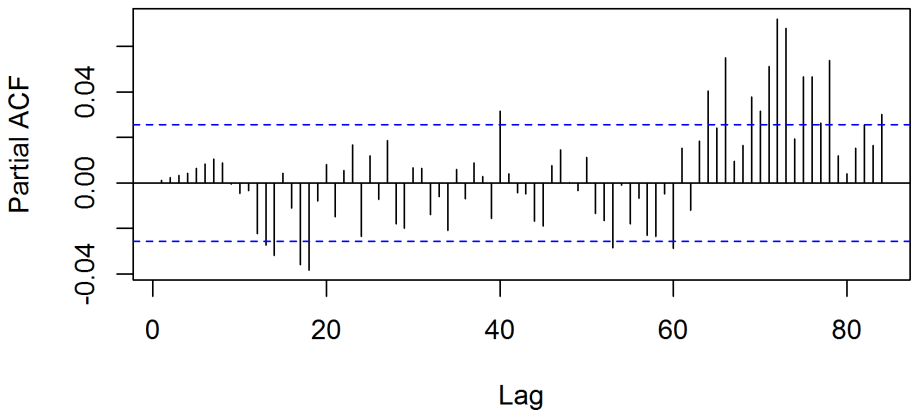


Figure 4.21: PACF of Surface Temperature Model Residuals

The QQPlot and the histogram of the residual of one-step ahead forecasting are shown in Figure 4.22 and 4.23. QQPlot is a common diagnostic graph to examine the normality of the residuals of linear regression models. If the residuals do not deviate far from the normal distribution, the points in their QQPlot should mostly stay on a straight line. Other patterns suggest non-normal distributions.

The pattern of the QQPlot of the residuals suggests a non-normal distribution, which has heavier tails. The distribution in the histogram is centered around zero with roughly symmetrical tails and the overall pattern is bell-shaped. This pattern of the residuals is acceptable for a linear regression model.

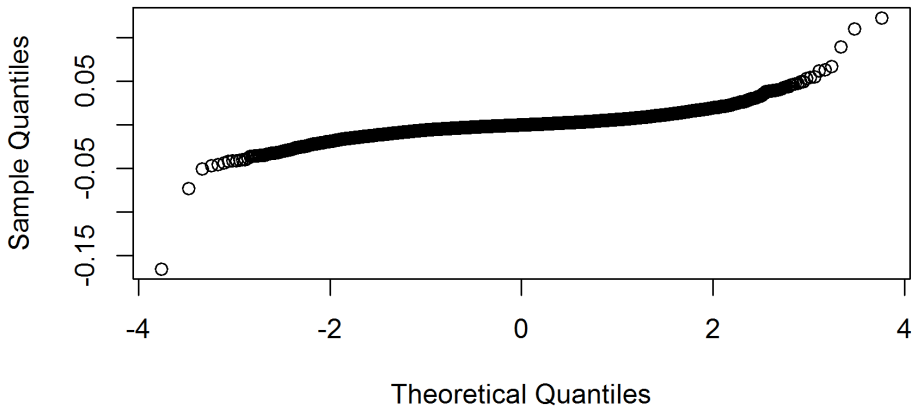


Figure 4.22: QQPlot of Surface Temperature Residuals – Calibration Data

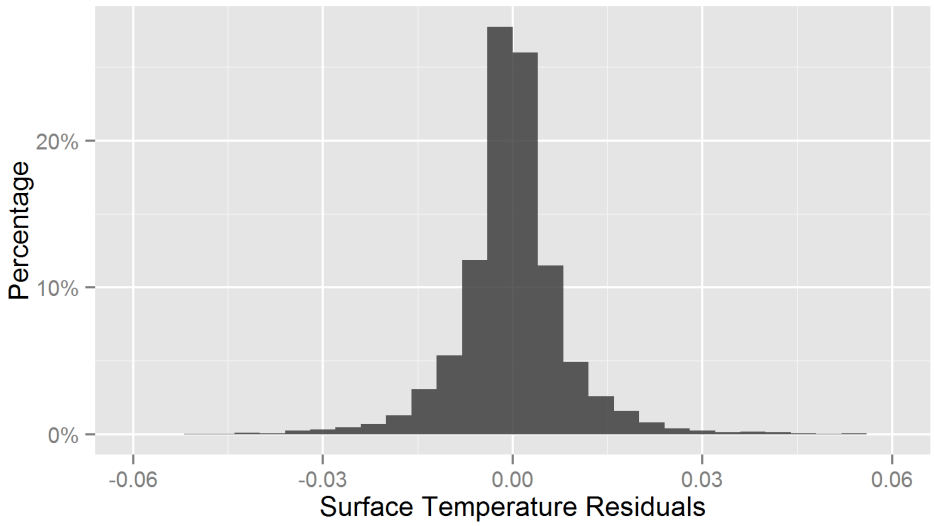


Figure 4.23: Histogram of Surface Temperature Residuals – Calibration Data

The calibrated model suggests that weather, traffic and pavement contaminant have some effects on road surface temperature changing, although some effects are hard to explain in a causal way. The validation results of this model are discussed in Section 4.8.1.

4.7.2 Friction Level Model

The friction level reported by the Vaisala sensor set is deterministically estimated using some empirical model by Vaisala. The purpose of this study is not to reversely engineer this model; therefore, in this study, FR is predicted using a linear function of $FR_{t-1}, FR_{t-2}, \ldots, FR_{t-p}$ and the same exogenous variables in the surface temperature model. As FR is not measured friction using friction meters, the calibration and validation results are for the purpose of testing and illustrating how the model framework can be applied to similar situations, especially when real friction measurement is available.

Similar to the ST model, the ARMAX$(p, 0, 0)$ model format is adopted with no differencing and MA components. The same sets of calibration and validation data for modelling ST are used for the friction level model.

The ACF and PACF graphs are respectively shown in Figure 4.24 and Figure 4.25. Both graphs show no periodical patterns. The ACF exponentially decreases and the PACF cuts off after lag 3, which suggests a typical autoregressive process. The models with $p = 1$ to 8 are all calibrated using the R package "forecast", and their AIC and BIC diagnostics are listed in Table 4.9.

Table 4.9: Comparison of Friction Model Diagnostics

p	AIC	BIC
1	-13640.9	-13525.21
2	-14385.17	-14263.7
3	-14610.48	-14483.22
4	-14697.79	-14564.74
5	-14735.09	-14596.26
6	-14744.67	-14600.06
7	-14743.34	-14592.95
8	-14744.32	-14588.14

As introduced in Chapter 3, lower AIC and BIC values suggest a better model fit. According to Table 4.9, ARMAX$(6, 0, 0)$ has the lowest AIC and BIC. The calibration results of this ARMAX$(6, 0, 0)$ model are summarized in Table 4.10, and the summaries of the models with other p values can be found in Appendix C.

Table 4.10: ARMAX Model for Friction Level

Intercept	AR1	AR2	AR3	AR4
0.3239	1.3148	-0.1336	-0.0626	-0.0280
(0.0185)	(0.0204)	(0.0339)	(0.0340)	(0.0338)
AR5	AR6			
-0.0358	-0.0697			
(0.0338)	(0.0205)			
WL_{t-1}	IL_{t-1}	SL_{t-1}		
0.0058	-0.0085	0.0023		
(0.0077)	(0.0061)	(0.0080)		
AT_t	$Precip_t$	$Traffic_t$	$Plow_t$	
0.1994	-0.0245	0.0146	0.0000	
(0.0184)	0.0060	(0.0086)	0.0012	
Salting_L0	Salting_L1	Salting_L2	Salting_L3	Salting_L4
-0.0027	-0.0022	-0.0048	-0.0095	-0.0074
(0.0026)	(0.0039)	(0.0049)	(0.0055)	(0.0058)
Salting_L5	Salting_L6	Salting_L7	Salting_L8	Salting_L9
-0.0047	-0.0025	-0.0019	-0.0014	-0.0021
(0.0058)	(0.0055)	(0.0049)	(0.0039)	(0.0026)
AIC		-14744.67		
BIC		-14600.06		
Log Likelihood		7397.34		

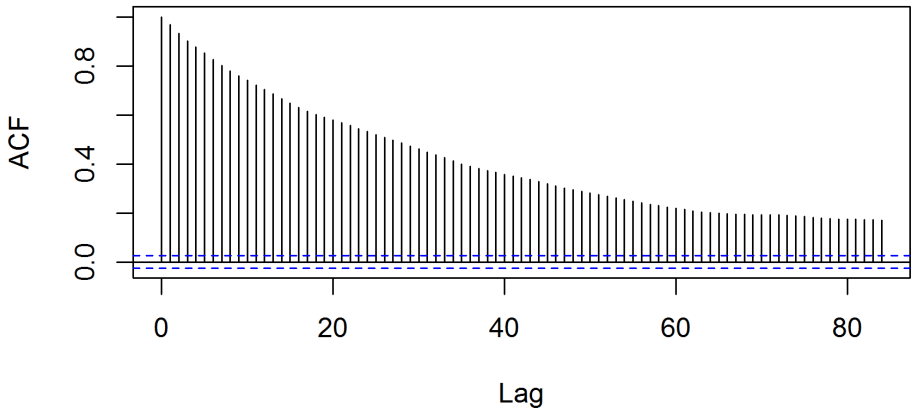


Figure 4.24: ACF of Friction Level

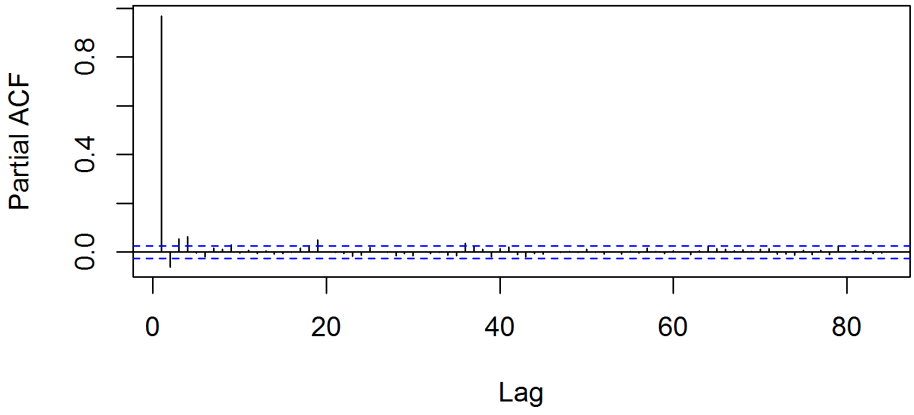


Figure 4.25: PACF of Friction Level

The ACF and PACF plots of the residuals are shown in Figure 4.26 and Figure 4.27. Except for a few marginally significant lags, most of auto-correlations have been captured by the model.

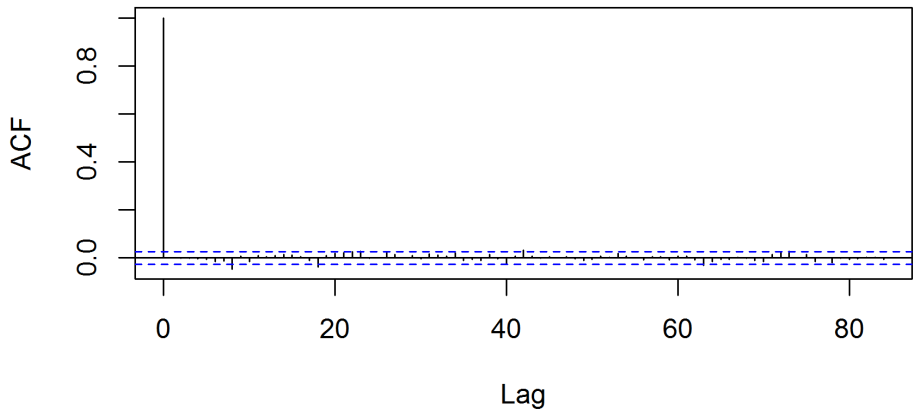


Figure 4.26: ACF of Friction Level Residuals – Calibration Data

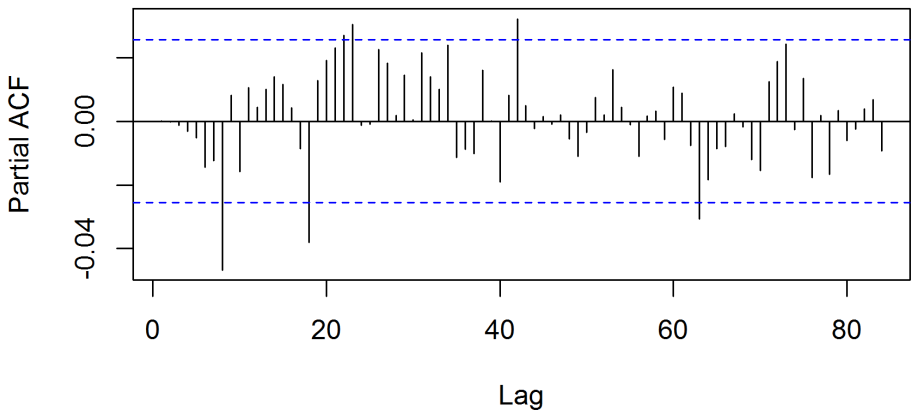


Figure 4.27: PACF of Friction Level Residuals – Calibration Data

As for the model coefficients, the following can be found.

- The AR terms at lag 1, 2 and 6 are significant. Including AR terms at lag 2 and 6 in the model breaks the strict Markovian assumption. Similar to the *ST* model, including those AR terms is to use them as the surrogate of some factors which cannot be included in the model. It should not be considered a serious challenge to the basic Markovian assumption.
- *AT* is significant and the magnitude of its coefficient is the second only to the AR term at lag 1, suggesting higher air temperature corresponds to higher friction level.
- *Precip* is significant and its coefficient is negative, which suggests more precipitation is correlated with lower friction levels. As most precipitation during the study period is snow, it can be safely explained as snowfalls causing a slippery road surface.
- *WL*, *IL*, *SL* at lag 1 are all insignificant. As *FR* at lag 1 has already entered the model, it could have masked the effect of *WL*, *IL*, *SL* at the same lag.
- *Traffic*, *Plow* and all salting operations are insignificant to predict friction, which is hard to explain. As the modelled friction levels are not real friction measurements, the resulting model could have difficulties for interpretation under certain circumstances.

The QQPlot and the histogram of the residual of one-step ahead forecasting is shown in Figure 4.28 and 4.29. The pattern of the QQPlot suggests a skewed non-normal distribution. The distribution in the histogram is roughly centered around zero with a bell shape, which is acceptable.

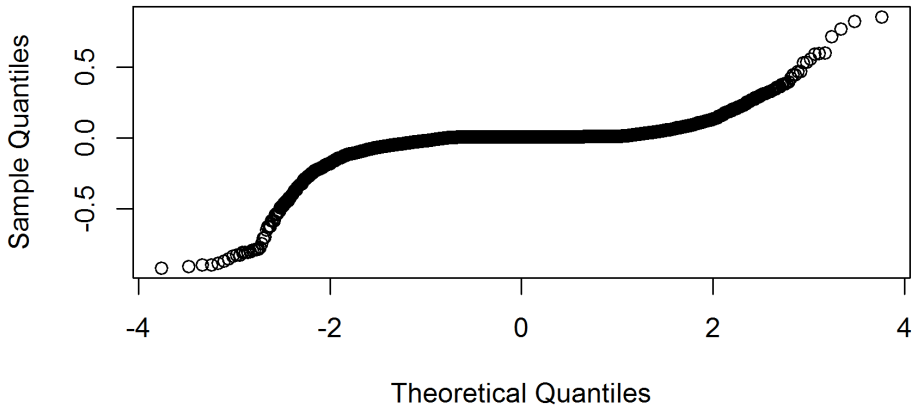


Figure 4.28: QQPlot of Friction Level Residuals – Calibration Data

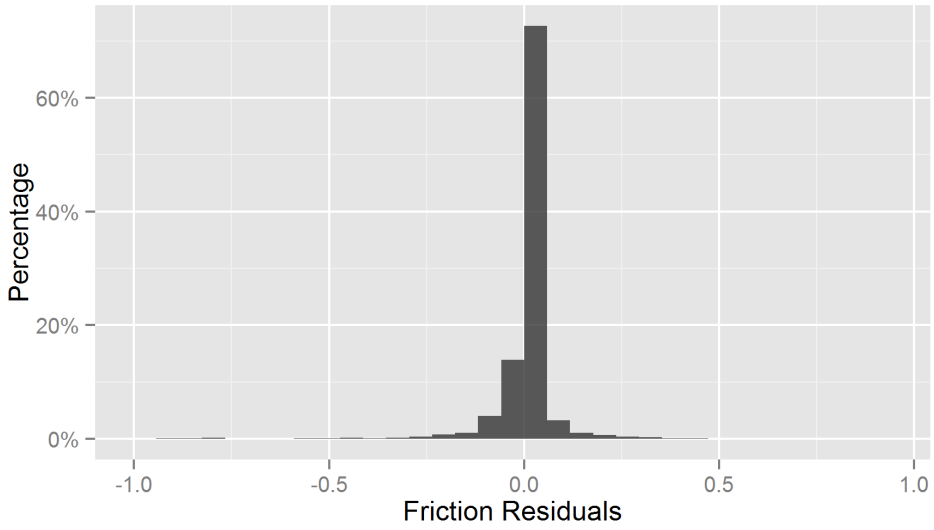


Figure 4.29: Histogram of Friction Level Residuals – Calibration Data

The calibrated model suggests that weather factors have direct effects on friction level changes. The effects of pavement contaminants could be masked by the effects of previous frictions in the model. Traffic and maintenance operations do not show clear effects. The validation results of this model are discussed in Section 4.8.2.

4.7.3 Contaminant Layer Model

According to the model framework in Section 4.2, the depths of three contaminant layers – WL, IL and SL – at time t can be predicted using their own previous values and the following exogenous variables:

- AT_t: average air temperature during $t-1$ to t
- $Precip_t$: average precipitation during $t-1$ to t
- $Traffic_t$: hourly traffic rate during $t-1$ to t
- $Plow_t$: plowing operation during $t-1$ to t
- $Salting\_L0 \sim Salting\_L9$: salting rates during $t-1$ to t ($L0$), $t-2$ to $t-1$ ($L1$), $\ldots$, $t-10$ to $t-9$ ($L9$).

The model takes the form of VARMAX(p, d, q) where $d = 0$ and $q = 0$. In both calibration and validation datasets, there are a lot of samples with no contaminant on the pavement and no precipitation. These samples usually correspond to the temporal gaps between snow events. If these samples are all included in model calibration and validation, the large proportion of zero values in the endogenous variables will introduce significant bias. So most of the samples with zero pavement contaminant and zero precipitation during those snow event gaps are cut out. Only a small proportion of such samples adjacent to the snow events are kept as safe buffers for lagging calculations in both model calibration and validation. The original dataset as a complete long time series is cut into a number of shorter time series in this way. Those shorter time series are again stitched together into a single long time series. With the low proportion of zero values in the safe buffers in this stitched time series, the VARMAX model is supposed to give comparatively unbiased results.

The same cut-and-stitch processing should also be applied to the data for the previous friction level modelling. As the friction data collected for our study are not real friction measurements, we add friction level modelling into this thesis study only for the purpose of testing and illustration as indicated in Section 4.7.2; therefore, this step of data processing is skipped for the friction level modelling. For the surface temperature modelling, this step of data processing should be avoided, as the resulting abrupt temperature (both surface and air temperatures) changes at the stitching points are actually invalid. Including those invalid temperatures into the dataset to directly model temperature changing process may introduce unexpected and unexplainable bias and unreliability into the resulting models.

The sample sizes of the original and the stitched time series are listed in Table 4.11. It can be seen that more than half of the samples in the original dataset have been removed.

Table 4.11: Sample Sizes of Original and Stitched Datasets

	Original	Stitched
Calibration	5863	2891
Validation	2403	819

For univariate ARMAX models, the scope of p value can be conveniently determined from ACF and PACF plots. For VARMAX models, this approach does not work well, especially when the endogenous variables are highly inter-correlated. Based on the model framework, the model is ideally strictly Markovian so that it is easy to explain. Therefore, a VARMAX$(1, 0, 0)$ is calibrated first, of which the detailed R output is listed in Appendix D and summarized in Table 4.12. The model is calibrated with the "VAR" function in the R package "vars" using the maximum likelihood approach.

Table 4.12: Contaminant Layer Depth Model

	WL Model	IL Model	SL Model
$(Const)$	-0.0088^{*}	0.0121^{**}	0.0047^{*}
	(0.0035)	(0.0038)	(0.0019)
WL_{t-1}	0.8903^{***}		
	(0.0084)		
IL_{t-1}	0.0292^{***}	0.8760^{***}	
	(0.0082)	(0.0088)	
SL_{t-1}	-0.0098^{**}	0.0277^{***}	0.9797^{***}
	(0.0038)	(0.0040)	(0.0036)
AT_t	0.0174^{**}	-0.0122^{*}	
	(0.0056)	(0.0059)	
$Precip_t$	0.0420^{***}		0.0303^{***}
	(0.0082)		(0.0086)
$Traffic_t$			-0.0073^{*}
			(0.0033)
$Plow_t$			
$Salting\_L1$		0.0176^{**}	
		(0.0066)	
$Salting\_L2$	0.0148^{*}		
	(0.0062)		
$Salting\_L3$	0.0159^{**}		
	(0.0062)		
$Salting\_L5$	0.0129^{*}		
	(0.0062)		
$Salting\_L6$		-0.0139^{*}	
		(0.0066)	
$Salting\_L7$		-0.0157^{*}	
		(0.0066)	
Multiple R-Squared	0.8597	0.8728	0.9719
Adjusted R-squared	0.8593	0.8725	0.9718
F-statistic	1961	2827	2.491e+04
p-value	$<$ 2.2e-16	$<$ 2.2e-16	2.2e-16

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

The ACF and CCF plots of model residuals are shown in Figure 4.30. Only the ACFs of *WL* residuals at lag 1 and *SL* residuals are significant, while other ACFs and CCFs are successfully captured by the model.

The VARMAX model consists of three submodels corresponding to *WL*, *IL* and *SL*. The adjusted R-square values of three submodels are respectively 0.8593, 0.8725 and 0.9718, which are quite high.

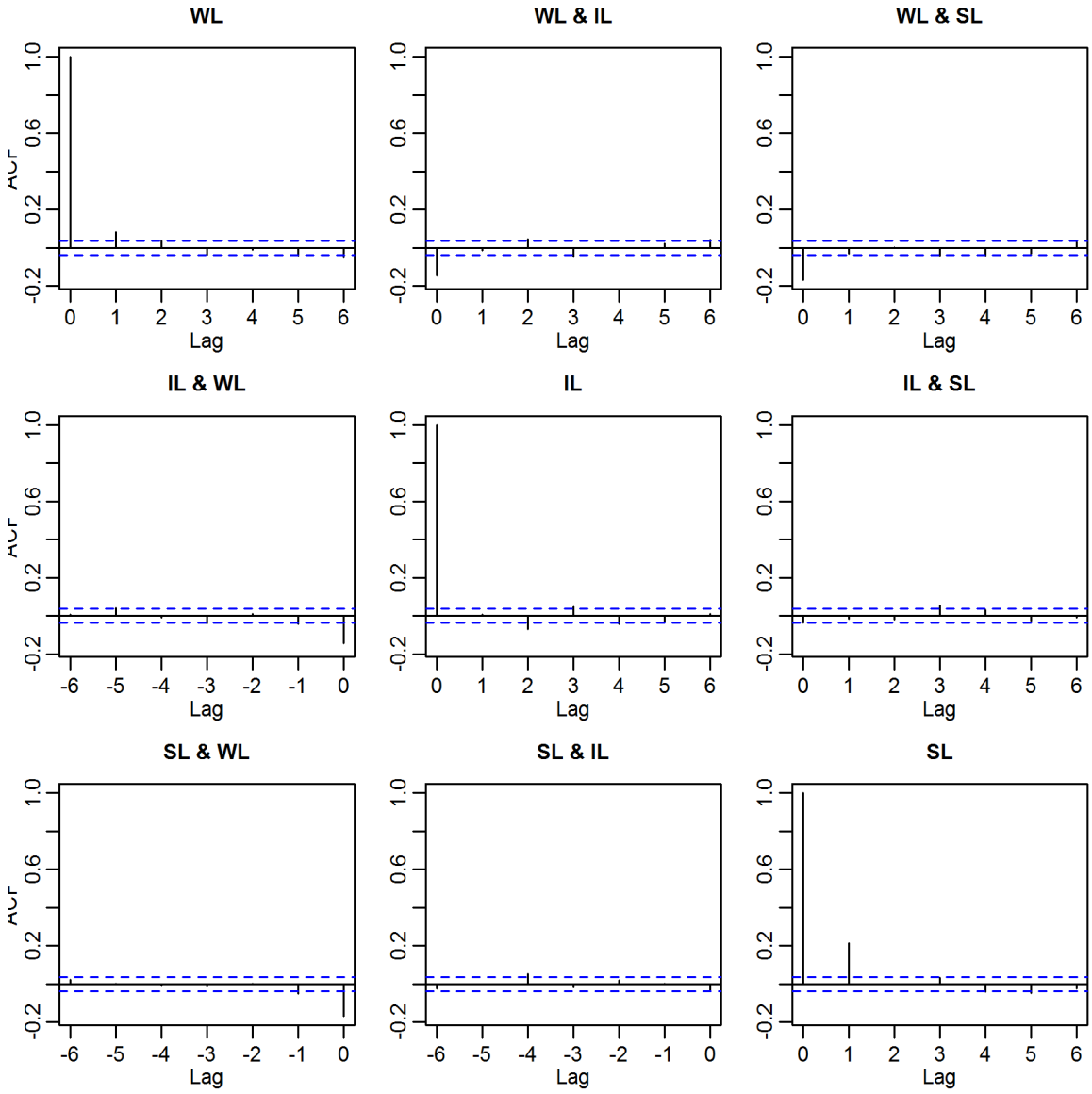


Figure 4.30: ACFs and CCFs of Contaminant Layer Residuals – Calibration Data

Water Layer Submodel

In the *WL* submodel, *WL* is auto-correlated positively as expected. It is also positively correlated with *IL* at time $t-1$ implying the process of ice melting. Unexpectedly, *WL* is negatively correlated with *SL* at last time interval, but the magnitude of the coefficient is very small.

AT is positively significant implying higher air temperature can melt more ice into water. *Precip* is also positively significant implying more precipitation can generate more water, which is also reasonable.

Salting at lag 2, 3 and 5 are all positively significant, implying more salt applied during the last 2 hours especially at lag 2 to 5 could generate more melted water. Salt applied at the current time interval and lag 1 is insignificant, consistent with the assumption that salt takes time to show its melting effect.

Ice Layer Submodel

In the *IL* submodel, *IL* is auto-correlated positively as expected. It is also positively correlated with *SL* of last time interval implying the snow on the pavement could be one source of the ice. *IL* is not correlated with *WL* of last time interval suggesting the refreezing process from water to ice rarely happens at the observation site and period.

AT is negatively significant implying a higher air temperature can melt more ice and decrease the depth of the ice layer, which is consistent with the findings in the *WL* submodel. *Precip* is insignificant suggesting that the precipitation does not directly affect the ice layer depth. The *SL* submodel shows that precipitation can increase the snow layer depth, while this *IL* submodel suggests that snow can be converted to ice. Therefore, it is likely that the precipitation increases the ice layer depth in an indirect manner.

Salt applied at lag 6 and 7 are negatively significant, implying more salt applied 2 hours to 3 hours previously, especially at lag 6 to 7 could decrease more ice. Salt applied at lag 1 is positively significant, which is hard to explain considering the time delay of the melting process. The aggregate salting effect of these three significant lags is however obviously negative, which again confirms salting operations can decrease the ice layer depth.

Snow Layer Submodel

In the *SL* submodel, *SL* is auto-correlated positively as expected, but it is not correlated with either *WL* or *IL* at the last time interval, suggesting that the transformation from

water or ice to snow rarely occurs on a roadway surface, which makes sense from a physical point of view..

Precip is positively significant as expected because snowfalls increase snow layer depth. *Traffic* is negatively significant, which can be interpreted as traffic compacts snow into ice or splashes the snow off the lane. This interpretation is just tentative and by no means conclusive.

The salt applied at previous time intervals does not show a significant effect on snow layer depth, which is unexpected and hard to explain. It might be an honest reflection of the truth that salt does not melt snow efficiently.

The QQPlots in Figure 4.31 and the histograms in Figure 4.32 of the residuals of three submodels show obvious non-normal patterns, but the residuals are concentrated around zero and the distribution patterns are all roughly bell-shaped.

This model is also cross-validated using the validation dataset and the results are discussed in Section 4.8.3.

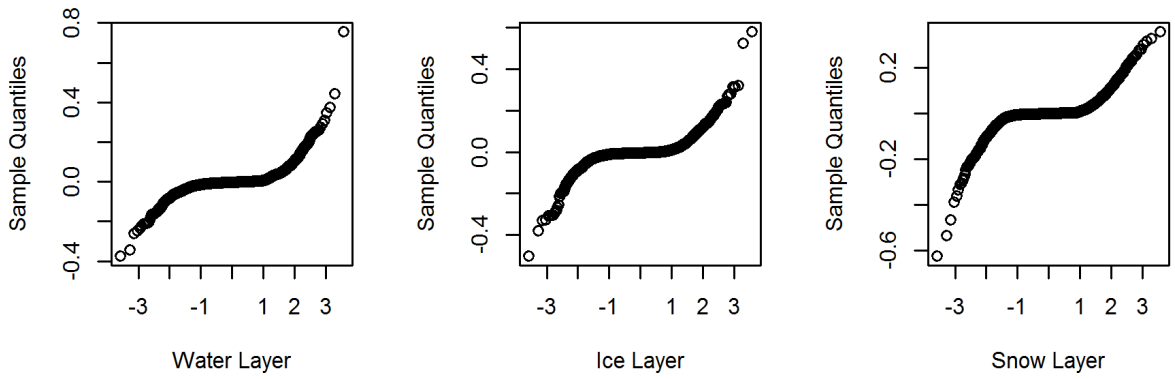


Figure 4.31: QQPlots of Contaminant Layer Residuals – Calibration Data

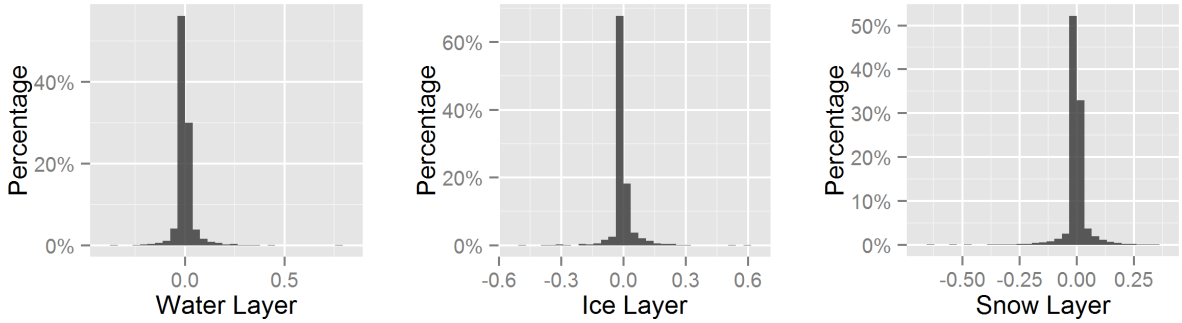


Figure 4.32: Histograms of Contaminant Layer Residuals – Calibration Data

Although the forecast made using the calibration data looks good, increasing p from 1 to a higher order like 2 could potentially improve the forecasting quality, as more than 1 AR term entering the model may catch the effect of some factors not included in the model. According to this idea, a VARMAX$(2, 0, 0)$ model is calibrated with the same set of exogenous variables as the VARMAX$(1, 0, 0)$ model. The R output of the VARMAX$(2, 0, 0)$ model is in Appendix E.

Some AR2 terms are significant, but we have not attempted to explain further as those terms are added to model some unexplainable hidden effects. The signs and significance of all exogenous variables are almost the same as the VARMAX$(1, 0, 0)$ model. The R-squared values of all three submodels are either not improved or just improved marginally. All these facts imply that the VARMAX$(2, 0, 0)$ model is very similar to the VARMAX$(1, 0, 0)$ model in terms of model structure and forecasting reliability.

Salt Residual Estimation

Salting operations appear as a very important factor in the contaminant layer model, but the calibrated VARMAX$(1,0,0)$ and VARMAX$(2,0,0)$ models do not directly quantify the effect of the salt residual. The only implication is that the amount of salt applied at a specific time lag has a fixed effect on current contaminant layer depths. This is an oversimplified assumption but could be the best in our case, as there is no direct or indirect measurement of salt residual. It is also applicable that the salt residual is estimated first and then enters the model in place of the applied salt. In this way, the effect of the salt residual can be straightforwardly quantified.

As no salt residual measurement is available in the dataset, a very simple decreasing function is assumed, which is

$$SaltRes_t = SaltApp_t + \frac{SaltRes_{t-1}}{e} \tag{4.29}$$

$SaltRes_t$ is the salt residual at time t. $SaltRes_{t-1}$ is the salt residual at time $t-1$. $SaltApp_t$ is the salt applied during the period from $t-1$ to t. It is assumed that the salt residual decreases exponentially with time but is not affected by other factors, like weather or traffic. As the salt residual data are not available, we have to stick with this oversimplified assumption.

A VARMAX$(1,0,0)$ model with salt residual as an explanatory variable is calibrated and the R output is listed in Appendix F. By comparing the coefficients of this model with the previous VARMAX$(1,0,0)$ model, it can be found that

- the signs and significance of all endogenous variables and exogenous variables in submodels are exactly the same;
- in the water layer submodels, the significant salting operations are at lag (2, 3, 5) with the signs (+, +, +), while the significant salt residuals are at lags (2, 4, 5) with the signs (+, -, +). The aggregate effect of the salt residuals is positive. This suggests that these two submodels are very similar in terms of model structure;
- in the ice layer submodels, the significant salting operations are at lags (1, 6, 7) with the signs (+, -, -), while the significant salt residuals are at lag (1, 6) with the signs (+, -). The signs of the aggregate effects of these two models are different. Overall these two submodels are very similar;

- in both snow layer submodels, the signs and significance of all endogenous variables and exogenous variables are the same.

The log-likelihood and R-squared values of these two models show almost no difference. As the decreasing function of the salt residual is very simple, the estimated salt residual is highly correlated to the applied salt. Therefore, replacing applied salt with salt residual in model calibration is expected to generate very similar results in terms of model structure.

Although estimating salt residual is not in the scope of this study due to the limitation of the data, it is potentially an important component of a more completed RSC forecasting model. It is suggested that such an estimation model should be studied and calibrated using the similar ARMAX analysis technique.

Interaction Effects

All ARMAX or VARMAX models in preceding sections only consider the main effects, and the interaction effects between explanatory variables are ignored due to reasons specified in Section 3.5.

With this proviso, a VARMAX model only considering the interaction effect of air temperature and salt residual is tentatively calibrated for contaminant layers and the details are in Appendix G. It can be found that

- the overall model quality is not improved by considering this interaction effect as suggested by the model log-likelihood and R-Squared values;
- the signs and significance of all weather and traffic main effects are all the same, implying a very similar model structure;
- the signs of some salting residual variables and interaction terms are difficult to explain. For example, the coefficient of the term $TempRes_L1$ in the snow layer submodel is positive, suggesting that higher air temperature with more salt residual at time lag 1 is correlated with more snow depth on the pavement at current time.

Although it is out of the scope of this study, a systematic exploratory analysis and modelling attempt should be made when the quality of the dataset allows.

4.8 Model Validation

4.8.1 Surface Temperature Model

Using the model calibrated in Section 4.7.1, the one-step ahead forecasts are made for both calibration and validation datasets and the observed and forecast values are plotted together in Figure 4.33.

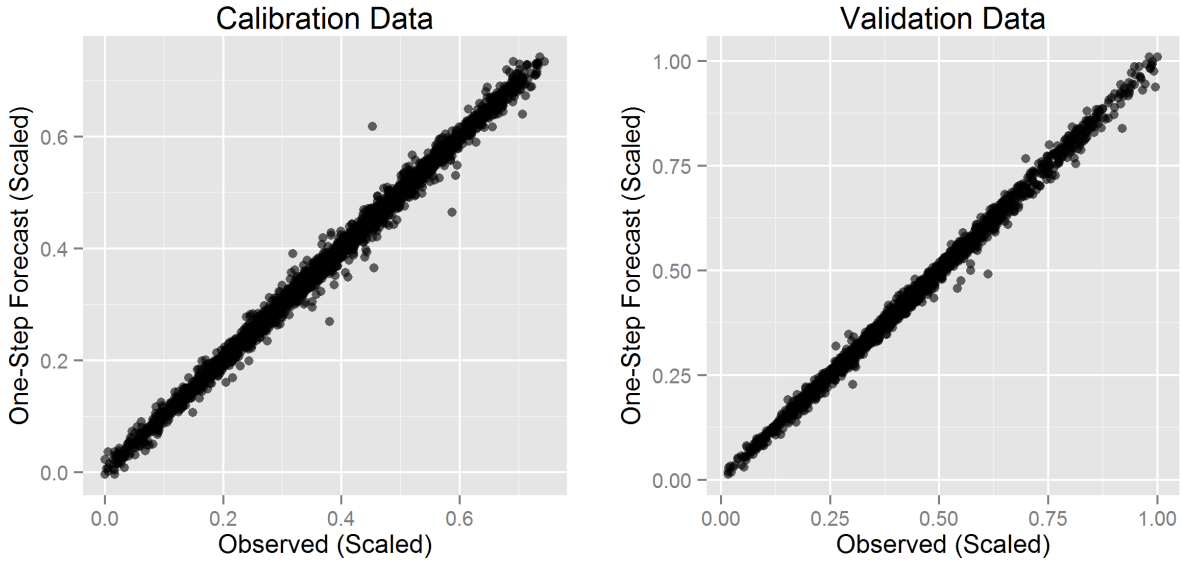


Figure 4.33: Surface Temperature Dotplots: Observed and One-Step Forecast

For both datasets, the one-step forecasts and observed values are mostly very similar; therefore, the points in both subplots of Figure 4.33 are overplotted within the narrow strips close to the lines of slope 1. Figure 4.34 shows the one-step forecasts and observed values as time series for the two datasets. Except for the two segments of missing data in the validation dataset, the forecasted and observed surface temperatures overlap each other most of the time for both datasets. Figure 4.33 and Figure 4.34 suggest that on average the model residuals are very small in magnitude for both calibration and validation datasets.

Figure 4.34: Surface Temperature Time Series: Observed vs. One-Step Forecast

The statistical summaries of one-step forecast residuals are listed in Table 4.13 for the calibration and validation datasets. The means of residuals are very close to zero. The forecast residuals of the validation dataset have a relatively larger standard deviation than those of the calibration dataset.

Table 4.13: One-Step Forecast Residuals of Surface Temperature Model

Dataset	Mean	Minimum	Maximum	Std	Std in °C
Calibration	0.0000	-0.1654	0.1223	0.0094	0.3459
Validation	0.0005	-0.0689	0.1208	0.0122	0.4490

The residual patterns shown in Figure 4.35 and 4.36 suggest an acceptable distribution, which is symmetrical, zero-centered and dispersed a little more highly than the residuals of the calibration data.

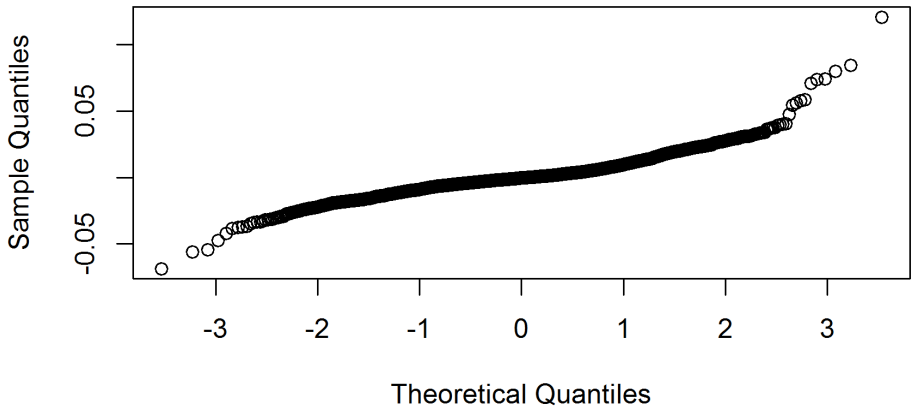


Figure 4.35: QQPlot of Surface Temperature Residuals – Validation Data

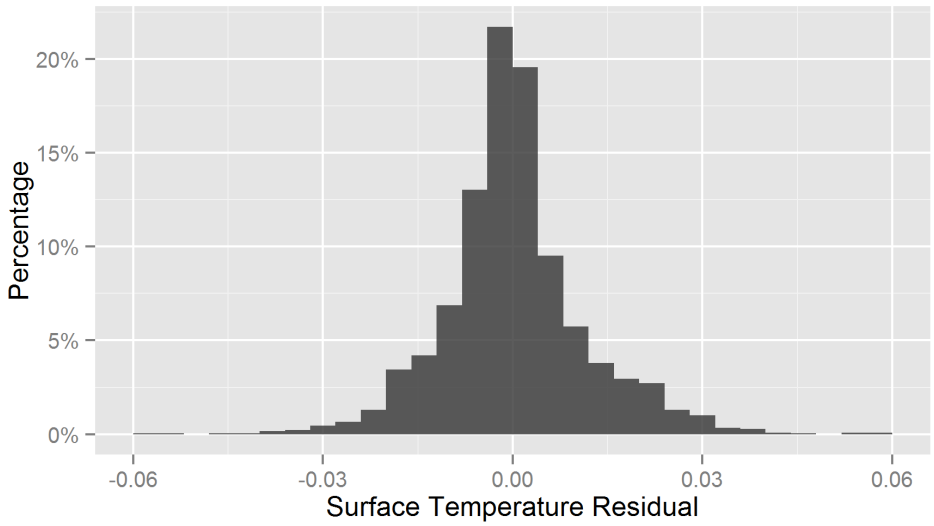


Figure 4.36: Histogram of Surface Temperature Residuals – Validation Data

The ACF and PACF plots of the residuals of the validation data are shown in Figure 4.37 and Figure 4.38. Both plots show some significant periodical patterns, especially the ACF plot. This suggests some seasonal components should be added and recalibrating a SARIMA model could solve the problem. Due to the following reasons, this study will not consider this extension:

- The sample size of this study is small and the covered period of the calibration dataset is less than 3 months. It is almost impossible to calibrate a reliable and practical SARIMA model. Ideally, more than one year of data should be used so that the seasonal and daily periodical components can be calibrated in a meaningful way.
- Even the model can not eliminate the periodical patterns of the validation data residuals; the resulting errors are generally very small in magnitude as suggested by the "Std in °C" column in Table 4.13. More than 95% of the residuals deviate less than 0.9°C from the true values. This high forecast accuracy eliminates the need for model recalibration.

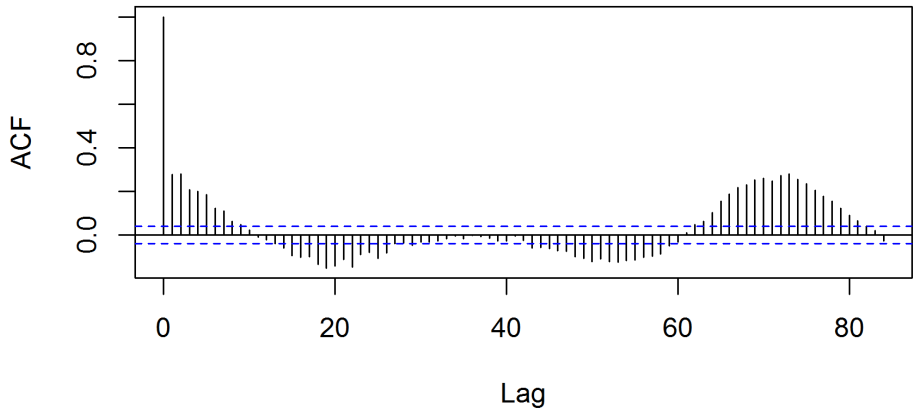


Figure 4.37: ACF of Surface Temperature Residuals – Validation Data

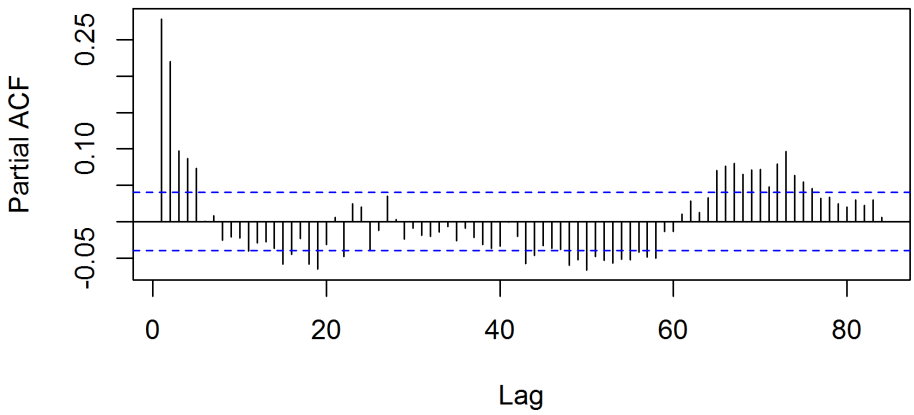


Figure 4.38: PACF of Surface Temperature Residuals – Validation Data

The one-step forecasting can be recursively made and chained together to make multi-step forecasts. In this way, short-term forecasting models can be used to make long-term forecasts. Figure 4.39 shows the recursive multi-step forecasting errors of the ST model for both calibration and validation datasets.

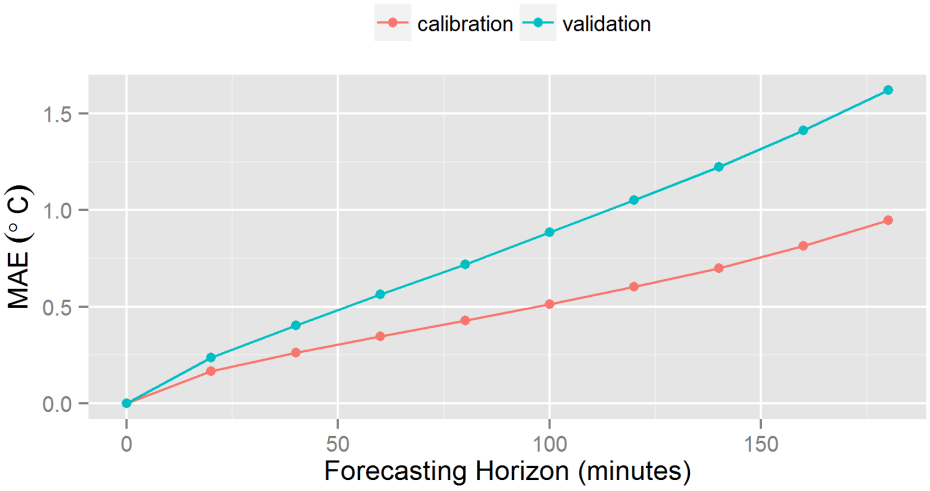


Figure 4.39: Multi-Step Forecasting Error – ST Model

The forecasting error is measured with the mean absolute error (MAE), which is given by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{4.30}$$

where n is the number of observations, $\hat{y}_i$ is the prediction, and y_i is the true value. The MAE error data shown in Figure 4.39 is listed in Appendix H.

As shown in Figure 4.39, longer forecasting horizons are associated with larger forecasting errors. The forecasting error of the calibration dataset is significantly smaller than that of the validation dataset at each forecasting horizon. The MAE of three-hour forecasts is less than 1°C for the calibration dataset, while larger than 1.5°C for the validation dataset. The average forecasting error within three hours is low for both datasets suggesting that the ST model can make relatively reliable forecasts on expanded horizons much longer than 20 minutes. This suggestion only makes sense when the conditions of weather, traffic and pavement contaminant can be accurately forecast, which is not always feasible.

4.8.2 Friction Model

Using the model calibrated in Section 4.7.2, the one-step ahead forecasts are made for both calibration and validation datasets and the observation and forecast values are plotted together in Figure 4.40.

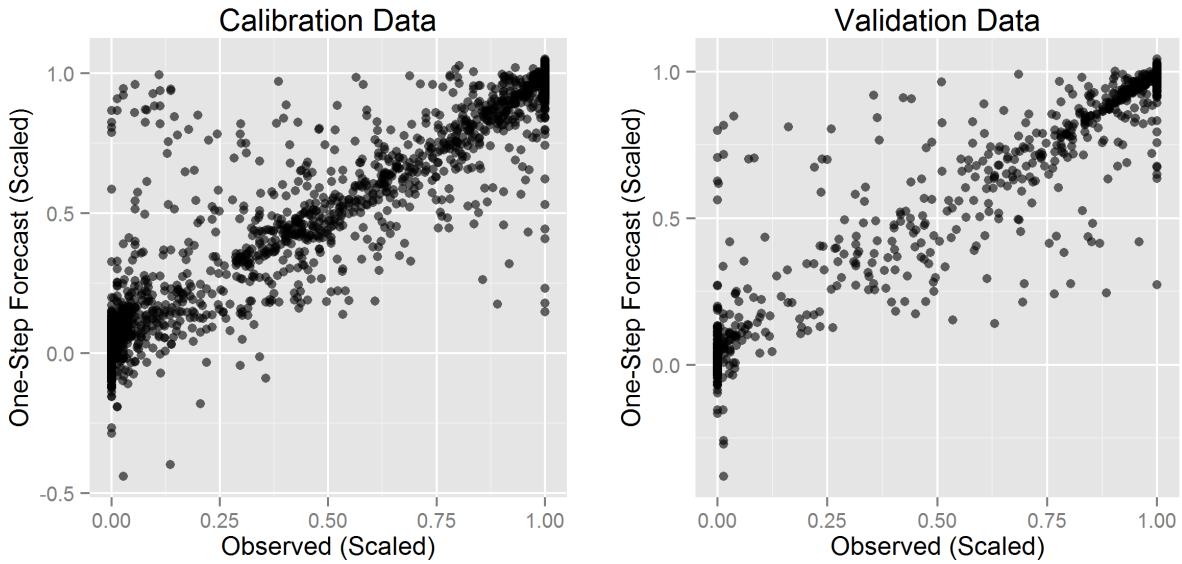


Figure 4.40: Friction Level Dotplots: Observed and One-Step Forecast

For both datasets, the one-step forecasts and observed values show strong correlation with relatively large dispersion. Figure 4.41 shows the one-step forecasts and observed values as time series for the two datasets. The forecast and observed friction levels are close to each other most of the time for both datasets. Figure 4.40 and 4.41 suggest that, on average, the model residuals are small in magnitude for both calibration and validation datasets.

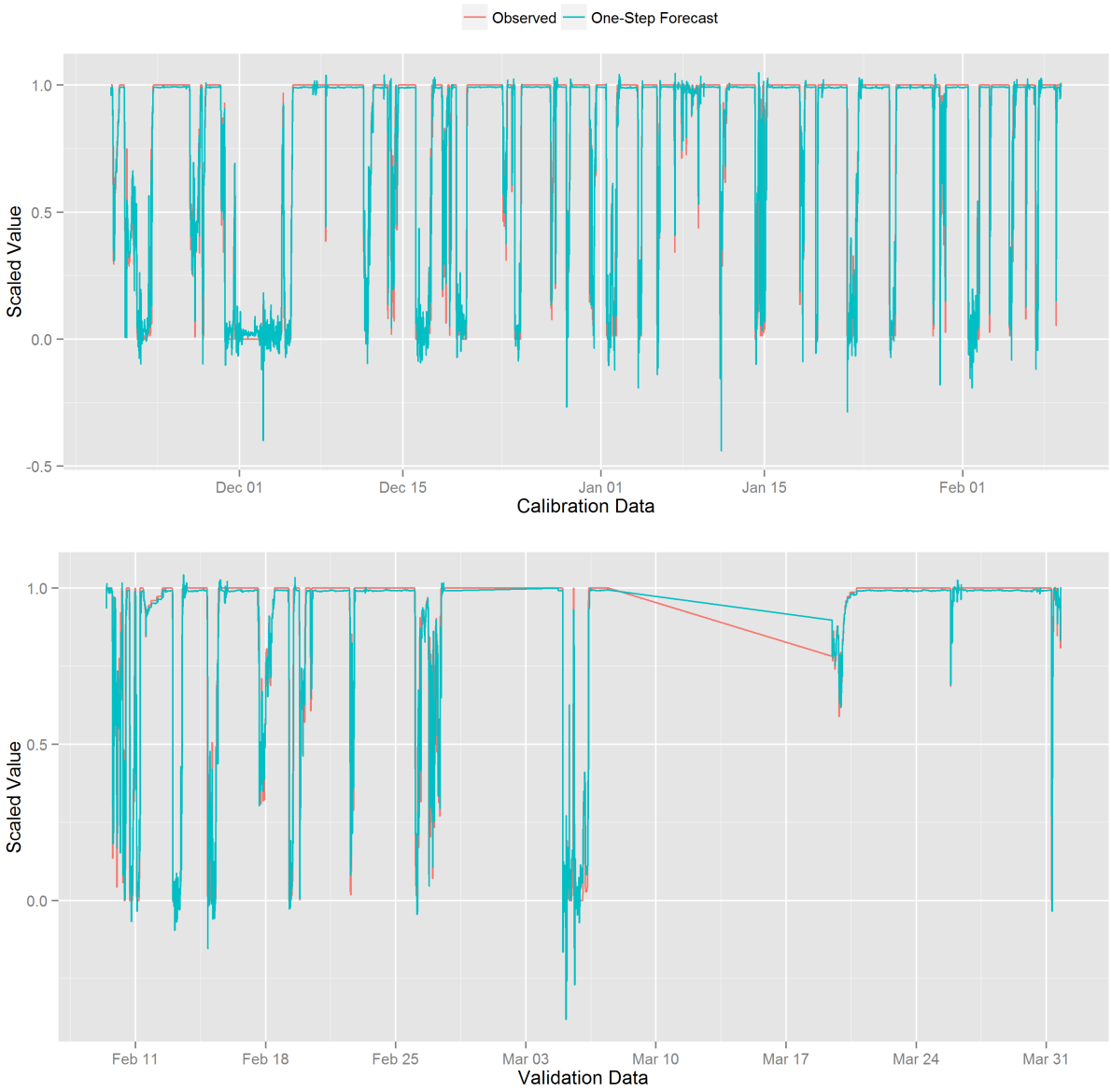


Figure 4.41: Friction Level Time Series: Observed and One-Step Forecast

The statistical summaries of one-step ahead forecast residuals are listed in Table 4.14. The means of both residuals are very close to zero. The residuals of the validation data have a little smaller standard deviation.

Table 4.14: One-Step Ahead Forecast Residuals of Friction Level

Dataset	Mean	Minimum	Maximum	Std	Std in Raw Scale
Calibration	0.000	-0.917	0.852	0.089	0.122
Validation	0.003	-0.812	0.728	0.086	0.118

The residual patterns shown in Figure 4.42 and 4.43 have an asymmetric distribution with a positive median. This suggests the forecast could be positively biased, but the magnitude of this bias is very small (less than 0.1).

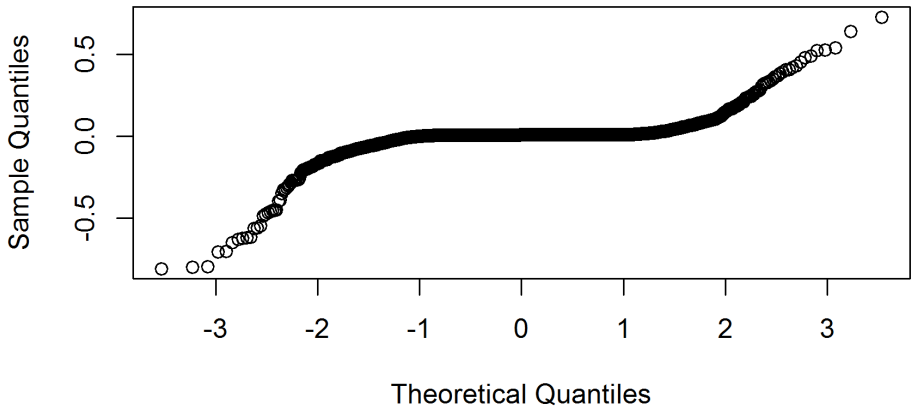


Figure 4.42: QQPlot of Friction Level Residuals – Validation Data

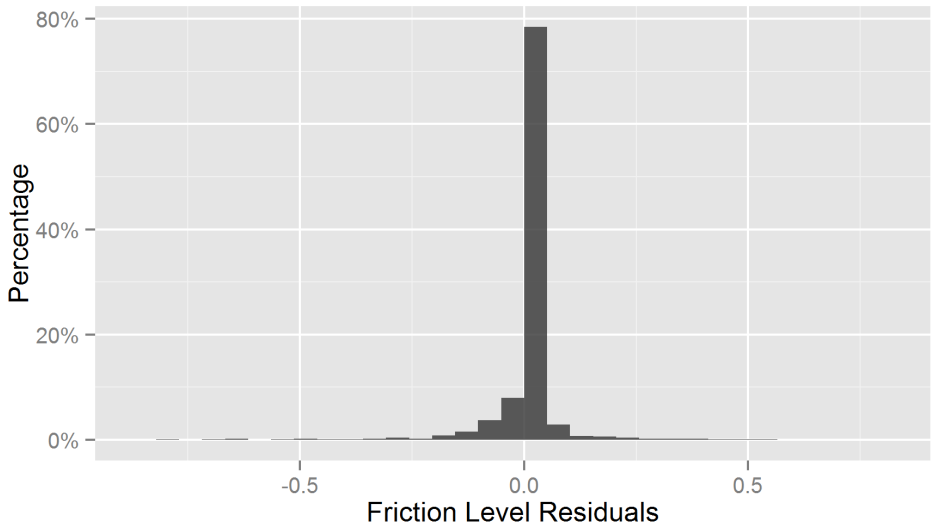


Figure 4.43: Histogram of Friction Level Residuals – Validation Data

The ACF and PACF plots of the residuals are shown in Figure 4.44 and Figure 4.45. There are only a few marginally significant lags and no periodical patterns, suggesting the model explains the auto-correlation very well for the validation data.

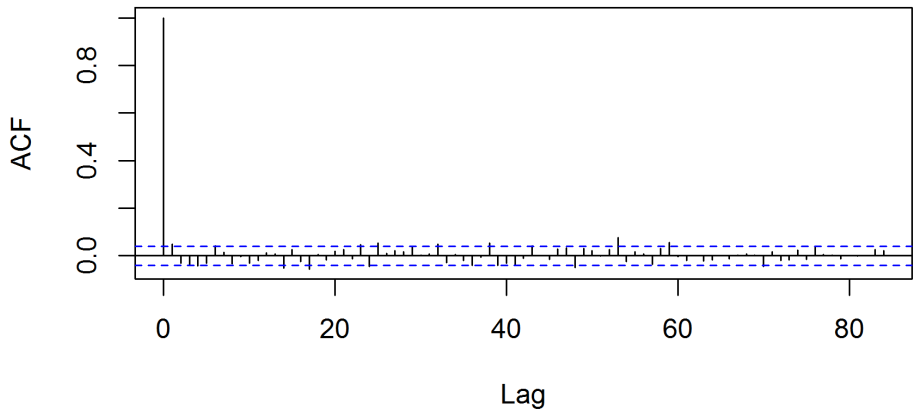


Figure 4.44: ACF of Friction Level Residuals – Validation Data

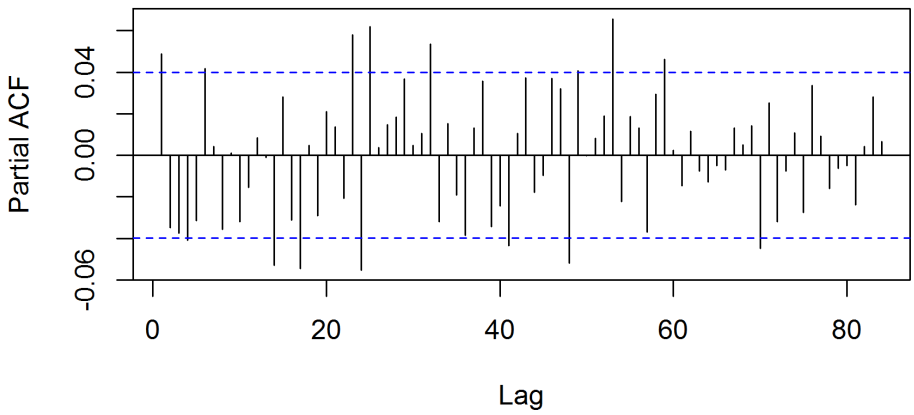


Figure 4.45: PACF of Friction Level Residuals – Validation Data

Figure 4.46 shows the recursive multi-step forecasting errors of the *FR* model for the calibration and validation datasets.

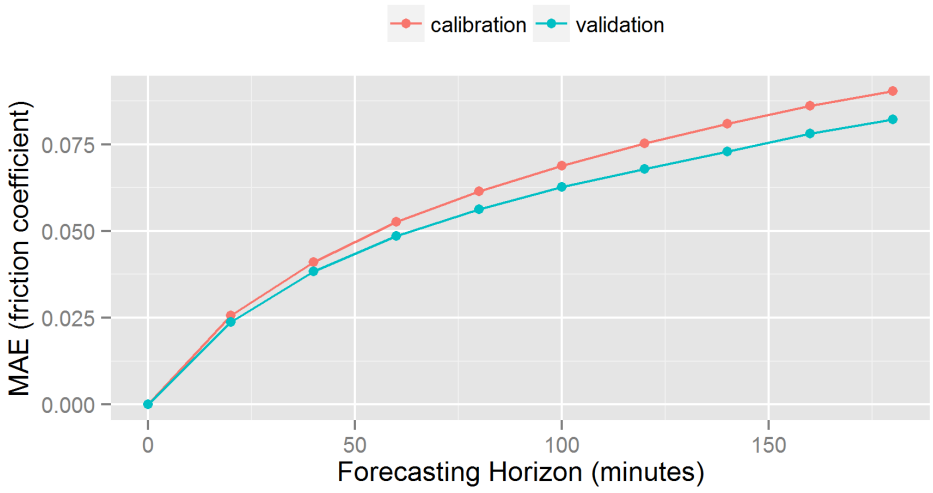


Figure 4.46: Multi-Step Forecasting Error – *FR* Model

The MAE data shown in Figure 4.46 are listed in Appendix H. As shown in Figure 4.46, longer forecasting horizons are associated with larger forecasting errors. Contrary to the common residual patterns found in statistical studies, the MAE of the calibration dataset is slightly larger than that of the validation dataset at each forecasting horizon. The three-hour forecasting MAEs for both datasets are smaller than 0.1 suggesting that the *FR* model can make quite reliable forecasts on longer horizons when the conditions of weather, traffic and pavement contaminant can be accurately forecast.

4.8.3 Contaminant Layers Model

Using the main effect VARMAX$(1, 0, 0)$ model calibrated in Section 4.7.3, the one-step ahead forecasts are made for both calibration and validation datasets and the observation and forecast values are plotted together in Figure 4.47, 4.48 and 4.49.

Figure 4.47: Water Layer Depth Dotplots: Observed and One-Step Forecast

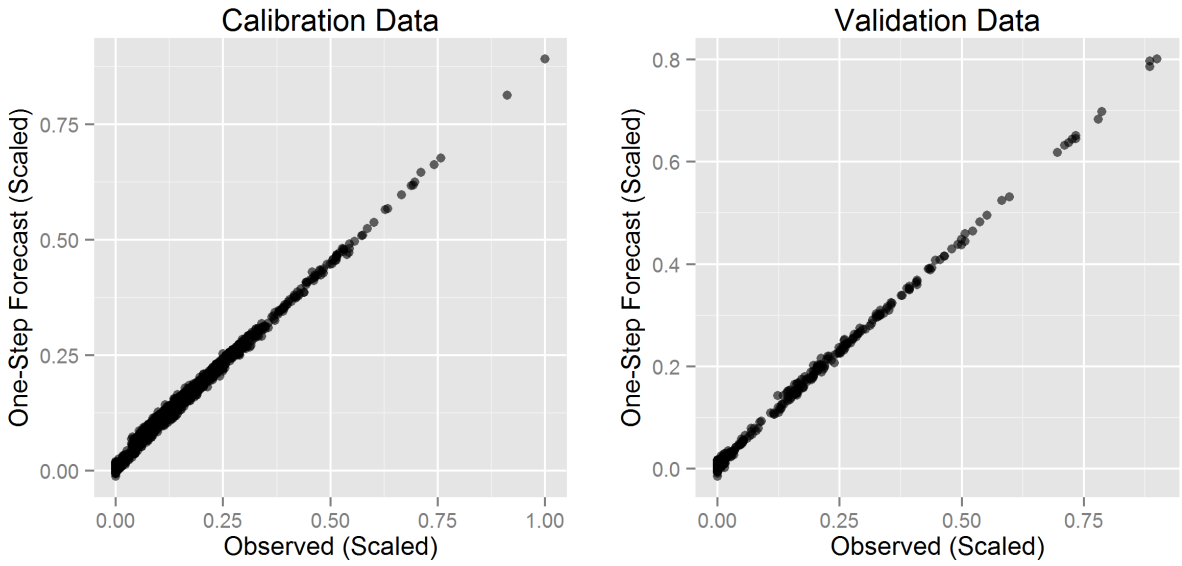


Figure 4.48: Ice Layer Depth Dotplots: Observed and One-Step Forecast

Figure 4.49: Snow Layer Depth Dotplots: Observed and One-Step Forecast

Except the *SL* submodel, the forecasts of the *WL* and *IL* submodels both have some systematic bias compared to the observations, as the corresponding figures show that the forecast values are slightly smaller than the observations on the whole, but the magnitude of this bias looks insubstantial. While the performances of the three submodels are all acceptable, Figure 4.50, 4.51 and 4.52 suggest that the *SL* submodel performs better than the *WL* and *IL* submodels as its one-step forecast values match up the observations more closely.

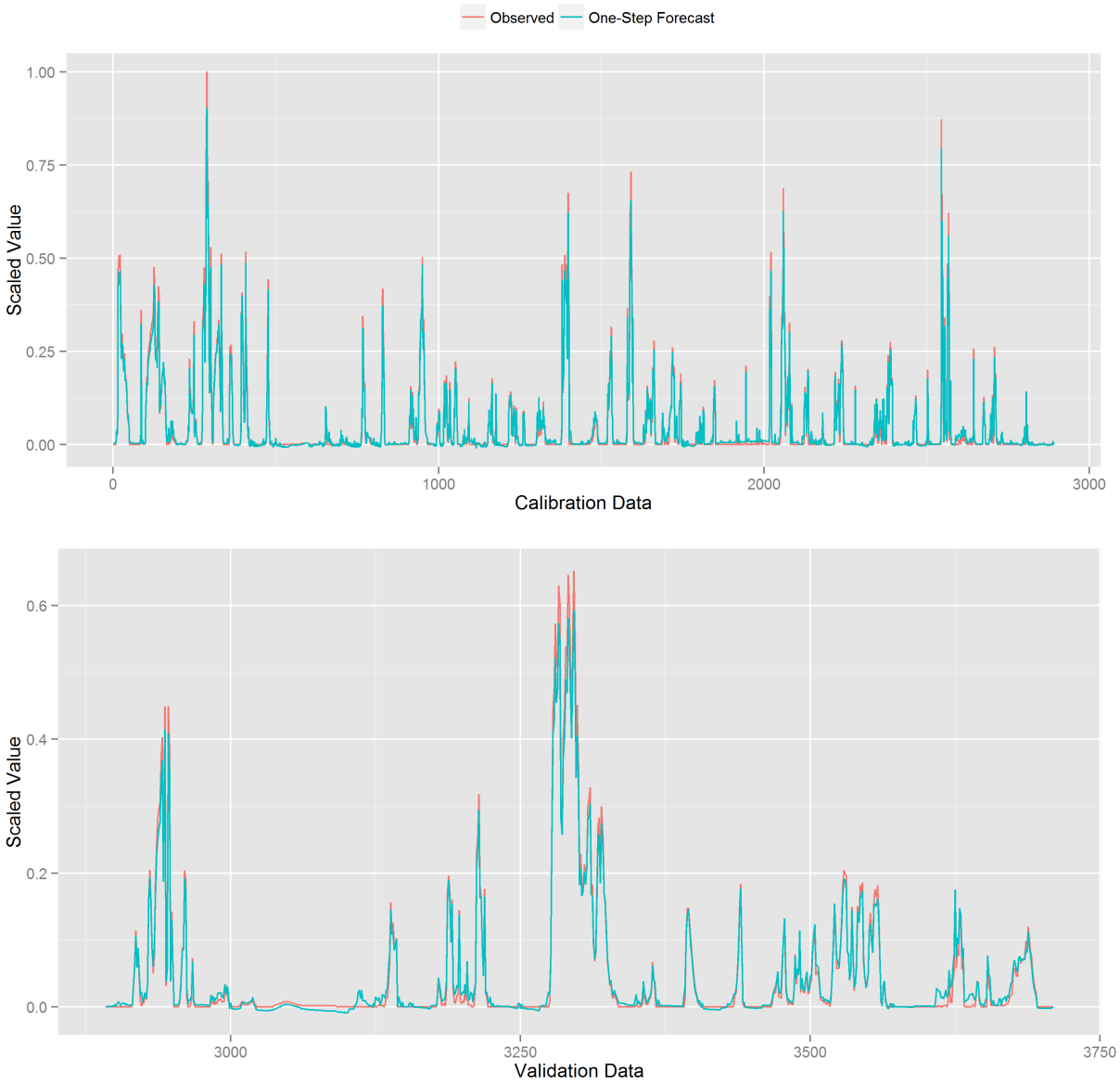


Figure 4.50: Water Layer Depth Time Series: Observed vs. One-Step Forecast

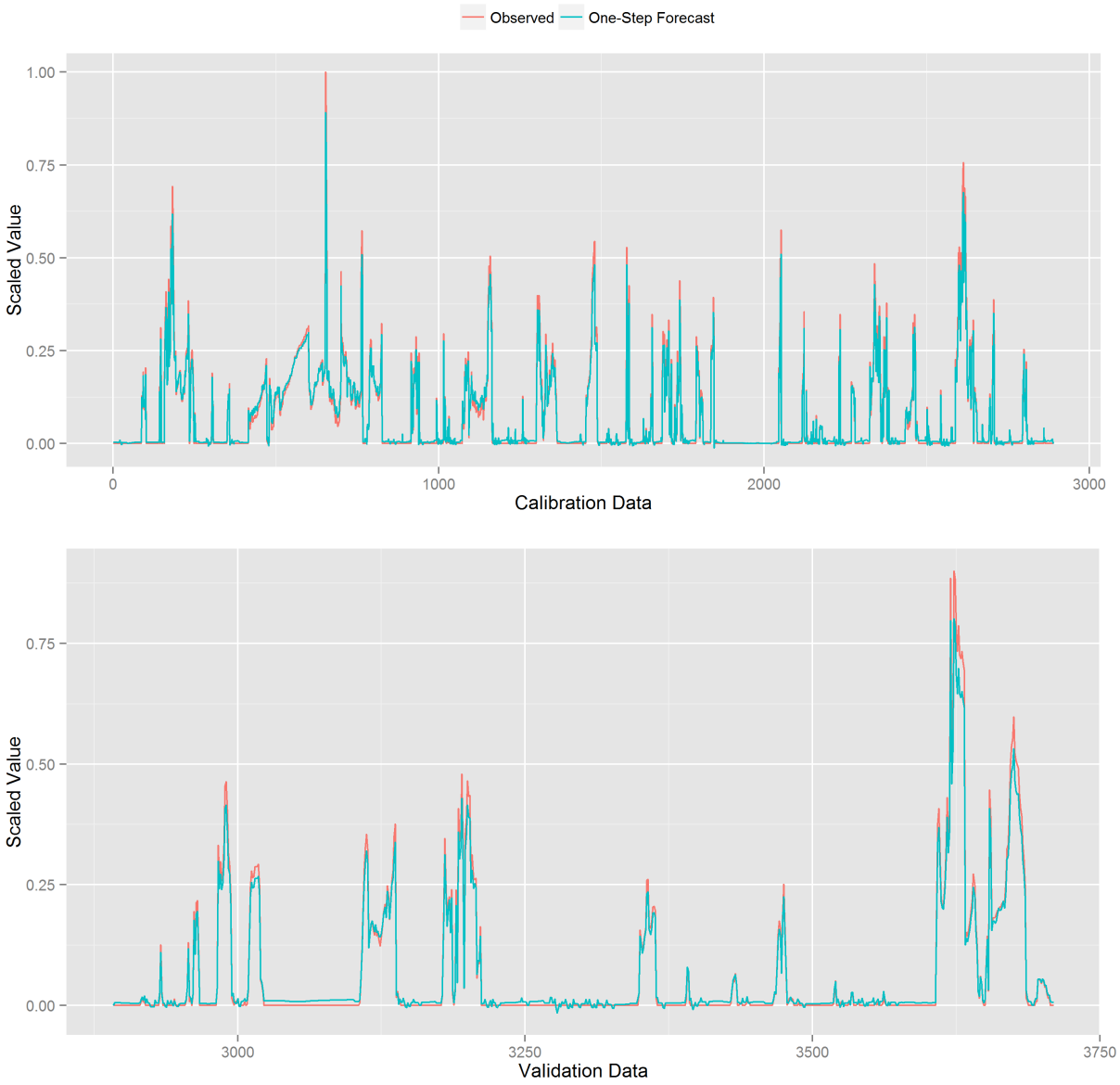


Figure 4.51: Ice Layer Depth Time Series: Observed vs. One-Step Forecast

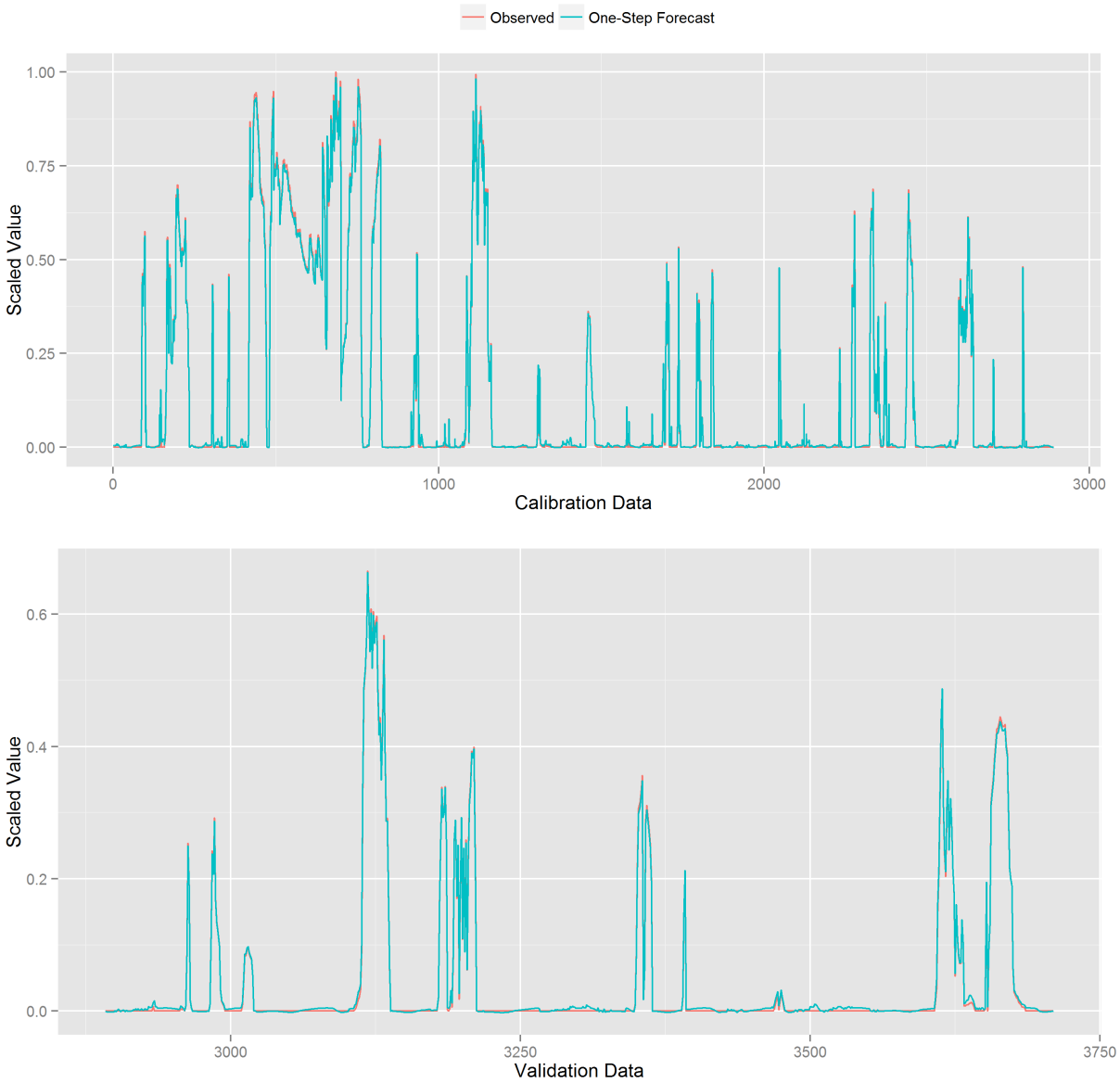


Figure 4.52: Snow Layer Depth Time Series: Observed vs. One-Step Forecast

The statistical summaries of one-step ahead forecast residuals are listed in Table 4.15. The means are all very close to zero. The residuals of the calibration and validation datasets of three submodels all have small standard deviations, implying they all have reliable forecasting quality.

Table 4.15: One-Step Ahead Forecast Residuals of Contaminant Layers

Dataset	Mean	Minimum	Maximum	Std	Std in mm
Calibration					
WL	0.000	-0.375	0.756	0.045	0.293
IL	0.000	-0.503	0.581	0.048	0.063
SL	0.000	-0.626	0.357	0.049	0.260
Validation					
WL	0.000	-0.028	0.064	0.012	0.078
IL	0.001	-0.019	0.099	0.017	0.022
SL	-0.001	-0.015	0.009	0.003	0.016

The patterns of residuals shown in Figure 4.53 and 4.54 have asymmetric distributions for all three contaminant layers. This suggests some prediction bias, but the magnitude of this bias is very small.

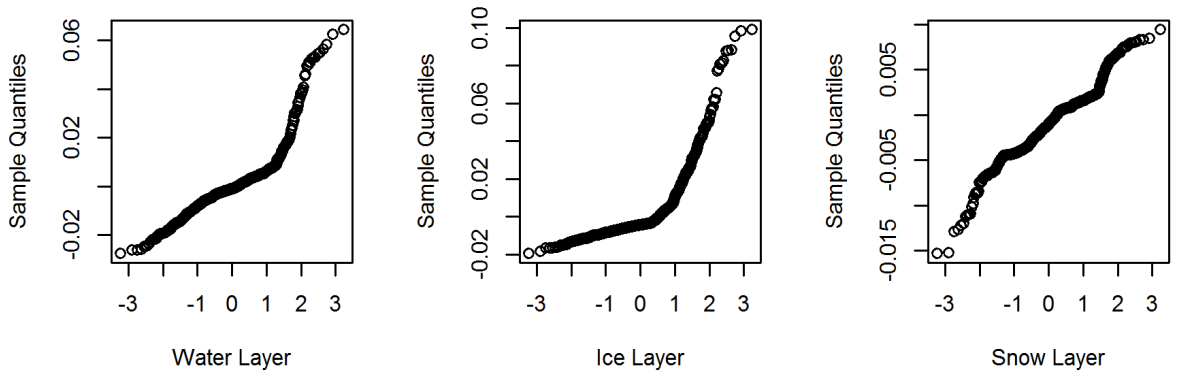


Figure 4.53: QQPlots of Contaminant Layer Residuals – Validation Data

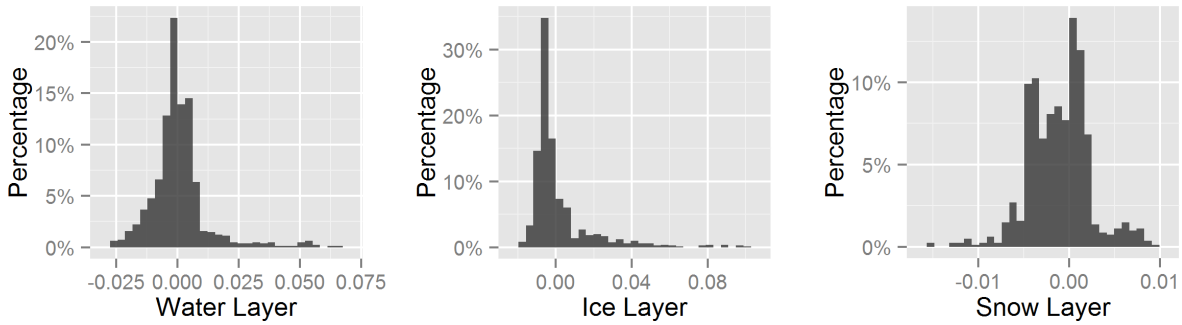


Figure 4.54: Histograms of Contaminant Layer Residuals – Validation Data

The ACF and CCF plots of the residuals of the validation data are shown in Figure 4.55. The residuals show both significant auto-correlation and cross-correlation at the first 6 lags, which suggests a structural change to the model, like adding MA components or differencing. As the magnitude of the residuals is already small, using a structurally complex model is unlikely to improve the overall forecasting quality, although the temporal correlation is dealt with in a more decent manner. When more data are available, preferably covering multiple winter seasons, more complex model structures should be tested to eliminate significant ACF and CCF lags in residuals.

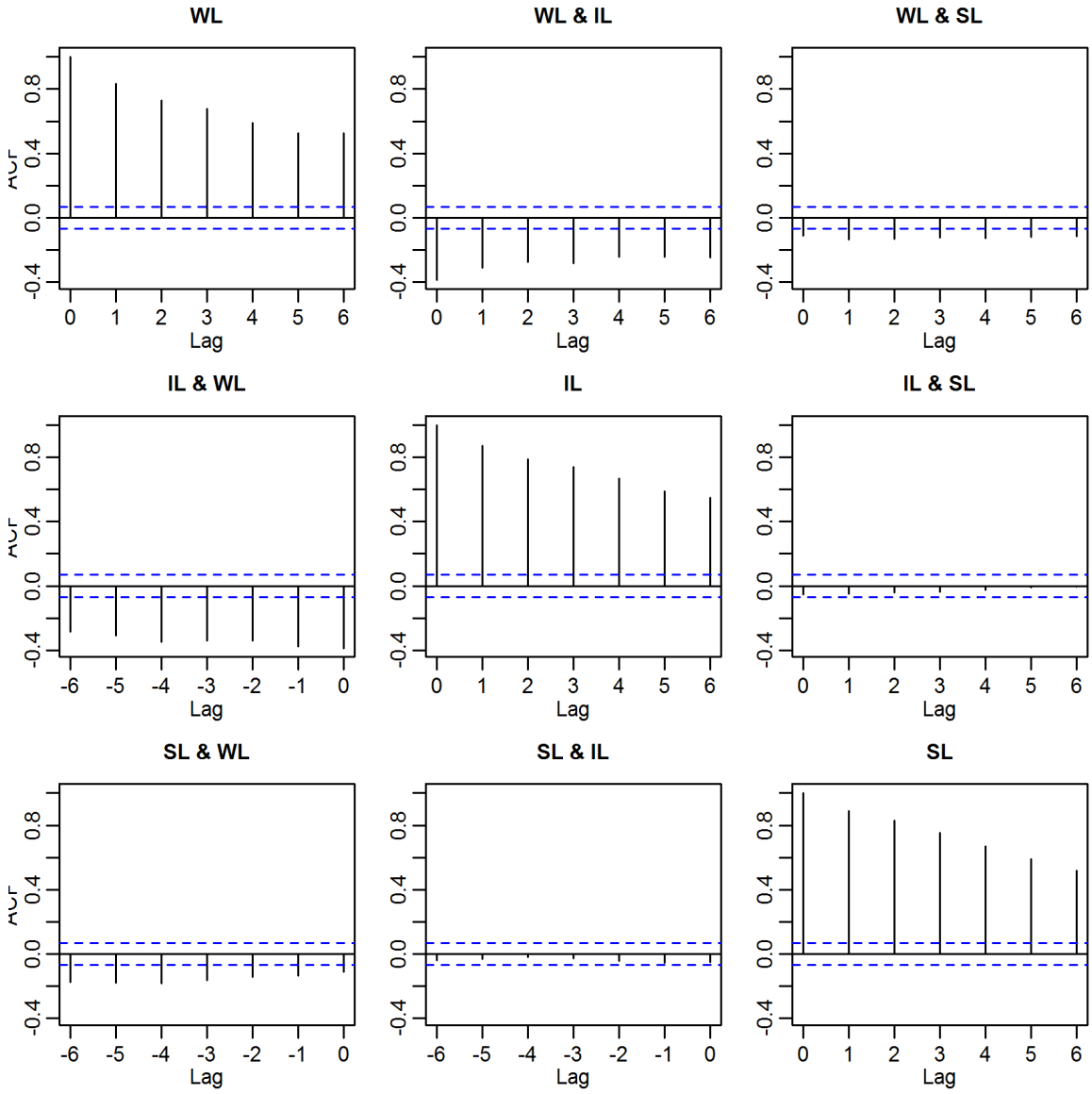


Figure 4.55: ACFs and CCFs of Contaminant Layer Residuals – Validation Data

The multi-step forecasting MAEs by the three submodels are shown in Figure 4.56, 4.57 and 4.58. The magnitudes of all MAEs are very small (less than 0.3mm, 0.08mm and 0.2mm respectively for three submodels) even for the three-hour horizon, and the *IL* submodel has the smallest MAE. This fact suggests that the three submodels can efficiently cooperate with each other, making reliable longer-term forecasts in a recursive manner. The MAE data shown in these three figures are listed in Appendix H.

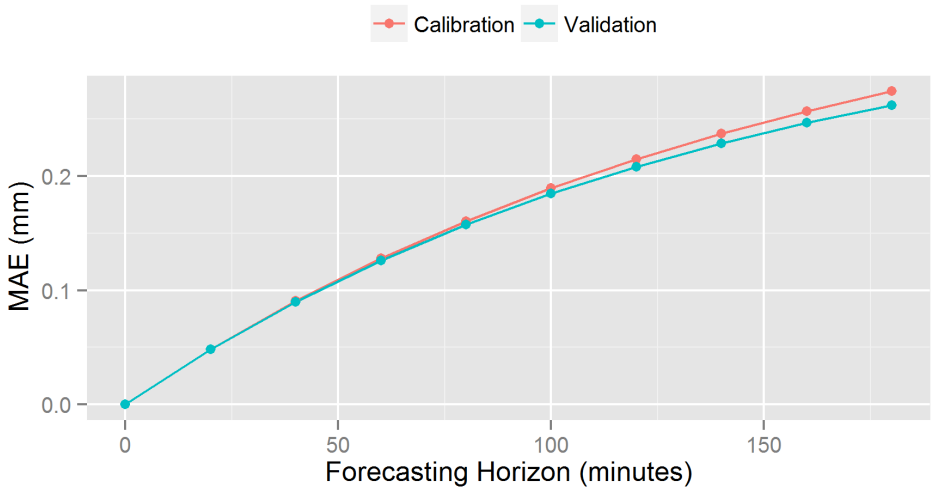


Figure 4.56: Multi-Step Forecasting Error – *WL* Submodel

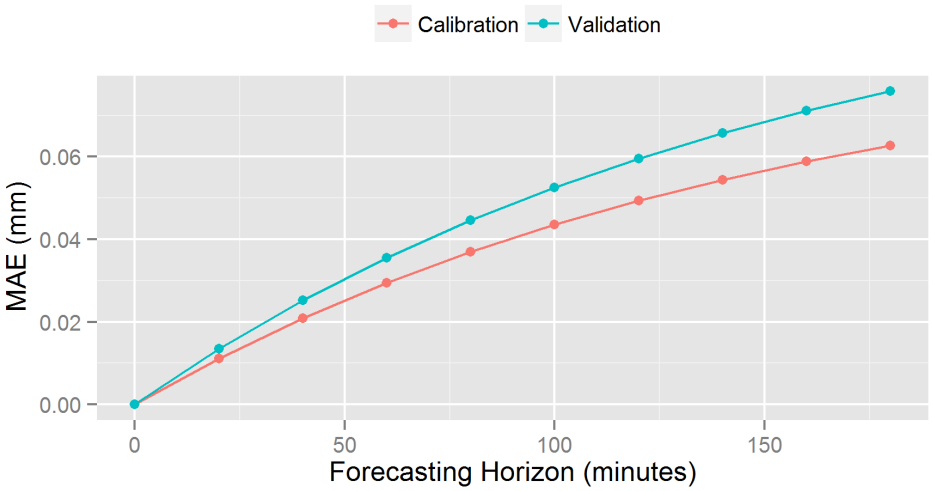


Figure 4.57: Multi-Step Forecasting Error – IL Submodel

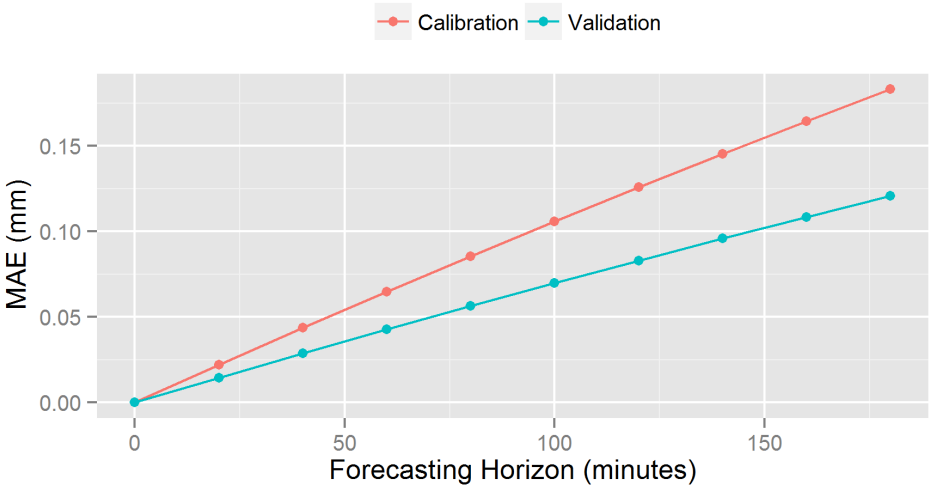


Figure 4.58: Multi-Step Forecasting Error – SL Submodel

4.9 Summary

In this chapter, a novel modelling framework for short-term RSC forecasting is proposed. Under this framework, the short-term changing process of the winter RSC measured by surface temperature, friction level and contaminant layer depths, is comprehensively explored and analyzed. A structure of forecasting models considering weather, traffic and maintenance operations are calibrated and validated using advanced time series analysis techniques. The forecasting errors of all the models are found to be small implying that both the proposed framework and the resulting solutions closely match the real-world observations.

Chapter 5

Conclusions and Future Work

Based on the findings and insights that have been gained in this research, this chapter makes general conclusions, highlights important contributions, points out major limitations, and suggests future research directions.

5.1 Major Conclusions

The main objective of this research is to address two challenging problems in winter roadway maintenance: reliable RSC estimation and accurate RSC forecasting.

For the RSC estimation problem, a set of aggregate measures of CFM was systematically studied and discreetly introduced into the modelling process based on their physical meanings. A structure of LR models were calibrated and validated using those aggregate measures, and the model performance was compared with the SVMs calibrated in parallel. The models were found satisfactory in terms of classification hit rate. The proposed modelling framework and methodology were also tested using vehicle speed data. Although this part of the work was not aimed at developing interpretable and applicable models, the modelling results have shown some success.

For the RSC forecasting problem, all influential factors including weather, traffic and maintenance operations were systematically analyzed in terms of their effects on RSC changing process. Several major RSC measures, such as surface temperature, friction level, and contaminant layer depths, were studied and modelled with the stochastic time series analysis approach. The resulting models are easy to interpret and capable of making highly accurate numerical forecasts.

5.2 Contributions

This research has made several significant contributions in two main areas, namely, RSC estimation and RSC forecasting. In the RSC estimation part, the major contributions of this research can be summarized as follows.

- The study has examined CFM from the perspectives of its aggregate measures, specifically probability distribution and spectral density features. The information contained in these features is for the first time systematically explored and interpreted in terms of their correlations with RSC.
- A new study framework is proposed to utilize the RSC information contained in those CFM aggregate measures to classify the RSC type. The modelling results have proved that this framework can significantly improve RSC estimation accuracy compared to existent models.
- The study methodology is directly aimed at capturing the uncertainty nature of friction measurements. A multi-level model structure is designed and all the single models are LR models which can make straight forward stochastic estimates. It is a new and efficient approach to capturing the overlapping structure of CFM ranges over different RSC types, a problem which has not been satisfactorily tackled by other field studies.
- Under a similar study framework, a set of RSC estimation models is also calibrated using speed data. Although the purpose is only to examine the possibilities, it is the first time that vehicle driving data are organized and analyzed under a comprehensive study framework for RSC estimation purposes.

In summary, for RSC estimation, this research has made significant contributions in introducing new explanatory features, proposing a well-grounded analysis framework, and developing efficient modelling methodologies.

The major contributions in the RSC forecasting part are summarized as follows.

- A rich set of numerical RSC measures is studied and the forecasting models are calibrated and validated for each one of them. It is the first study to have developed highly efficient models with such a complete view.

- A new study framework has been designed to consider all important conditional factors, including weather, traffic and maintenance operations. This design can reliably capture the evolution pattern of the winter RSC.
- The maintenance operations, especially saltings, are handled by loosening the strict Markovian assumption, i.e., a history instead of one single time interval of salting operations is considered. In this way, the variation of snow/ice melting speed caused by both residual salt amounts and salt/contaminant mixing states is incorporated in the forecasting model, which enables accurate short-term forecasting for contaminant layers. This approach practically circumvents a major limitation of previous studies and makes the post-salting RSC forecasting more reliable and accurate.
- Under the proposed study framework, several advanced time series modelling methodologies are introduced into the analysis, which can capture the highly complex interactions between RSC measures and conditional factors at once. Those methodologies, especially the univariate and multivariate ARIMA methods, are for the first time applied to the winter RSC evolution process.
- With the innovative time series analysis methods, the calibrated models are simple in structure, easy to interpret and mostly consistent with physical knowledge. Compared to contemporary studies, such models provide extra flexibility and economy for refactory, tuning and deployment.
- All the modelled RSC measures are numerical and the forecast errors are relatively very small, suggesting the empirical approach could be an efficient alternative to the physical approach. With the well-designed modelling methods, the resulting empirical models as calibrated in our study can change the role of empirical models from a long-term low-temporal-resolution maintenance planning facility into a high-temporal-resolution high-accuracy decision support and simulation tool.

In summary, for the RSC forecasting problem, this research has made several significant breakthroughs in modelling methodology and analysis techniques. It is the first study that clearly shows the power of empirical forecasting models.

5.3 Limitations

This thesis achieves its objectives satisfactorily and is expected to contribute significantly to the research community. However, there are several limitations. For the RSC estimation part:

- Due to the data aggregation process, the available sample size for some RSC types become very small, causing some difficulties in both model calibration and validation.
- All the test runs are done on a single highway, so without careful validation, the modelling results may not be directly applicable to other highways of different surface textures, although the effect of the latter is expected to be small.
- Although it is proposed that relevant environmental factors should be considered as possible predictors, our collected dataset is short of such data. Inclusion of these weather related factors could largely improve RSC estimation reliability.
- RSC types were classified manually by a couple of university students based on video images. The students received a short training by the thesis author only; therefore, the sample classification may be subject to personal bias and classification errors.
- Only the measurements of one type of continuous friction meter are used in this study, so the findings can not be generalized to other friction meters without further study.

For the RSC forecasting part:

- Only one winter season of data is available, which may affect the application scope of the resulting models. The ideal data should cover as many types of snow storms as possible, and thus large variances of conditional factors can be captured.
- As the data on plowing operations and traffic are not completely accurate, the impact of this inaccuracy needs to be evaluated in future study.
- Only one study site maintained by the same contractor is tested in this research. As a result, the maintenance data are limited in terms of maintenance strategies, methods and materials being applied. Furthermore, the effects of local factors, like road geometries, roadside geospatial profiles, and road materials, which could significantly affect RSC, can not be studied in this research. Therefore the resulting models need to be further validated before being applied to other sites.
- As the study is focused on spot-wise forecasting, its applicability to forecasting the RSC of a whole road section needs to be evaluated in future study. This is a major limitation of all models based on the data collected by static RWIS stations.

5.4 Future Directions

With all its contributions and limitations, the proposed research work can be extended in many ways, and first and foremost is the availability of more sample data.

The sample problem in this study is not caused by limited resources or non-carefully designed data collection plans, but by the modelling methodologies. CFM data were collected by well trained MTO staff and the total length of the test runs is enough for developing mean-friction based models. However, after the RSC type designation and aggregation processing, which was proposed after data collection, the sample size turns out to be not sufficiently large as expected.

The data collection and integration for RSC forecasting has even more challenges. MTO's maintenance database did not keep any hourly or more granular operation records, which are critical for successful short-term RSC forecast modelling. Accurate plowing and salting information were not always in the AVL dataset as well. In addition, there were very few Vaisala stations installed in Ontario, so it seriously affects the amount of available weather and RSC data.

The only way to overcome such sampling problems in future study is to promote better communications and interactions between government agents, contractors and researchers. Ideally, the proposed modelling frameworks are integrated with available decision support or performance monitoring systems, so that the whole process of data collection, data storage and processing, and model tuning can be seamlessly and dynamically realized. Besides this long-term goal, there are several conspicuous ways to extend current works. For RSC estimation:

- Collect CFM data on different classes of highways and road surface textures covering more winter seasons. The modelling results should also be validated across different friction meters. As for the modelling techniques, based on new datasets, different aggregation intervals and aggregate measures can be tested.
- Using driving data to estimate RSC is very attractive. Our study only tests the speed data, but other driving data like 3-axle accelerations could work much better, .

For RSC forecasting:

- RSC, weather, traffic and maintenance data of more winter seasons from more sites should be collected. These sites should cover multiple patrol routes maintained by different contractors.

- Although this study uses the data from a static RWIS station, specific RWIS sensors, like the ones used in this study, can be mounted on a driving vehicle and working mobile (Saarikivi, 2012). In this way, spatially continuous RSC measures can be collected, and thus the modelling framework can be extended to make forecasts for a whole road section.
- The performance of the empirical model developed under our framework should be compared with those physical models that may be able to tackle the lagging effect of salting more effectively.

5.5 Final Remarks

In this thesis research, the research problems are addressed mainly through statistical modelling methodologies. The collected data are processed and analyzed to obtain extra knowledge about the studied physical processes. The resulting models are the outcome of this data-mining progress and can be used to solve the relevant engineering problems. This approach is considered suitable for modelling physical processes with significant uncertainty, like friction measurements, or of highly complex mechanism, like RSC changes in time.

This research provides a foundation to researchers in winter road maintenance for expanding the existing toolsets in operation decision support and performance measure systems. The proposed research framework for RSC estimation may evolve in a natural progression towards developing more reliable and robust winter road monitoring systems. For the RSC forecasting problem, the modelling framework and techniques relying on ARIMA analysis are for the first time systematically tested. The modelling methodology coupled with advanced RSC sensing technology could evolve into sophisticated RSC forecasting systems.

Some major findings of this thesis research could stimulate many interesting studies. For example, our study has tested the possibility of using vehicle driving data to estimate winter RSCs. As such data are ubiquitously available with the help of telematics devices and smart phones, significant research success in this area could make each vehicle on the road become a mobile RSC probe.

APPENDICES

Appendix A

Vaisala Sensor Set Specifications

Vaisala Non-Invasive Pavement Surface State Sensor DSC111 (Spectro)

Features and Benefits

- Non-invasive remote surface state sensing
- Spectroscopic measuring principle, identifying the presence and thickness of:
  - Water
  - Ice
  - Slush
  - SnowFrost
- Unique measurement of road friction or grip
- Accurate and stable measurement results even with intense traffic
- Eye-safe laser technology
- Easy installation and service
- Low maintenance costs
- Weather-proof, durable design
- Easy integration with Vaisala ROSA Road Weather Station, or can operate as a stand-alone solution with solar/gsm options
- NTCIP compliant when used with Vaisala DMC interface card

The unique Spectro sensor eliminates the service disruption which was previously associated with the installation of a road weather station. The remote installation means that there is no requirement to slot-cut the surface or close the road. The sensor may be installed in a remote location on a pole adjacent to the road, or as an addition to the Vaisala ROSA Road Weather Station.

The spectroscopic measuring principle enables accurate measurement of the amounts of water, ice, and snow. Water and ice are measured independently of each other, enabling Spectro to accurately report the surface state.

Spectro provides an accurate measure of the presence of ice crystals well before they cause the road to be slick. The Winter Maintenance engineer is therefore able to monitor even the smallest changes in grip in order to take the appropriate remedial action.

The water reading is useful for advanced warning of hydro-planing.

Spectro has proven its capabilities during three years of intensive field testing in collaboration with Vaisala customers.

Together with Cyclo, which remotely measures surface temperature, Spectro forms a versatile stand-alone weather station.

Vaisala Inc., Boulder Operations
P.O. Box 3659, Boulder CO 80307-3659, USA
Email: trafficweather@vaisala.com • www.vaisala.com
Switchboard +1 303 499 1701 • Customer Support Toll Free +1 800 408 9457 • Fax +1 303 262 4058

TECHNICAL DATA **DSC111 (Spectro)**

ELECTRICAL	
Power supply	9 ... 30 VDC
Power consumption for operation	0.7 W above 5°C max 1.9 W below 5°C
Power consumption for lens heaters	0 ... 4 W user adjustable
Interfaces	RS-485 isolated, RS-232
Connectors	3 x M12 (5 pins)
1:	RS-485 and power, male
2:	RS-232, male
3:	RS-485 and power, female
	Extension connector for the DST111
Cables	3 m, 10 m, 25 m
	One end without connector
	0,6 m extension cable to the DST111

ENVIRONMENTAL	
Operating temperature	-40°C/°F...+60°C/140°F
Operating humidity	0 ... 100 % RH
CE Compliant	IEC(EN)-61326
Safety	Eye-safe, Laser class 1
Vibration	IEC 60721-3-3

INSTALLATION	
Measuring distance	6 ... 50 ft
Measuring area	Diam. 10" at 30ft
Installation angle from the horizontal line	30 ... 85°
Fits onto the standard sensor arm DM32ARM with cross-section of 40 mm x 40 mm	

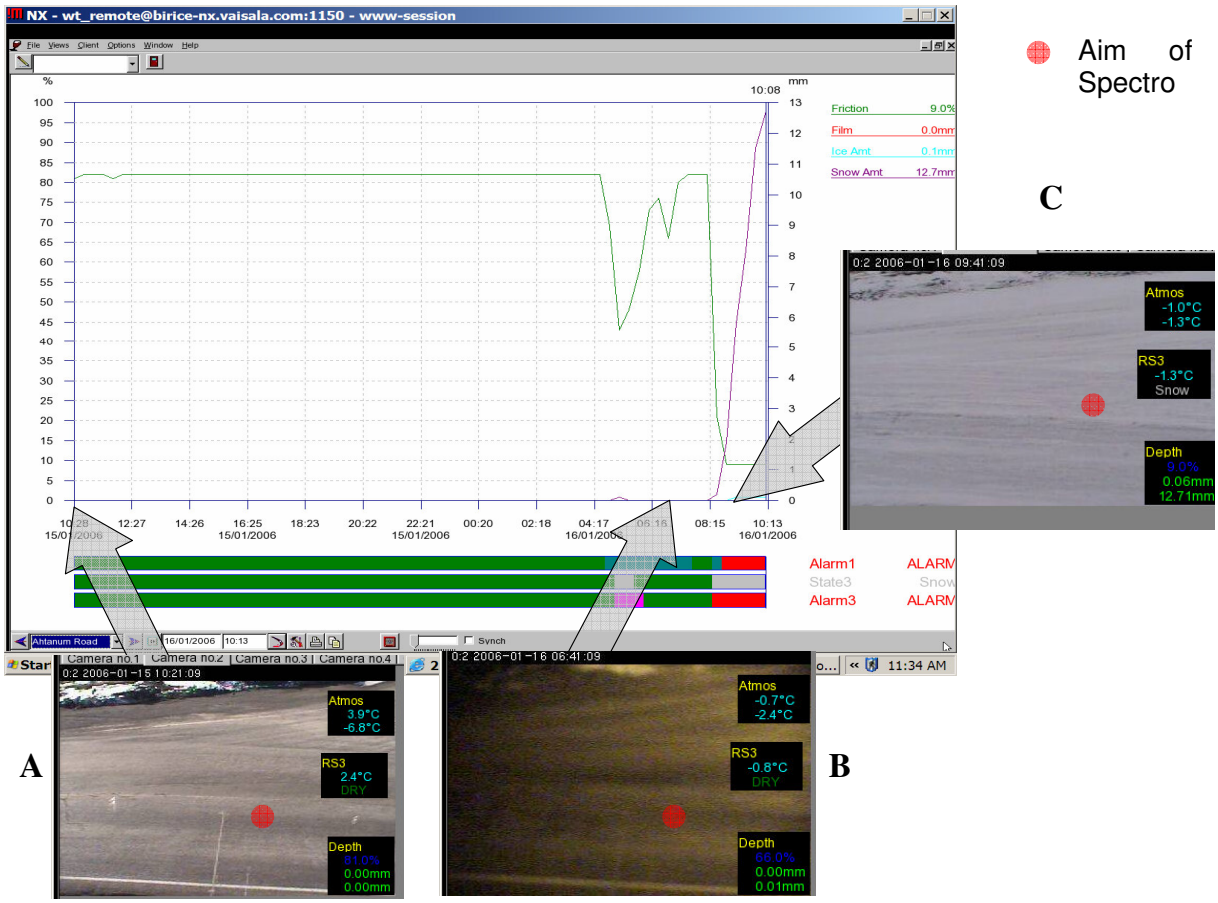
MEASURING RANGE	
Layer thickness*	
Water	0.00 ... 2 mm
Ice	0.00 ... 2 mm
Snow	0.00 ... 20 mm
Resolution	0.01 mm
Friction coefficient	0.01 ... 1.00
Resolution	0.01 units
Surface states	Dry, Moist, Wet, Snow/Frost, Ice, Slush

MECHANICAL	
Dimensions (cm)	46 x 21 x 14
Weight	7 lb / 3.7 kg

Units can be displayed in inches

Vaisala Inc., Boulder Operations
P.O. Box 3659, Boulder CO 80307-3659, USA
Email: trafficweather@vaisala.com • www.vaisala.com
Switchboard +1 303 499 1701 • Customer Support Toll Free +1 800 408 9457 • Fax +1 303 262 4058

Example of Spectro display



The graph shows Spectro data from 1028am on January 15th to 1013am on January 16th 2006. The green line (friction or grip) remains constant at 82% indicating a dry road (see camera image A) until the first light snow shower occurs at 4.20am.
From 8.15am snow accumulates rapidly, as shown by the purple line, which increases to a water equivalent level of about half an inch (12.7mm) at 1013am. This corresponds with a reduction in friction (or grip) as the green line drops to about 10%.
Camera image A shows a dry pavement with spectro aimed just to the right of the embedded sensor. Images B & C are aimed at the same location, but at later times, showing the snow accumulation.

Vaisala Inc., Boulder Operations
P.O. Box 3659, Boulder CO 80307-3659, USA
Email: trafficweather@vaisala.com • www.vaisala.com
Switchboard +1 303 499 1701 • Customer Support Toll Free +1 800 408 9457 • Fax +1 303 262 4058

Vaisala Non-Invasive Pavement Surface Temperature Sensor DST111 (Cyclo)

Features and Benefits

- Non-invasive remote temperature measurement
- Unique correction of the error caused by the emissivity of the road surface, negating the need for emissivity adjustment
- Easy installation and service
- Low maintenance costs
- No internal moving parts
- Stable measurement results even with intense traffic
- Weather-proof, durable design
- Assessment of air temperature and humidity
- Easy integration with Vaisala ROSA Road Weather Station
- Capability to act as stand-alone device in remote locations with solar/gsm options
- NTCIP compliant when used in conjunction with Vaisala DMC interface card

The unique Cyclo sensor provides a non-invasive alternative to measuring pavement surface temperature. By measuring the infrared radiation emitted by the surface and applying intelligent signal processing, Cyclo provides a reliable remote surface temperature measurement.

Cyclo provides reliable results in conditions where most of the commercially available infrared sensors fail. At night time, when the road surface is cooling under a clear sky, conventional infrared sensors provide an error of up to -3°C (-7°F) due to emissivity conditions of the road surface. Cyclo compensates for this error by its unique design.

Installation of Cyclo is easy, requiring no slot cutting or closure of the road. Supplied with solar/gsm options, the sensor is ideal for stand-alone operation in remote/in-fill locations and on bridge decks (when used in conjunction with DSC111 – Spectro). The sensor is simply installed on a mast, or existing structure beside the road.

Cyclo can also be installed alongside an existing Vaisala ROSA Road Weather Station.

Together with Spectro, which measures surface state, Cyclo forms a versatile stand-alone weather station.

Vaisala Inc., Boulder Operations
P.O. Box 3659, Boulder CO 80307-3659, USA
Email: trafficweather@vaisala.com • www.vaisala.com
Switchboard +1 303 499 1701 • Customer Support Toll Free +1 800 408 9457 • Fax +1 303 262 4058

Appendix B

Traffic Estimation Model

Table B.1: Traffic Estimation Model for Dunvegan

	Traff Estimation Model $ln(Hourly\_Traffic\_Dunvegan)$
$(Intercept)$	1.68^{***}
	(0.22)
$ln(Hourly\_Traffic\_Ottawa)$	0.17^{**}
	(0.06)
$ln^2(Hourly\_Traffic\_Ottawa)$	0.04^{***}
	(0.00)
Multiple R-squared	0.87
Adjusted R-squared	0.87
F-statistic	3.55e+04 on 2 and 10959 DF
p-value	< 2.2e-16

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Appendix C

ARMAX$(p, 0, 0)$ Models for Friction Level

C.1 ARMAX$(1, 0, 0)$

```
Series: endo_vars
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1  intercept    Plow      AT   Precip  Traffic  Salting_L0
      0.9905     0.1185  0.0001  0.5576  -0.0225   0.0507     -0.0037
s.e.  0.0027     0.0306  0.0018  0.0193   0.0087   0.0069      0.0036
      Salting_L1  Salting_L2  Salting_L3  Salting_L4  Salting_L5  Salting_L6
         -0.0041     -0.0085     -0.0151     -0.0143     -0.0113     -0.0096
s.e.      0.0047      0.0056      0.0061      0.0063      0.0063      0.0061
      Salting_L7  Salting_L8  Salting_L9      WL       IL       SL
         -0.0089     -0.0059     -0.0048  0.0077  -0.0086  -0.0016
s.e.      0.0056      0.0047      0.0034  0.0108   0.0086   0.0110

sigma^2 estimated as 0.0001969:  log likelihood=6840.45
AIC=-13640.9   AICc=-13640.54   BIC=-13525.21
```

C.2 ARMAX$(2, 0, 0)$

```
Series: endo_vars
ARIMA(2,0,0) with non-zero mean

Coefficients:
         ar1      ar2  intercept    Plow      AT   Precip  Traffic  Salting_L0
      1.5659  -0.5747     0.2726  0.0001  0.2754  -0.0236   0.0383     -0.0036
s.e.  0.0183   0.0182     0.0298  0.0011  0.0201   0.0057   0.0102      0.0030
      Salting_L1  Salting_L2  Salting_L3  Salting_L4  Salting_L5  Salting_L6
         -0.0041     -0.0078     -0.0133     -0.0114     -0.0085     -0.0058
s.e.      0.0052      0.0070      0.0082      0.0088      0.0088      0.0082
      Salting_L7  Salting_L8  Salting_L9      WL       IL      SL
         -0.0043     -0.0023     -0.0024  0.0047  -0.0082  0.0040
s.e.      0.0070      0.0051      0.0029  0.0079   0.0061  0.0082

sigma^2 estimated as 0.0001443:  log likelihood=7213.59
AIC=-14385.17   AICc=-14384.78   BIC=-14263.7
```

C.3 ARMAX$(3, 0, 0)$

```
Series: endo_vars
ARIMA(3,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3  intercept     Plow      AT   Precip  Traffic
      1.4078  -0.1053  -0.3128     0.3236  -0.0001  0.1926  -0.0225   0.0225
s.e.  0.0194   0.0349   0.0201     0.0250   0.0012  0.0189   0.0059   0.0099
      Salting_L0  Salting_L1  Salting_L2  Salting_L3  Salting_L4  Salting_L5
         -0.0026     -0.0027     -0.0056     -0.0103     -0.0080     -0.0051
s.e.      0.0028      0.0043      0.0059      0.0069      0.0075      0.0075
      Salting_L6  Salting_L7  Salting_L8  Salting_L9      WL       IL      SL
         -0.0026     -0.0015     -0.0007     -0.0016  0.0037  -0.0062  0.0025
s.e.      0.0069      0.0059      0.0043      0.0027  0.0075   0.0060  0.0079

sigma^2 estimated as 0.0001312:  log likelihood=7327.24
```

```
AIC=-14610.48   AICc=-14610.05   BIC=-14483.22
```

C.4 ARMAX$(4, 0, 0)$

```
Series: endo_vars
ARIMA(4,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3      ar4  intercept    Plow      AT   Precip
      1.3483  -0.1243  -0.0448  -0.1914     0.3277  0.0000  0.1889  -0.0245
s.e.  0.0200   0.0344   0.0343   0.0201     0.0215  0.0012  0.0185   0.0059
      Traffic  Salting_L0  Salting_L1  Salting_L2  Salting_L3  Salting_L4
       0.0178     -0.0026     -0.0022     -0.0049     -0.0094     -0.0071
s.e.   0.0092      0.0027      0.0040      0.0052      0.0060      0.0065
      Salting_L5  Salting_L6  Salting_L7  Salting_L8  Salting_L9      WL
         -0.0042     -0.0019     -0.0011      -5e-04     -0.0016  0.0050
s.e.      0.0065      0.0060      0.0052       4e-03      0.0026  0.0077
           IL      SL
      -0.0072  0.0037
s.e.   0.0061  0.0080

sigma^2 estimated as 0.0001264:  log likelihood=7371.89
AIC=-14697.79   AICc=-14697.32   BIC=-14564.74
```

C.5 ARMAX$(5, 0, 0)$

```
Series: endo_vars
ARIMA(5,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3      ar4      ar5  intercept    Plow      AT
      1.3237  -0.1312  -0.0593  -0.0191  -0.1279     0.3260  0.0000  0.1943
s.e.  0.0202   0.0340   0.0340   0.0338   0.0203     0.0195  0.0012  0.0183
       Precip  Traffic  Salting_L0  Salting_L1  Salting_L2  Salting_L3
      -0.0245   0.0158     -0.0027     -0.0023     -0.0048     -0.0095
```

```
s.e.   0.0060   0.0088      0.0027      0.0039      0.0049      0.0056
      Salting_L4  Salting_L5  Salting_L6  Salting_L7  Salting_L8  Salting_L9
         -0.0074     -0.0047     -0.0024     -0.0017     -0.0012     -0.0022
s.e.      0.0060      0.0060      0.0056      0.0049      0.0039      0.0026
          WL       IL      SL
      0.0054  -0.0092  0.0023
s.e.  0.0078   0.0061  0.0080

sigma^2 estimated as 0.0001244:  log likelihood=7391.55
AIC=-14735.09   AICc=-14734.59   BIC=-14596.26
```

C.6 ARMAX$(6, 0, 0)$

```
Series: endo_vars
ARIMA(6,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3      ar4      ar5      ar6  intercept    Plow
      1.3148  -0.1336  -0.0626  -0.0280  -0.0358  -0.0697     0.3239  0.0000
s.e.  0.0204   0.0339   0.0340   0.0338   0.0338   0.0205     0.0185  0.0012
          AT   Precip  Traffic  Salting_L0  Salting_L1  Salting_L2  Salting_L3
      0.1994  -0.0245   0.0146     -0.0027     -0.0022     -0.0048     -0.0095
s.e.  0.0184   0.0060   0.0086      0.0026      0.0039      0.0049      0.0055
      Salting_L4  Salting_L5  Salting_L6  Salting_L7  Salting_L8  Salting_L9
         -0.0074     -0.0047     -0.0025     -0.0019     -0.0014     -0.0021
s.e.      0.0058      0.0058      0.0055      0.0049      0.0039      0.0026
          WL       IL      SL
      0.0058  -0.0085  0.0023
s.e.  0.0077   0.0061  0.0080

sigma^2 estimated as 0.0001238:  log likelihood=7397.34
AIC=-14744.67   AICc=-14744.12   BIC=-14600.06
```

C.7 ARMAX$(7, 0, 0)$

```
Series: endo_vars
ARIMA(7,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3      ar4      ar5      ar6      ar7  intercept
      1.3136  -0.1348  -0.0625  -0.0290  -0.0378  -0.0475  -0.0171     0.3223
s.e.  0.0204   0.0340   0.0340   0.0339   0.0339   0.0339   0.0209     0.0184
        Plow      AT   Precip  Traffic  Salting_L0  Salting_L1  Salting_L2
      0.0000  0.2026  -0.0246   0.0143     -0.0027     -0.0023     -0.0048
s.e.  0.0012  0.0189   0.0060   0.0085      0.0027      0.0039      0.0049
      Salting_L3  Salting_L4  Salting_L5  Salting_L6  Salting_L7  Salting_L8
         -0.0095     -0.0074     -0.0047     -0.0026     -0.0019     -0.0013
s.e.      0.0055      0.0058      0.0058      0.0055      0.0049      0.0039
      Salting_L9      WL       IL      SL
         -0.0021  0.0059  -0.0083  0.0024
s.e.      0.0026  0.0078   0.0062  0.0080

sigma^2 estimated as 0.0001237:  log likelihood=7397.67
AIC=-14743.34   AICc=-14742.75   BIC=-14592.95
```

C.8 ARMAX$(8, 0, 0)$

```
Series: endo_vars
ARIMA(8,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3      ar4      ar5     ar6     ar7      ar8
      1.3129  -0.1367  -0.0637  -0.0299  -0.0400  -0.052  0.0289  -0.0353
s.e.  0.0204   0.0340   0.0340   0.0339   0.0339   0.034  0.0339   0.0204
      intercept    Plow      AT   Precip  Traffic  Salting_L0  Salting_L1
          0.322  0.0000  0.2037  -0.0245   0.0139     -0.0027     -0.0024
s.e.      0.018  0.0012  0.0189   0.0060   0.0085      0.0027      0.0039
      Salting_L2  Salting_L3  Salting_L4  Salting_L5  Salting_L6  Salting_L7
         -0.0050     -0.0096     -0.0075     -0.0048     -0.0027     -0.0019
```

```
s.e.      0.0049      0.0055      0.0058      0.0058      0.0055      0.0049
      Salting_L8  Salting_L9      WL       IL      SL
         -0.0013     -0.0021  0.0057  -0.0081  0.0019
s.e.      0.0039      0.0026  0.0078   0.0062  0.0081

sigma^2 estimated as 0.0001236:  log likelihood=7399.16
AIC=-14744.32   AICc=-14743.69   BIC=-14588.14
```

Appendix D

VARMAX$(1, 0, 0)$ Model for Contaminant Layers

```
VAR Estimation Results:
========================
Endogenous variables: WL, IL, SL
Deterministic variables: const
Sample size: 2890
Log Likelihood: 14206.308
Roots of the characteristic polynomial:
0.9797 0.8903 0.876
Call:
VAR(y = endo_vars, p = 1, type = "const", exogen = exog_vars)


Estimation results for equation WL:
===================================
WL = WL.l1 + IL.l1 + SL.l1 + const + AT + Precip + Salting_L2 + Salting_L3
 + Salting_L5

            Estimate Std. Error t value Pr(>|t|)
WL.l1       0.890253   0.008365 106.429  < 2e-16 ***
IL.l1       0.029202   0.008229   3.549 0.000393 ***
SL.l1      -0.009805   0.003779  -2.594 0.009526 **
```

```
const      -0.008769   0.003522  -2.490 0.012842 *
AT          0.017443   0.005642   3.092 0.002010 **
Precip      0.042038   0.008244   5.099 3.63e-07 ***
Salting_L2  0.014806   0.006208   2.385 0.017147 *
Salting_L3  0.015935   0.006160   2.587 0.009733 **
Salting_L5  0.012904   0.006185   2.086 0.037036 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04511 on 2881 degrees of freedom
Multiple R-Squared: 0.8597,Adjusted R-squared: 0.8593
F-statistic:  1961 on 9 and 2881 DF,  p-value: < 2.2e-16


Estimation results for equation IL:
===================================
IL = IL.l1 + SL.l1 + const + AT + Salting_L1 + Salting_L6 + Salting_L7

            Estimate Std. Error t value Pr(>|t|)
IL.l1       0.875983   0.008771  99.878  < 2e-16 ***
SL.l1       0.027676   0.004007   6.908 6.04e-12 ***
const       0.012106   0.003767   3.214  0.00133 **
AT         -0.012193   0.005889  -2.070  0.03851 *
Salting_L1  0.017582   0.006592   2.667  0.00769 **
Salting_L6 -0.013902   0.006596  -2.108  0.03516 *
Salting_L7 -0.015685   0.006595  -2.378  0.01746 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04841 on 2883 degrees of freedom
Multiple R-Squared: 0.8728,Adjusted R-squared: 0.8725
F-statistic:  2827 on 7 and 2883 DF,  p-value: < 2.2e-16


Estimation results for equation SL:
===================================
```

```
SL = SL.l1 + const + Precip + Traffic

         Estimate Std. Error t value Pr(>|t|)
SL.l1    0.979746   0.003601 272.057  < 2e-16 ***
const    0.004730   0.001937   2.442 0.014657 *
Precip   0.030264   0.008554   3.538 0.000409 ***
Traffic -0.007286   0.003313  -2.199 0.027945 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04903 on 2886 degrees of freedom
Multiple R-Squared: 0.9719,Adjusted R-squared: 0.9718
F-statistic: 2.491e+04 on 4 and 2886 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
           WL         IL         SL
WL  0.0020417 -3.179e-04 -3.752e-04
IL -0.0003179  2.352e-03 -8.885e-05
SL -0.0003752 -8.885e-05  2.416e-03

Correlation matrix of residuals:
        WL       IL       SL
WL  1.0000 -0.14505 -0.16892
IL -0.1450  1.00000 -0.03727
SL -0.1689 -0.03727  1.00000
```

Appendix E

VARMAX$(2, 0, 0)$ Model for Contaminant Layers

```
VAR Estimation Results:
=========================
Endogenous variables: WL, IL, SL
Deterministic variables: const
Sample size: 2889
Log Likelihood: 14279.983
Roots of the characteristic polynomial:
0.967 0.8708 0.8708 0.2275 0.1074     0
Call:
VAR(y = endo_vars, p = 2, type = "const", exogen = exog_vars)


Estimation results for equation WL:
===================================
WL = WL.l1 + IL.l1 + SL.l1 + WL.l2 + const + AT + Precip + Salting_L2 +
Salting_L3 + Salting_L5

            Estimate Std. Error t value Pr(>|t|)
WL.l1       0.972389   0.018544  52.436  < 2e-16 ***
IL.l1       0.027874   0.008201   3.399 0.000686 ***
SL.l1      -0.010088   0.003765  -2.679 0.007419 **
```

```
WL.l2      -0.091578   0.018469  -4.958 7.52e-07 ***
const      -0.009291   0.003510  -2.647 0.008167 **
AT          0.019227   0.005632   3.414 0.000650 ***
Precip      0.042289   0.008212   5.150 2.79e-07 ***
Salting_L2  0.015985   0.006189   2.583 0.009845 **
Salting_L3  0.015380   0.006137   2.506 0.012261 *
Salting_L5  0.014847   0.006173   2.405 0.016232 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04494 on 2879 degrees of freedom
Multiple R-Squared: 0.8609,Adjusted R-squared: 0.8604
F-statistic:  1782 on 10 and 2879 DF,  p-value: < 2.2e-16


Estimation results for equation IL:
===================================
IL = WL.l1 + IL.l1 + WL.l2 + SL.l2 + const + Salting_L1 + Salting_L7

            Estimate Std. Error t value Pr(>|t|)
WL.l1      -0.053732   0.019767  -2.718  0.00660 **
IL.l1       0.876482   0.008733 100.368  < 2e-16 ***
WL.l2       0.043206   0.019829   2.179  0.02942 *
SL.l2       0.028885   0.004009   7.205 7.40e-13 ***
const       0.004888   0.001253   3.902 9.76e-05 ***
Salting_L1  0.017994   0.006556   2.745  0.00609 **
Salting_L7 -0.013356   0.006560  -2.036  0.04184 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04841 on 2882 degrees of freedom
Multiple R-Squared: 0.8728,Adjusted R-squared: 0.8725
F-statistic:  2825 on 7 and 2882 DF,  p-value: < 2.2e-16


Estimation results for equation SL:
```

```
====================================
SL = SL.l1 + SL.l2 + const + AT + Precip

        Estimate Std. Error t value Pr(>|t|)
SL.l1   1.194545   0.018144  65.837  < 2e-16 ***
SL.l2  -0.220004   0.018135 -12.131  < 2e-16 ***
const   0.009507   0.003503   2.714 0.006682 **
AT     -0.012927   0.005727  -2.257 0.024069 *
Precip  0.028774   0.008404   3.424 0.000626 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04786 on 2884 degrees of freedom
Multiple R-Squared: 0.9732,Adjusted R-squared: 0.9732
F-statistic: 2.095e+04 on 5 and 2884 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
           WL         IL         SL
WL  0.0020273 -3.128e-04 -3.569e-04
IL -0.0003128  2.355e-03 -7.958e-05
SL -0.0003569 -7.958e-05  2.303e-03

Correlation matrix of residuals:
        WL       IL       SL
WL  1.0000 -0.14315 -0.16519
IL -0.1431  1.00000 -0.03417
SL -0.1652 -0.03417  1.00000
```

Appendix F

VARMAX$(1, 0, 0)$ Model for Contaminant Layers With Salt Residual

```
VAR Estimation Results:
========================
Endogenous variables: WL, IL, SL
Deterministic variables: const
Sample size: 2890
Log Likelihood: 14204.607
Roots of the characteristic polynomial:
0.9797 0.8919 0.8776
Call:
VAR(y = endo_vars, p = 1, type = "const", exogen = exog_vars)


Estimation results for equation WL:
===================================
WL = WL.l1 + IL.l1 + SL.l1 + const + AT + Precip + SaltRes_L2 + SaltRes_L4
+ SaltRes_L5

            Estimate Std. Error t value Pr(>|t|)
WL.l1       0.891853   0.008397 106.206  < 2e-16 ***
```

```
IL.l1       0.030191   0.008230   3.668 0.000249 ***
SL.l1      -0.009864   0.003778  -2.611 0.009083 **
const      -0.007307   0.003593  -2.034 0.042070 *
AT          0.015839   0.005692   2.783 0.005424 **
Precip      0.041997   0.008242   5.095 3.71e-07 ***
SaltRes_L2  0.016377   0.006009   2.726 0.006459 **
SaltRes_L4 -0.014790   0.006161  -2.400 0.016437 *
SaltRes_L5  0.015097   0.006227   2.424 0.015402 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04511 on 2881 degrees of freedom
Multiple R-Squared: 0.8598,Adjusted R-squared: 0.8593
F-statistic:  1962 on 9 and 2881 DF,  p-value: < 2.2e-16


Estimation results for equation IL:
===================================
IL = IL.l1 + SL.l1 + const + SaltRes_L1 + SaltRes_L6

            Estimate Std. Error t value Pr(>|t|)
IL.l1       0.877582   0.008693 100.955  < 2e-16 ***
SL.l1       0.029347   0.003925   7.478 9.99e-14 ***
const       0.004176   0.001178   3.544 0.000401 ***
SaltRes_L1  0.019436   0.006298   3.086 0.002047 **
SaltRes_L6 -0.015682   0.006291  -2.493 0.012727 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04843 on 2885 degrees of freedom
Multiple R-Squared: 0.8726,Adjusted R-squared: 0.8724
F-statistic:  3952 on 5 and 2885 DF,  p-value: < 2.2e-16


Estimation results for equation SL:
===================================
```

```
SL = SL.l1 + const + Precip + Traffic

         Estimate Std. Error t value Pr(>|t|)
SL.l1    0.979746   0.003601 272.057  < 2e-16 ***
const    0.004730   0.001937   2.442 0.014657 *
Precip   0.030264   0.008554   3.538 0.000409 ***
Traffic -0.007286   0.003313  -2.199 0.027945 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04903 on 2886 degrees of freedom
Multiple R-Squared: 0.9719,Adjusted R-squared: 0.9718
F-statistic: 2.491e+04 on 4 and 2886 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
           WL         IL         SL
WL  0.0020408 -3.204e-04 -3.756e-04
IL -0.0003204  2.356e-03 -8.537e-05
SL -0.0003756 -8.537e-05  2.416e-03

Correlation matrix of residuals:
        WL       IL       SL
WL  1.0000 -0.14609 -0.16913
IL -0.1461  1.00000 -0.03578
SL -0.1691 -0.03578  1.00000
```

Appendix G

VARMAX$(1, 0, 0)$ Model with Interaction Effect

```
VAR Estimation Results:
========================
Endogenous variables: WL, IL, SL
Deterministic variables: const
Sample size: 2890
Log Likelihood: 14213.274
Roots of the characteristic polynomial:
0.9769 0.8919 0.8744
Call:
VAR(y = endo_vars, p = 1, type = "const", exogen = exog_vars)


Estimation results for equation WL:
===================================
WL = WL.l1 + IL.l1 + SL.l1 + AT + Precip + SaltRes_L2 + SaltRes_L4
 + SaltRes_L5 + TempRes_L3

            Estimate Std. Error t value Pr(>|t|)
WL.l1       0.891878   0.008385 106.371  < 2e-16 ***
IL.l1       0.030509   0.008193   3.724  0.00020 ***
SL.l1      -0.009394   0.003782  -2.484  0.01304 *
```

```
AT          0.018537   0.005802   3.195  0.00141 **
Precip      0.041792   0.008240   5.072 4.19e-07 ***
SaltRes_L2  0.015235   0.005933   2.568  0.01028 *
SaltRes_L4 -0.015963   0.006095  -2.619  0.00886 **
SaltRes_L5  0.015447   0.006225   2.481  0.01314 *
TempRes_L3 -0.010237   0.004122  -2.484  0.01306 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04509 on 2881 degrees of freedom
Multiple R-Squared: 0.8599,Adjusted R-squared: 0.8594
F-statistic:  1964 on 9 and 2881 DF,  p-value: < 2.2e-16


Estimation results for equation IL:
===================================
IL = IL.l1 + SL.l1 + AT + SaltRes_L1 + SaltRes_L5 + SaltRes_L6 + TempRes_L5

            Estimate Std. Error t value Pr(>|t|)
IL.l1       0.874438   0.008787  99.515  < 2e-16 ***
SL.l1       0.027450   0.004022   6.825 1.07e-11 ***
AT         -0.012857   0.006268  -2.051  0.04035 *
SaltRes_L1  0.016929   0.006381   2.653  0.00801 **
SaltRes_L5  0.014604   0.006735   2.168  0.03022 *
SaltRes_L6 -0.019450   0.006605  -2.945  0.00326 **
TempRes_L5  0.013218   0.004515   2.928  0.00344 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.0484 on 2883 degrees of freedom
Multiple R-Squared: 0.8729,Adjusted R-squared: 0.8726
F-statistic:  2828 on 7 and 2883 DF,  p-value: < 2.2e-16


Estimation results for equation SL:
===================================
```

```
SL = SL.l1 + AT + Precip + Traffic + TempRes_L1

            Estimate Std. Error t value Pr(>|t|)
SL.l1       0.976854   0.003724 262.317  < 2e-16 ***
AT         -0.015235   0.006151  -2.477 0.013306 *
Precip      0.032048   0.008604   3.725 0.000199 ***
Traffic    -0.007467   0.003296  -2.265 0.023557 *
TempRes_L1  0.016018   0.004546   3.523 0.000433 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.04898 on 2885 degrees of freedom
Multiple R-Squared: 0.9719,Adjusted R-squared: 0.9719
F-statistic: 1.998e+04 on 5 and 2885 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
           WL         IL         SL
WL  0.0020465 -3.217e-04 -3.771e-04
IL -0.0003217  2.360e-03 -8.807e-05
SL -0.0003771 -8.807e-05  2.418e-03

Correlation matrix of residuals:
        WL       IL       SL
WL  1.0000 -0.14638 -0.16954
IL -0.1464  1.00000 -0.03687
SL -0.1695 -0.03687  1.00000
```

Appendix H

Multi-Step Forecasting MAE of RSC Models

H.1 Surface Temperature

	ForecastHorizon	MAE	Dataset
1	0	0.0000000	calibration
2	20	0.1678591	calibration
3	40	0.2624593	calibration
4	60	0.3466855	calibration
5	80	0.4301734	calibration
6	100	0.5127732	calibration
7	120	0.6023059	calibration
8	140	0.7004046	calibration
9	160	0.8159428	calibration
10	180	0.9479579	calibration
11	0	0.0000000	validation
12	20	0.2360460	validation
13	40	0.4043886	validation
14	60	0.5648082	validation
15	80	0.7203856	validation
16	100	0.8839116	validation
17	120	1.0509589	validation
18	140	1.2223720	validation

19	160	1.4109004	validation
20	180	1.6196865	validation

H.2 Friction Level

	ForecastHorizon	MAE	Dataset
1	0	0.00000000	calibration
2	20	0.02556735	calibration
3	40	0.04099866	calibration
4	60	0.05261701	calibration
5	80	0.06143010	calibration
6	100	0.06884161	calibration
7	120	0.07530223	calibration
8	140	0.08091218	calibration
9	160	0.08610781	calibration
10	180	0.09029798	calibration
11	0	0.00000000	validation
12	20	0.02369263	validation
13	40	0.03829639	validation
14	60	0.04860900	validation
15	80	0.05619228	validation
16	100	0.06266128	validation
17	120	0.06796160	validation
18	140	0.07289253	validation
19	160	0.07817058	validation
20	180	0.08212058	validation

H.3 Contaminant Layer Depths

H.3.1 Water Layer

	ForecastHorizon	MAE	Dataset
1	0	0.00000000	Calibration
2	20	0.04831811	Calibration
3	40	0.09069000	Calibration

4	60	0.12788749	Calibration
5	80	0.16057673	Calibration
6	100	0.18934859	Calibration
7	120	0.21469402	Calibration
8	140	0.23703425	Calibration
9	160	0.25674917	Calibration
10	180	0.27421109	Calibration
11	0	0.00000000	Validation
12	20	0.04816841	Validation
13	40	0.08987051	Validation
14	60	0.12597640	Validation
15	80	0.15728750	Validation
16	100	0.18446759	Validation
17	120	0.20808177	Validation
18	140	0.22866300	Validation
19	160	0.24654359	Validation
20	180	0.26215783	Validation

H.3.2 Ice Layer

	ForecastHorizon	MAE	Dataset
1	0	0.00000000	Calibration
2	20	0.01114412	Calibration
3	40	0.02090849	Calibration
4	60	0.02946494	Calibration
5	80	0.03696389	Calibration
6	100	0.04353738	Calibration
7	120	0.04930138	Calibration
8	140	0.05435855	Calibration
9	160	0.05879816	Calibration
10	180	0.06269608	Calibration
11	0	0.00000000	Validation
12	20	0.01343640	Validation
13	40	0.02522161	Validation
14	60	0.03556038	Validation
15	80	0.04463192	Validation
16	100	0.05259296	Validation

17	120	0.05958095	Validation
18	140	0.06571627	Validation
19	160	0.07110438	Validation
20	180	0.07583766	Validation

H.3.3 Snow Layer

	ForecastHorizon	MAE	Dataset
1	0	0.00000000	Calibration
2	20	0.02204404	Calibration
3	40	0.04364159	Calibration
4	60	0.06480171	Calibration
5	80	0.08553326	Calibration
6	100	0.10584490	Calibration
7	120	0.12574516	Calibration
8	140	0.14524235	Calibration
9	160	0.16434465	Calibration
10	180	0.18306005	Calibration
11	0	0.00000000	Validation
12	20	0.01454050	Validation
13	40	0.02878650	Validation
14	60	0.04274396	Validation
15	80	0.05641873	Validation
16	100	0.06981653	Validation
17	120	0.08294296	Validation
18	140	0.09580354	Validation
19	160	0.10840364	Validation
20	180	0.12074853	Validation

References

Akaike, H. (1974, December 06). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, *19*(6), 716–723. Retrieved from `http://dx.doi.org/10.1109/TAC.1974.1100705` doi: 10.1109/TAC.1974.1100705

Al-Qadi, I., Loulizi, A., Flintsch, G., Roosevelt, D., Decker, R., Wambold, J., & Nixon, W. (2002). *Feasibility of Using Friction Indicators to Improve Winter Maintenance Operations and Mobility* (Tech. Rep.). (NCHRP Web Document 53 NCHRP Project 6-14 Contractor Final Report)

Andresen, A., & Wambold, J. (1998). Modified Pavement Friction Models for Winter. In *Roadsurf Symposium.*

Andrey, J., Brenning, A., Kirchhoff, D., Mills, B., & Perchanok, M. (2008). Modeling Winter Maintenance Activities Using Classification Trees. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board.

Andrle, S., Kroeger, D., & Sinhaa, R. (2003). *Deploying The Maintenance Decision Supoort System (MDSS) In Iowa* (Tech. Rep.). Center for Transportation Research and Education Iowa, Department of Transportation. (CTRE 02-129 Final Report CTRE Project 02-129 Transportation Research Board)

Antvik, G. (1997). *History of Friction Measurements at Airports* (Tech. Rep.).

Arsenault, S. (2007). Development of a Decision Support System for Winter Maintenance. In *2007 Annual Conference and Exhibition of the Transportation Association of Canada: Transportation - An Economic Enabler (Les Transports: Un Levier Economique).* Transportation Association of Canada. (French)

Askelson, M. (2008). The Pavement Precipitation Accumulation Estimation System. In *Fourth National Conference on Surface Transportation Weather; Seventh International*

Symposium on Snow Removal and Ice Control Technology. Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-017)

Axelson, L. (1994). Automatic Measurement Of Road Status. In *Ninth PIARC International Winter Road Congress.* Bundesministerium fur Wirtschaftliche Angelegenheiten.

Axelson, L. (1995). Winter Road Maintenance System. In *Maintenance Management,Transportation Research Board Conference Proceedings 5.* Transportation Research Board.

Axelson, L., & Malmberg, Ã. (1991). *Expert-System Based Road Salting, A Preliminary Study and A Prototype* (Tech. Rep.). Swedish National Road Administration.

Bachmann, T. (1998). *Wechselwirkungen im Prozess Der Reibung Zwischen Reifen und Fahrbahn.* Unpublished doctoral dissertation, Fahrzeugtechnic tu Darmstadt (FZD).

Ben-Akiva, M., & Lerman, S. (1985). *Discrete Choice Analysis.* The MIT Press.

Blomqvist, G., & Gustafsson, M. (2004). Patterns fo Residual Salt on Road Surface: Case Study. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Box, G. E. P., & Jenkins, G. (1990). *Time Series Analysis, Forecasting and Control.* Holden-Day, Incorporated.

Breiman, L. (1984). *Classification and regression trees.* Wadsworth International Group. Retrieved from `http://books.google.ca/books?id=uxPvAAAAMAAJ`

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

Brown, B., & Baass, K. (1997). Seasonal Variation in Frequencies and Rates of Highway Accidents as Function of Sereciry. *Transportation Research Record*(1581), 59-65.

Buchanan, F., & Gwartz, S. (2005). Road Weather Information Systems at the Ministry of Transportation, Ontario. In *2005 Annual Conference of the Transportation Association of Canada.* Transportation Association of Canada.

Burtwell, M. (2001). Assessment of the Performance of Prewetted Salt for Snow Removal and Ice Control. *Transportation Research Record*(1741), 68-74.

Canadian Strategic Highway Research Program. (2000). *Anti-Icing And RWIS Technology in Canada* (Tech. Rep.). C-SHRP.

Carmichael, C., Gallus, W., Temeyer, B., & Bryden, M. (2004). A Winter Weather Index for Estimating Winter Roadway Maintenance Costs in the Midwest. *Journal of Applied Meteorology*, *43*, 1783-1790.

Chambers, J., & Hastie, T. (1992). *Statistical models in S.* Wadsworth & Brooks/Cole Advanced Books & Software.

Cluett, C., & Jenq, J. (2007). *A Case Study of the Maintenance Decision Support System (MDSS) in Maine* (Tech. Rep.). Battelle Seattle Research Center Federal Highway Administration. (FHWA-JPO-08-001 DTFH61-02-C-00134)

Comford, D., & Thomes, J. (1996). A Comparison Between Special Winter Indices and Expenditure on Winter Road Maintenance in Scotland. *International Journal of Climatology*, *16*, 339-357.

Comfort, G. (1994). *Laboratory Tests of the Performance Of Highway De-icing Chemicals And Winter Sand On Compacted Snow and Ice* (Tech. Rep.). Fleet Technology Ltd.

Comfort, G., & Dinovitzer, A. (1996). *Field Tests of Winter Sands on Packed Snow and Bare Ice Surfaces* (Tech. Rep.). Fleet Technology Ltd.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. In *Ieee transactions on information theory.*

Cowe Falls, L., Jurgens, R., & Chan, J. (2008). Effect of Winter Events on Highway Performance in the Province of Alberta. In *Transportation research board 87th annual meeting.*

Crevier, L.-P., & Delage, Y. (2001). METRo: A New Model for Road-Condition Forecasting in Canada. *Journal of Applied Meteorology*, *40*, 2026-2037.

Datla, S., & Sharma, S. (2008). Association of Highway Traffic Volumes with Cold and Snow and Their Interactions. In *Transportation Research Board 87th Annual Meeting.* Transportation Research Board.

Decker, R., Bignell, J., Lambertsen, C., & Porter, K. (2001). Measuring Efficientcy of Winter Maintenance Practices. *Transportation Research Record*(1741).

Faraway, J. J. (2002). *Practical Regression and Anova using R.* Retrieved from http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf

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

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

Finnish National Road Administration. (2001). *Winter Maintenance Policy in Finland* (Tech. Rep.). FNRA.

FISHER, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*(2), 179–188. Retrieved from http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x doi: 10.1111/j.1469-1809.1936.tb02137.x

Fu, L., Feng, F., & Perchanok, M. (2008). Probabilistic Models for Discriminating Road Surface Conditions Based on Friction Measurements. In *Transportation research board 87th annual meeting.* Transportation Research Board.

Fu, L., & Perchanok, M. (2006). Effects of Winter Weather and Maintenance Treatments on Highway Safety. In *Transportation research board 85th annual meeting.* Transportation Research Board.

Fu, L., Sooklall, R., & Perchanok, M. (2006). *Effectiveness of Alternative Chemicals for Snow Removal on Highways* (No. 1948). Transportation Research Board.

Fukuzawa, Y., Kajiya, Y., & Kobayashi, T. (1997). Development of Road Freezing Forecast Method. In *52nd Annual Conference of Japan Society of Civil Engineers.*

Gini, C. W. (1971). Variability and Mutability, Contribution to the Study of Statistical Distributions and Relations, Studi Economico-Giuridici della R. Universita de Cagliari (1912). Reviewed in: Light, RJ, Margolin, BH: An Analysis of Variance for Categorical Data. *J. Amer. Stat. Assoc*, *66*.

Gnanadesikan, R. (1997). *Methods for statistical data analysis of multivariate observations.* Wiley. Retrieved from http://books.google.ca/books?id=FtcSzLqeFbQC

Goodwin, L. (2002). *Analysis of Weather-Related Crashes on U.S. Highways* (Tech. Rep.). Mitrerek Systems.

Gustavsson, T., & Bogren, J. (2012). BiFi - Bearing information through vehicle intelligence. In *SIRWEC 2012 Program.* The Standing International Road Weather Commission.

Hanbali, R. (1994). Economic Impact of Winter Road Maintenance on Road Users. *Transportation Research Record*(1442).

Hanbali, R., & Kuemmel, D. (1993). Traffic Volume Deductions Due to Winter Storm Conditions. *Transportation Research Record*(1387), 159-164.

Hastie, T., & Tibshirani, R. (1990). *Generalized Additive Models: T. J. Hastie and R.J. Tibshirani.* Chapman & Hall.

Henry, J. (2000). *Evaluation of Pavement Friction Characteristics (NCHRP synthesis 291)* (Tech. Rep.). Transportation Research Board.

Hodgins, B., Dinning, M., & Wismer, D. (2011). *Winter Road Condition Terminology User Guide* (Tech. Rep.). Transportation Association of Canada.

Hulme, M. (1982). A New Winter Index and Geographical Variations in Winter Weather. *Journal of Meteorology*, *7*, 294-330.

Hunt, C., Mitchell, G., & Richardson, W. (2004). Field Persistence of Anti-Icing Sodium Chloride Residuals. In *Sixth International Symposium on Snow Removal and Ice Control Technology.*

Hyndman, R. J., & Khandakar, Y. (2008, 7 29). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, *27*(3), 1–22. Retrieved from `http://www.jstatsoft.org/v27/i03`

ITS/Win Research Program. (2008). *Road Surface Classifications and the Icy Road Forecasting System* (Tech. Rep.). Civil Engineering Research Institute of Hokkaido.

Jenkins, G., & Watts, D. (1968). *Spectral Analysis And Its Applications.* Holden-Day.

Johnson, R., & Wichern, D. (2007). *Applied Multivariate Statistical Analysis.* Pearson Prentice Hall.

Jones, B. (2003). Road Maintenance Costs in Canada in Winter. *Weather and Transportation in Canada, Department of Geography Publication Series*(55).

Kantardzic, M. (2002). *Data mining: Concepts, models, methods and algorithms*. New York, NY, USA: John Wiley & Sons, Inc.

Karray, F., & Silva, C. (2004). *Soft Computing and Intelligent Systems Design*. Addison Wesley.

Katko, K. (1993). Goals and Methods of Winter Maintenance in Finland. *Transportation Research Record*(1387).

Kawai, S., Furukane, T., Shibata, K., & Horita, Y. (2012, jan.). Automatic distinction of road surface conditions in road images at night-time using pca and mahalanobis weighting. In *Consumer electronics (icce), 2012 ieee international conference on* (p. 231 -232). doi: 10.1109/ICCE.2012.6161842

Kecman, V. (2001). *Learning and soft computing: Support vector machines, neural networks, and fuzzy logic models*. Mit Press. Retrieved from `http://books.google.ca/books?id=W5SAhUqBVYoC`

Ketcham, S., Minsk, L., Blackburn, R., & Fleege, E. (1996). *Manual of Practice for An Effective Anti-Icing Program: A Guild for Highway Winter Maintenance Personnel.* Cold Regions Research and Engineering Laboratory (Federal Highway Administration). (The Office of Engineering, Office of Engineering RD, and Office of Technology Applications sponsored and managed this project as part of the Federal Highway Administration's TE 28 Anti-Icing Technology Program. FHWA-RD-95-202 DTFH61-93-Y-00123)

Klein-Paste, A. (2008). Physical Processes That Affect Runway Surface Conditions During Winter Time: A Conceptual Model. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-014)

Knapp, K., & Smithson, L. (2000). Winter Storm Event Volume Impact Analysis Using Multiple-Source Archived Monitoring Data. *Transportation Research Record*(1700).

Kummer, H. (1996). *Unified Theory of Rubber and Tire Friction* (Tech. Rep.). The Pennsylvania State University, Department of Mechanical Engineering.

Lee, C., Loh, W., Qin, X., & Sproul, M. (2008). Development of New Performance Measure for Winter Maintenance Using Vehicle Speed Data. In *Transportation Research Board 87th Annual Meeting.* Transportation Research Board.

Lefebvre, M. (2007). *Applied Stochastic Processes*. Springer.

Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*. Springer-Verlag.

Lysbakken, K., & Norem, H. (2008). The Amount of Salt on Road Surfaces After Salt Application: A Discussion of Mechanisms and Parameters. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.*

Mahoney III, W., Bernstein, B., Wolff, J., Linden, S., Myers, W., Hallowell, R., ... Burkheimer, D. (2005). FHWA's Maintenance Decision Support System Project: Results and Recommendations. *Transportation Research Record*.

Marchetti, M., Duval, R., Livet, J., Francois, M., Pecquenard, P., & Guillevic, P. (2008). Numerical Evolution of Deicer Amount on Pavement Surface: Traffic and Weather Influences. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.*

McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. Chapman and Hall. Retrieved from `http://books.google.ca/books?id=h9kFH2_FfBkC`

Mc-Cullagh, P., & Nelder, J. (1999). *Generalized Linear Models* (Second ed.). Chapman & Hall/CRC.

McCullouch, B., Belter, D., Konieczny, T., & McClellan, T. (2004). Indiana Winter Severity Index. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Mechler, H., Groãÿ, J., & Schmitz-Hãbsch, A. (2000). *IRS-20/21 (Intelligent Road Sensor)* (Tech. Rep.). Lufft.

Mickens, R. (1991). *Difference Equations* (Second ed.). Taylor & Francis.

Ministry of Transportation Onatrio. (2003). *Maintenance Best Practice - Summary, MBP-701* (Tech. Rep.). MTO.

Ministry of Transportation Ontario. (1993 to 2009). *Ontario Road Safety Annual Reports* (Tech. Rep.). MTO.

Ministry of Transportation Ontario. (2003). *Maitenance Management Information System* (Tech. Rep.). MTO.

Ministry of Transportation Ontario. (2007). *Road and Weather Information Sheet: Ottawa Area Office.*

Minsk, L. (1998). *Snow and Ice Control Manual for Transportation Facilities.* McGraw-Hill Companies, Incorporated.

Möller, S. (2008). Winter Model: Road Condition Submodel. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-011)

Morin, D., & Perchanok, M. (2003). *Road Salt Use in Canada (In: Weaterh and Transportaiton in Canada)* (Tech. Rep.). University of Waterloo, Canada. (Publication of University of Waterloo, Canada, 55)

NCHRP. (2004). *Snow and Ice Control: Guidelines for Material and Methods* (Tech. Rep.). Transportation Research Board of the National Academies.

Nixon, W. (1998a). *Friction as A Tool for Winter Maintenance.* Iowa State University.

Nixon, W. (1998b). *The Potential Of Friction As A Tool For Winter Maintenance* (Tech. Rep.). University of Iowa, Iowa City Iowa Department of Transportation. (IDOT Project TR-400 IIHR Tech Rept No. 392 Final Report)

Nixon, W., & Qiu, L. (2005). Developing a Storm Severity Index. *Transportation Research Record*(1911).

Nordstrom, O. (1998). Development and Validation of BV14, A New Twin Track Fixed Slip Friction Tester for Winter Road Maintenance Monitoring in Sweden. In *Xth PIARC International Winter Road Congress* (Vol. 3). PTRC Education and Research Services Limited.

Norrman, J. (2000). Slipperiness on Roads - An Expert System Classification. *Meteorological Applications*, *7*(1), 27.

Norwegian Ministry of Transport and Communication. (2003). *Standard for Maintenance and Operations, Handbook111* (Tech. Rep.). NMTC.

Öberg, G., & Gregersen, N. (1991). MINSALT : an experiment with unsalted road in Västerbotten county. *VTI meddelande 636*.

Omer, R., & Fu, L. (2010, sept.). An automatic image recognition system for winter road surface condition classification. In *Intelligent transportation systems (itsc), 2010 13th international ieee conference on* (p. 1375 -1379). doi: 10.1109/ITSC.2010.5625290

Perchanok, M. (1994). *Evaluation of a Video System for Remote Monitoring of Winter Road Surface Conditions* (Tech. Rep.). Ministry of Transportation Ontario.

Perchanok, M. (1998). Friction Based Quality Assurance Measures for Snow Clearing Operations. In *Xth PIARC International Winter Road Congress* (Vol. 3). PTRC Education and Research Services Limited. (Transportation Research Board)

Perchanok, M. (2002a). *Decision Support System for Winter Maintenance: The DART Database* (Tech. Rep.). Ministry of Transportation Ontario.

Perchanok, M. (2002b). Patchiness of Snow Cover and Its Relation to Quality Assurance in Winter Operations. In *New Challenges for Winter Road Service. XIth International Winter Road Congress.* World Road Association - PIARC.

Perchanok, M. (2007). *Variation of Snow Cover and Extrapolation of RWIS Data Along a Highway Maintenance Route* (Tech. Rep.). Ministry of Transportation Ontario.

Perchanok, M. (2008). An Approach to Terrain Classification to Improve Road Condition Forecasts of Maintenance Decision Support Systems. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-039)

Perchanok, M., Manning, D., & ARMSTRONG, J. (1991). *Highway De-Icers: Standards, Practice, and Research in the Province of Ontario* (Tech. Rep.). Ministry of Transportation Ontario.

Perchanok, M., Raven, R., Stephenson, D., & Foreman, S. (1995). *Field Trials of Prewetted Sand for Winter Maintenance* (Tech. Rep.). Ministry of Transportation Ontario.

Petty, K., Mahoney III, W., Cowie, J., Dumont, A., & Myers, W. (2008). Providing Winter Road Maintenance Guidance: An Update of the Federal Highway Administration Maintenance Decision Support System. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-019)

Pfaff, B. (2008, 7 29). Var, svar and svec models: Implementation within r package vars. *Journal of Statistical Software*, *27*(4), 1–32. Retrieved from `http://www.jstatsoft.org/v27/i04`

Pilli-Sihvola, Y., & Lahesmaa, J. (2004). Projected state of road weather monitoring in Finland in 2007. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Pilli-Sihvola, Y., Toivonen, K., Haavasoja, T., Haavisto, V., & Nylander, P. (2006). New Approach to Road Weather: Measuring Slipperiness. In *SIRWEC.*

Pisano, P., Stern, A., Mahoney, W., Myers, W., & Burkheimer, D. (2004). Winter Road Maintenance Decision Support System Project: Overview and Status. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Qiu, L., & Nixon, W. (2008a). Effects of Adverse Weather on Traffic Crashes: Systematic Review and Meta-Analysis. In *Transportation Research Board 87th Annual Meeting.* Transportation Research Board. (08-2320)

Qiu, L., & Nixon, W. (2008b). Modeling the Causal Relationships Between Winter Highway Maintenance, Adverse Weather and Mobility. In *Transportation Research Board 87th Annual Meeting.* Transportation Research Board. (08-3101)

R Core Team. (2012). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/` (ISBN 3-900051-07-0)

Reinsel, G. (2003). *Elements of Multivariate Time Series Analysis.* Springer.

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

Saarikivi, P. (2012). Development of mobile optical remote road condition monitoring in Finland. In *SIRWEC 2012 Program.* The Standing International Road Weather Commission.

Sandberg, U. (1997). *Influence of Road Surface Texture on Traffic Characteristics Related to Environment, Economy and Safety* (Tech. Rep.). Swedish National Road and Transport Research Institute.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. Retrieved from `http://dx.doi.org/10.2307/2958889` doi: 10.2307/2958889

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell Systems Technical Journal*, *30*, 50–64.

Shao, J., & Jones, S. (2012). Area Forecast Model for Winter Road Maintenance over a Road Network. In *SIRWEC 2012 Program.* The Standing International Road Weather Commission.

Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and Its Applications With R Examples*. Springer. (ISBN 978-0-387-29317-2)

Sooklall, R., Fu, L., & Perchanok, M. (2006). Effectiveness of Pre-wetting Strategy for Snow and Ice Control on Highways. In *Annual Conference & Exhibition of the Transportation Association of Canada, 2006. Congres et exposition annuels de l'Association des transport du Canada, 2006.* Transportation Association of Canada.

Suggett, J., Hadayeghi, A., Andrey, J., Mills, B., & Leach, G. (2007). Development of A Winter Severity Index for Salt Management in Canada. In *Transportation Research Board 86th Annual Meeting.* Transportation Research Board. (07-2426)

Suggett, J., Hadayeghi, A., Mills, B., Andrey, J., & Leach, G. (2007). *Development of Winter Severity Indicator Models for Canadian Winter Road Maintenance* (Tech. Rep.). Transportation Association of Canada.

Takahashi, N., Tokunaga, R., & Asano, M. (2008). Development of Winter Maintenance Support System: Overview and Status. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-022)

The National Center for Atmospheric Research. (2005). *MDSS Prototype Release-5 Technical Description* (Tech. Rep.). The National Center for Atmospheric Research.

The National Center for Atmospheric Research. (2008). *MDSS Prototype Release-5.0 Technical Description Version 1.1* (Tech. Rep.). Federal Highway Administration (FHWA) Road Weather Management Program.

Thornes, J. (1993). Cost-Effective Snow and Ice Control For The 1990s. *Transportation Research Record*(1387).

Tiao, G., & Tsay, R. (1990). Model specification in multivariate time series (with discussion). *Journal of the Royal Statistical Society*.

Toivonen, K., & Kantonen, J. (2004). Finnishi Road Administration's Web Road Weather Project. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Tokunaga, R., Funahashi, M., Takahashi, N., & Asano, M. (2008). A Feasibility Study on Friction for Winter Road Management. In *Fourth National Conference on Surface Transportation Weather; Seventh International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C126 Snow08-020)

Transportation Association of Canada. (1999). *Salt Management Guide* (Tech. Rep.). TAC.

Transportation Association of Canada. (2008). *Road Salt and Snow And Ice Control Primer* (Tech. Rep.). TAC.

Transportation Research Board. (2000). *Highway Capacity Manual.*

Usman, T., Fu, L., & Miranda-Moreno, L. F. (2010). Quantifying safety benefit of winter road maintenance: Accident frequency modeling. *Accident Analysis Prevention*, *42*(6), 1878 - 1887. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0001457510001442` doi: 10.1016/j.aap.2010.05.008

Vaa, T. (2001). Methods for Measuring and Reporting Winter Maintenance Activities. *Transportation Research Record, Advances and Issues in Snow-Removal and Ice-Control Technology*(1741).

Vapnik, V. (1999). *The nature of statistical learning theory.* Springer. Retrieved from `http://books.google.ca/books?id=sna9BaxVbj8C`

Venables, W., & Ripley, B. (1999). *Modern applied statistics with s-plus.* Springer. Retrieved from `http://books.google.ca/books?id=JZhzQgAACAAJ`

Venalainen, A. (2001). Estimation of Road Salt Use Based on Winter Air Temperature. *Meteorological Applications*, *8*, 333–338.

Venalainen, A., & Markku, K. (2003). Estimation of WRM Costs Using Climate Data. *Meteorological Applications*, *10*(69-73).

Wallman, C. (2001). Vehicle Speed and Flow in Various Winter Road Conditions. In *Ninth Maintenance Management Conference.* Transportation Research Board. (Transportation Research Board Conference Proceedings 23)

Wallman, C. (2004). THE Winter Model: A Winter Maitenance Management System. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Wallman, C., & Åström, H. (2001). *Friction Measurement Methods and the Correlation between Road Friction and Traffic Safety: A Literature Review* (Tech. Rep.). VTI, the Swedish National Road and Transport Research Institute.

Wallman, C., Peter, W., & Gudrun, O. (1997). *Effects of Winter Road Maintenance* (Tech. Rep.). VTI, the Swedish National Road and Transport Research Institute.

Wambold, J., Antle, C., Henry, J., & Rado, Z. (1995). *International PIARC Experiment to Compare and Harmonize Texture and Skid Resistance Measurements, Final report* (Tech. Rep.). Permanent International Association of Road Congresses (PIARC).

Yamagiwa, Y., Kajiya, Y., & Uemura, T. (2004). Development of An Advanced Winter Road Management Support System for Great Sapporo. In *Sixth International Symposium on Snow Removal and Ice Control Technology.* Transportation Research Board. (Transportation Research E-Circular E-C063 SNOW04-005 SNOW04-039)

Zadeh, L. (1953). Theory of filtering. *Journal of the Society for Industrial and Applied Mathematics*, *1*(1), 35-51.