



Department of Civil & Environmental
Engineering
University of Alberta

Student Paper Competition

Advances in Winter Road Surface Conditions Monitoring and Forecasting via Deep Learning and Geostatistics

By:

Mingjian Wu, M.Sc. Graduate Research Assistant

Department of Civil and Environmental Engineering, University of Alberta

Edmonton, T6G 1H9, Tel: 587.937.7024

Email: mingjian.wu@ualberta.ca

February 26th, 2022

Contents

Preface	3
Acknowledgement	4
1. Introduction.....	5
2. Methodology	6
2.1 The Generic Framework	6
2.2 Regression Kriging	7
2.3 Deep Learning.....	7
3. Case Study	8
3.1 Study Area	8
3.2 Data Collection and Processing	9
3.3 Continuous Mapping of Road Surface Index.....	10
4. Conclusions and Future Recommendations	15
References.....	17

Preface

Work presented in this paper is either accepted, published or is under-review for publication in various journals and conferences in the areas of transportation engineering.

Peer-reviewed Book Chapter (In-Press)

1. **Wu, M.**, Kwon, T. J., Fu, L. & Wong, H.A. (2021). Advances in Sustainable Winter Road Maintenance and Management for Future Smart Cities. Advanced Structural Sensing and Monitoring Systems. *Elsevier* (In-Press)

Peer-reviewed Journal Papers (Published)

2. Gu, L., **Wu, M.**, & Kwon, T. J. (2020). An enhanced spatial statistical method for continuous monitoring of winter road surface conditions. *Canadian Journal of Civil Engineering*, 47(10), 1154-1165.
3. Biswas, S., **Wu, M.**, Kwon, T. J. & Melles, S. J. (2019). Use of Topography, Weather Zones, and Semivariogram Parameters to Optimize Road Weather Information System Station Density across Large Spatial Scales. *Transportation Research Record*, DOI:0361198119846467.

Peer-reviewed Journal Papers (Under-Review)

1. **Wu, M.** and Kwon, T. J., An Automatic Architecture Designing Approach of Convolutional Neural Networks for Road Surface Conditions Image Recognition: Tradeoff between Accuracy and Efficiency. Submitted to *Journal of Sensors*. Jul 2021.
2. **Wu, M.**, Kwon, T. J. and Fu, L., Spatial Mapping of Winter Road Surface Conditions via Hybrid Geostatistical Techniques. Submitted to *Journal of Cold Regions Engineering*. Jul 2020.

Refereed conference papers and presentations

1. **Wu, M.**, Kwon, T. J. and Fu, L. (2021). Continuous Mapping of Winter Road Surface Conditions via Deep Learning and Geostatistics. Accepted for presentation at the 101st Annual Meeting of Transportation Research Board, Washington, D.C., Jan 2022.
2. **Wu, M.** and Kwon, T. J. (2021). An Automatic Architecture Designing Approach of Convolutional Neural Networks for Road Surface Conditions Image Recognition Tradeoff between Accuracy and Efficiency. Accepted for presentation at the 101st Annual Meeting of Transportation Research Board, Washington, D.C., Jan 2022.

Acknowledgement

The presented work was done under the supervision of Dr. Tae J. Kwon Basyouny from University of Alberta and Dr. Liping Fu from University of Waterloo, without their guidance and help, it is impossible to get all these done. I would also like to thank Tina Greenfield of Iowa Department of Transportation (DOT) for providing valuable data essential for this study. I would also like to acknowledge all the help and support from the Aurora Project Team, along with Zach Hans and Neal Hawkins of Iowa State University, and Khyle Clute of Iowa DOT for providing continued support to the project. This research is funded by the Aurora Program and partially funded by Natural Sciences and Engineering Research Council of Canada (NSERC).

1. Introduction

In countries with severe winter seasons, transportation agencies often face significant challenges in meeting the safety and mobility needs of road users. Inclement weather events, such as frequent snow, sleet, ice, and frost events result in high variance in road surface conditions (RSC) over space and time (1,2,3,4,5). Poor driving conditions as a result of inclement weather is something that municipalities must contend with due to it being a safety risk for road users. If proper precautions are not taken by municipalities, winter collisions become a constant problem that threaten the lives of motorists.

Given the extensive records of winter incidences, transportation agencies strive to provide the best condition possible by mobilizing various winter road maintenance (WRM) activities. WRM operations such as plowing, deicing, and sanding operations are essential for mitigating the chance and/or severity of collisions. However, due to the vast spatial distances covered by highway networks and the uncertain nature of weather events, it is difficult to monitor and predict the fluctuating RSC, making both public travel and WRM extremely challenging. To help support and facilitate winter maintenance decisions, an advanced intelligent transportation system (ITS) technology known as the road weather information systems (RWIS), available in both stationary and mobile forms are deployed throughout many road networks around the world. RWIS (shown in Figure 1) is an integrated system that includes environmental sensors for measuring temperatures, humidity, pavement conditions, etc. Using RWIS, highway maintenance authorities can collect real-time and near-future road weather and surface conditions data to improve the efficiency and effectiveness of their maintenance operations. Furthermore, most RWIS nowadays are equipped with cameras that provide users with a direct view of the road segment; however, determining the RSC via these cameras is still being done manually, which prevents the full utilization of these rich image-based road condition data. Despite its many benefits, RWIS's main disadvantage lies in its high installation and operation costs, municipalities are thus only able to deploy them to a limited number of locations, hampering its coverage and effectiveness. Because of the coverage issue, transportation agencies have been looking for ways to infer road weather and surface conditions using spatially scarce RWIS information to maximize their return on investments.



Figure 1 Typical RWIS. Left: Stationary RWIS. Right: Mobile RWIS.

Although some prior studies helped provide insight into how RSC varies over space and its estimation process, they suffer from one major limitation - the models were developed to provide only site-specific RSC rather than for an entire segment of road. Having continuous RSC information over a road network is

critical for safety and for effective winter maintenance services that produce a high level of service.

Therefore, the main purpose of this study is to develop a methodological framework that is tailored specifically to continuously map road surface index (RSI) which is a surrogate measure of road surface friction level, by incorporating two concurrent state-of-art techniques - *deep learning* (DL) and *geostatistics* (i.e., kriging). The novelty of the method proposed herein lies in its unique ability to infer RSC conditions between different pairs of existing RWIS stations by incorporating dash camera images and RWIS stationary measurements as sole input – the ability that all highway maintenance agencies have long been looking for to make their current WRM practices more efficient and sustainable.

2. Methodology

2.1 The Generic Framework

As illustrated previously, the primary objective of this study is to develop a methodological framework that combines CNN and kriging for RSI estimation. More specifically, our proposed approach focuses on the following aspects:

- Using CNN to automatically classify images collected from dash cameras into their corresponding RSC category,
- Creating an image processing technique that converts each classified RSC category into an RSI numerical value between 0 and 1, reflecting the slipperiness of the road surface, and this will be detailed in the case study,
- Employing regression kriging (RK), a variant of kriging, to continuously map RSI values for all unmeasured locations along the highway stretch.

A schematic diagram showing this framework is given in Figure 2.

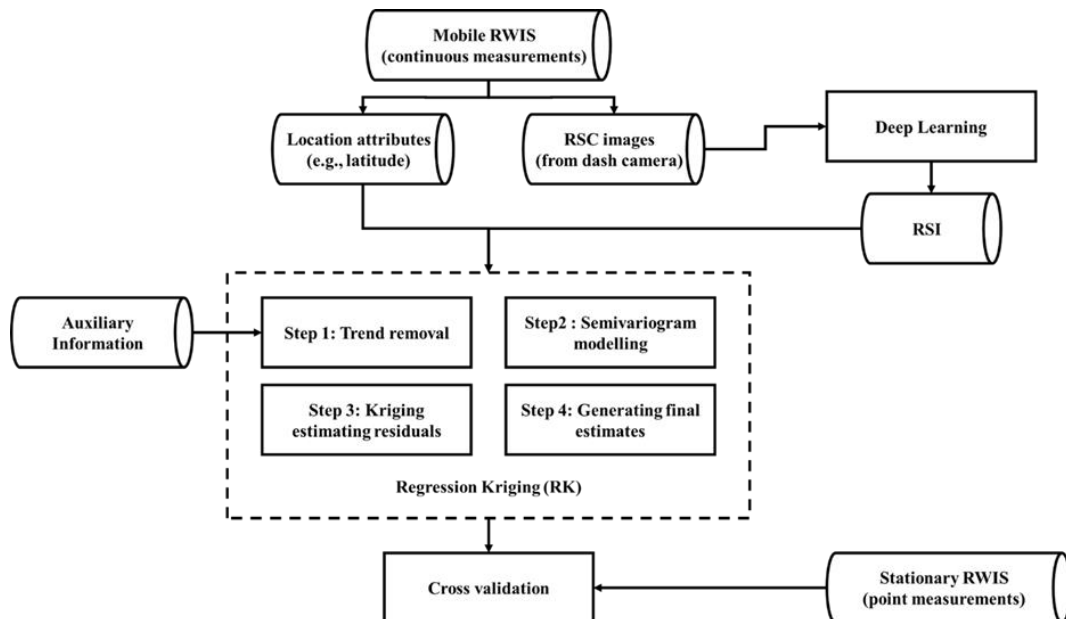


Figure 2 Proposed Framework: Estimating RSI Using RK and DL

2.2 Regression Kriging

Like all other kriging methods, regression kriging (RK) provides the best linear unbiased estimates (BLUE) for variables that tend to vary over space. It combines both deterministic and stochastic components of random variables under investigation, not only does it provide estimates but also estimation uncertainty at unknown locations based on the set of known observations. An example of RK process is illustrated in Figure 3, while its detailed procedures can be found in Figure 2 (Step 1 – Step 2). As can be seen from the figure, the deterministic part of variation is first estimated by a multiple linear regression (MLR) model to remove the possible trend, then the residuals can be interpolated by kriging which characterizes and quantifies the underlying spatial structure of the observed measurements. The estimated residuals are then added back to the regression results to generate the final estimations.

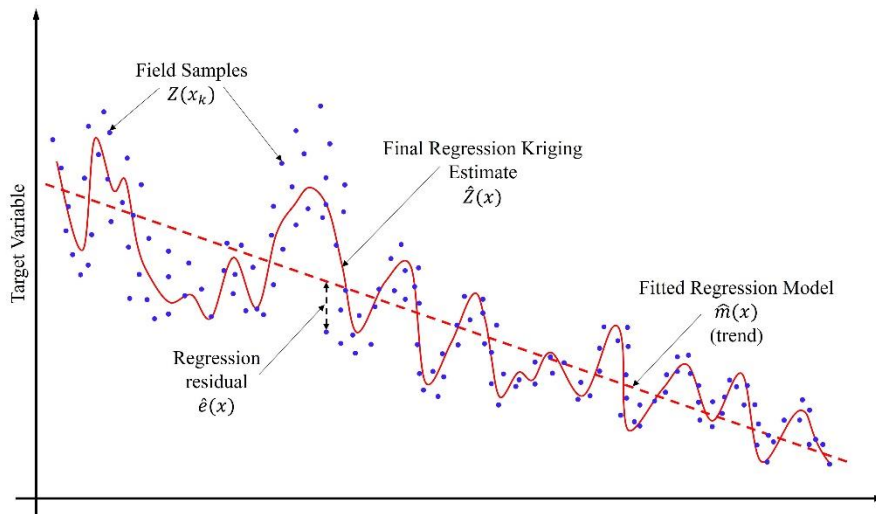


Figure 3 An example of regression kriging process

2.3 Deep Learning

Although RK is effective in estimating values at unmonitored areas, its value to WRM authorities is the highest when real-time point measurements of RSC are obtained in a timely manner, so that it can be used as inputs in the RK technique to provide the most updated and relevant information.

DL models are evolved from a simpler type of model called the multi-layer perceptron, which takes an input vector and uses consecutive groups of non-linear functions called layers to produce higher-level representations of the input data. In contrast, DL models use many layers, where each succeeding layer takes the output of the previous layer as its input. Most DL models fall into the supervised learning category because their goal is to make the model produce a desired output mapping based on the input observations. In our case we expect the DL (i.e., CNN) model to receive images from dash cameras attached to mobile RWIS stations.

Figure 4 shows a simplified CNN architecture used for image classification. The input image is converted into a three-dimensional matrix, which then moves through each of the model's layers until it becomes a vector of probabilities, with the highest probability corresponding to the most likely RSC category for that image.

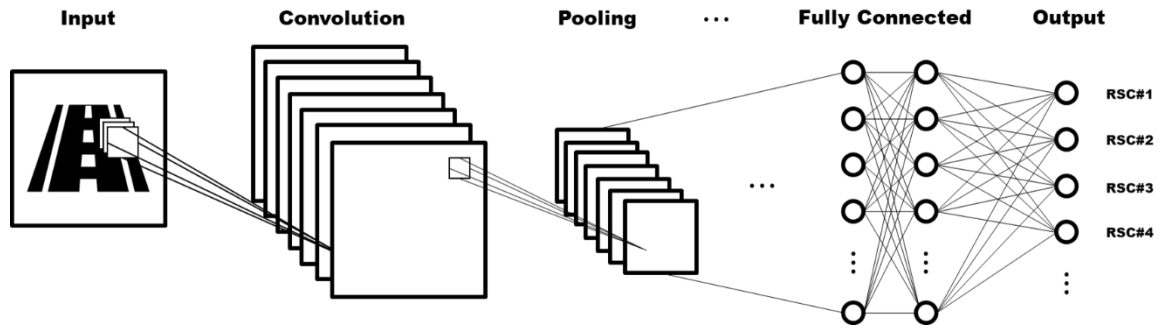


Figure 4 The generic architecture of a CNN for image classification

3. Case Study

To evaluate the feasibility of RK and DL for estimating RSC along highways, this section provides a real-world case study example. Each case study follows the procedures outlined in the methodology section, the descriptions and results associated with each part are also presented here.

3.1 Study Area

Since RSC can be easily influenced by a variety of factors, the selected study area must consider and control any geographical, topographical and traffic related effects. In addition, considering the fact that harsh weather events often affect long stretches of the road, and that WRM activities can be performed synchronously, the spatial coverage also needs to be considered in the study area selection process. Therefore, the southern section of Interstate 35 (I-35) within Iowa, US that is completely paved with asphalt, was selected for this study. This study area can be seen in Figure 5.

The selected road segment is approximately 118 km long. Within this section are three stationary RWIS stations. RSC images were collected by the automated vehicle location (AVL) system, which can be considered as a stand-in version of a heavily simplified mobile RWIS. The other data (e.g., geographical and topographical features, etc.) were obtained from the three stationary RWIS stations along with other data sources.

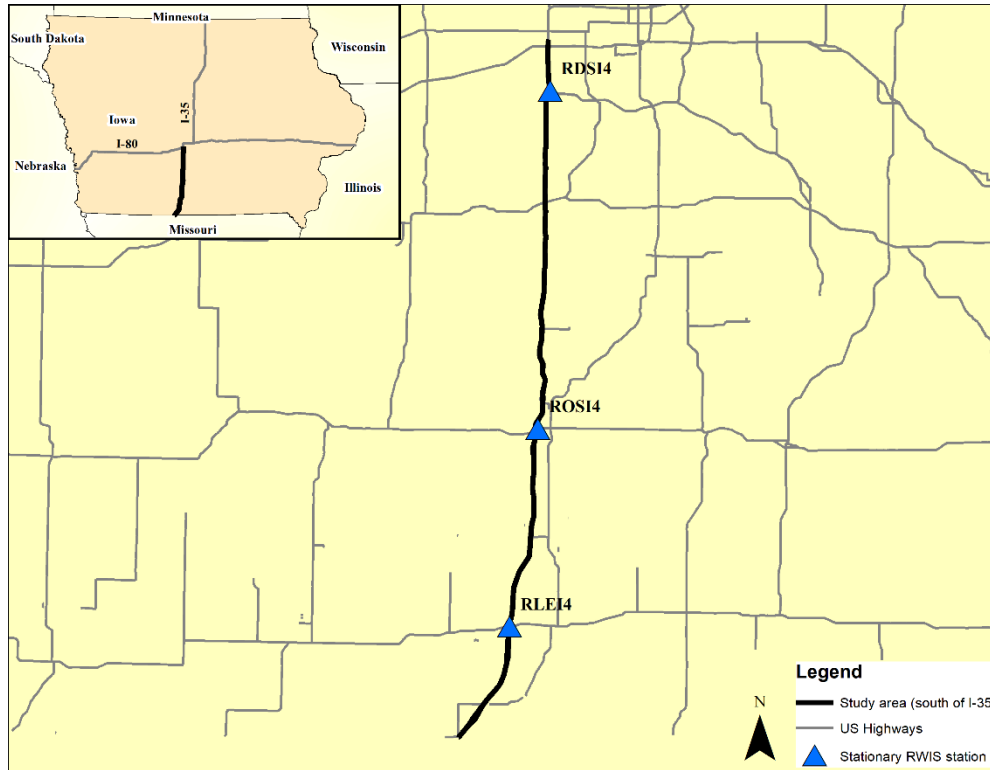





Figure 5 Study area - southern section of I-35 in Iowa

3.2 Data Collection and Processing

AVL makes use of the global positioning system (GPS) to track the locations of their winter maintenance vehicle fleet. And for temperature, a regular thermometer plus an infer-red thermometer to measure the air temperatures and RST every 10 seconds along the roadway. In addition, a standard vehicle-mounted dash camera is utilized to record the RSC images along the roadway typically every 5-10 minutes. These RSC images were manually labelled according to the classification scheme shown in Table 1, which are then used to train the DL model and validate its accuracy. Since RSI values were not directly collected by AVL, RSC categories are converted into RSI values before spatially mapping the RSI for unvisited areas. The conversion method can be found later in this section. Based on the provided dataset, there were 34 hourly observations qualified for the following analysis. RSC categories (labelled RSC images) can be found in Figure 6. Finally, all the data were linked by location and time using ArcGIS, and then aggregated into hourly base data sets for geostatistics modelling.

Table 1 Definition of different RSC categories

Sample Image	RSC Description/Definition	RSC Category
	<p>At least 3 meters of the pavement cross-section in all lanes clear of snow or ice.</p>	<p>Bare Pavement</p>

	<p>Only part of wheel path is clear of snow or ice.</p>	<p>Partially Snow Covered</p>
	<p>No wheel path clear of snow or ice.</p>	<p>Fully Snow Covered</p>
	<p>Not recognizable because of too blurry, too dark or too much light.</p>	<p>Undefined</p>

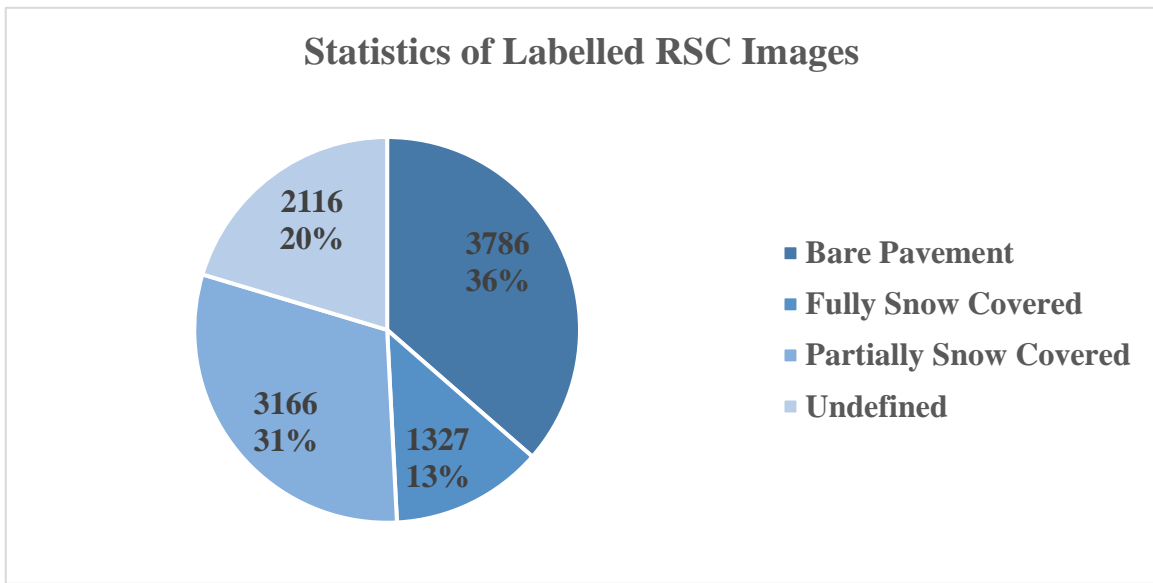


Figure 6 Statistics of Labelled RSC Images

3.3 Continuous Mapping of Road Surface Index

To spatially map RSI, as illustrated in Figure 2, DL and RK were both utilized. Since the RSI values were not directly measured but determined using dash camera images, the DL model was used to automate the process of classifying the RSC, then using an image thresholding technique (to be further discussed in the forthcoming sections), RSI values were calculated based on the DL image classifications. Finally, the RK model was adopted to estimate RSI for those unvisited areas (i.e., locations where RSC images were not available).

a) Road Surface Condition Image Recognition via Deep Learning

Following procedures described in the methodology section, a relatively simple architected CNN model was developed using Tensorflow API (6) and trained on Compute Canada with 32G GPU (7). Table 2 shows the architecture of the developed CNN. Since the main purpose of this study is not to investigate the architecture of CNN, detailed procedures of how this model was developed is not included in this paper. Figure 7 shows the training (90% of all data) and validation (10% of all data) accuracy of the developed model, along with a confusion matrix depicting the validation accuracy of each image classification category. The confusion matrix is a specific table format that visualizes the performance of the CNN model. The number of correct and incorrect predictions are summarized with normalized values (i.e., percentage) and broken down by category. The values in the diagonal line represent the prediction accuracy. For example, in the first row of the confusion matrix, 0.94 means the CNN model correctly classified 94% of unseen bare pavement images into the Bare category, but incorrectly classified 0% (0), 5.3% (0.053) and 0.26% (0.0026) of them into other three categories.

Table 2 Architecture of the developed CNN

Layer Type	Output Size	Number of Parameters
Input Layer	224×224×3	0
Convolutional	224×224×16	448
Max-Pooling	111×111×16	0
Convolutional	109×109×32	4640
Max-Pooling	54×54×32	0
Convolutional	52×52×64	18496
Max-Pooling	26×26×64	0
Convolutional	24×24×128	73856
Max-Pooling	12×12×128	0
Convolutional	10×10×256	295168
Max-Pooling	5×5×256	0
Dropout	5×5×256	0
Flatten	6400	0
Dense	1000	6401000
Dropout	1000	0
Dense	1000	1001000
Dropout	1000	0

Dense (Output Layer)	4	4004
----------------------	---	------

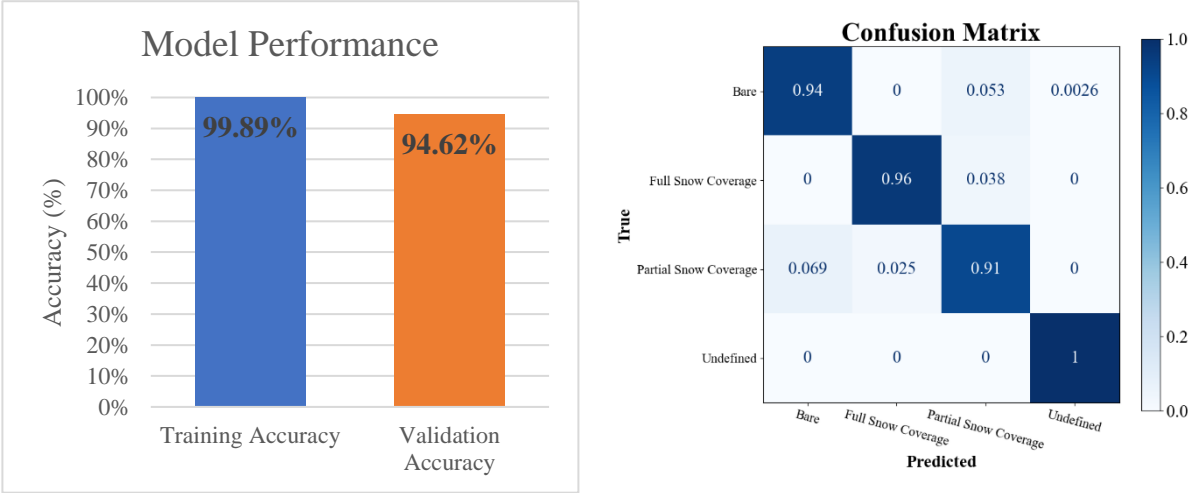


Figure 7 Model performance. Left: Training Accuracy and Validation Accuracy. Right: Confusion Matrix

Results showed that the developed CNN model achieved a high level of training and validation accuracy. This was supported by the confusion matrix, which displayed that the model could distinguish between different RSC categories, as validation accuracy for each category is over 90%. Based on these results, it can be said that the developed CNN model is able to accurately label RSC based on road images.

As the value of RSI is not collected by the AVL directly but is a function of snow and ice coverage level, RSC category predicted by CNN must be first converted into RSI before continuous spatial mapping. In this step, a risk-based RSI values summarized by Fu et al. (8) was adopted. In their original work, there are a total of seven RSC categories. In comparison, we only have three. However, we can convert our three RSC categories into the seven categories defined in the original literature. It is important to note that if the camera images were classified as fully snow covered, there is no way for the model to distinguish between packed and unpacked snow; for this reason, packed and unpacked were grouped together. In addition, we also omitted the icy category entirely, as we do not have a method of determination such as a thermal image or road friction value. RSI of each RSC category is shown below in Table 3.

Table 3 RSC categories and corresponding RSI values

Original			Customized (i.e., our case)	
RSC Category	Road Surface Index (RSI)			RSC Category
	Min.	Max.	Avg.	
Bare and Dry	0.9	1	0.95	Bare Pavement

Bare and Wet	0.8	0.9	0.85	
Slushy	0.7	0.8	0.75	Partially Snow Covered
Partly Snow Covered	0.5	0.7	0.6	
Snow Covered	0.3	0.5	0.4	Fully Snow Covered
Snow Packed	0.2	0.3	0.25	
Icy	0.05	0.2	0.125	NA

Since the customized RSC category shown in Table 3 were based on a one-to-one reflection, using a single value (e.g., mean value of the included original RSC categories) to represent each customized RSC category will inevitably eliminate a lot of details. For example, in the image labelling process, pavements with one wheel path visible, two-wheel paths visible, etc. are all categorized into the partially snow-covered category; however, their RSI can be very different. To solve this issue, the image thresholding (9,10) technique is adopted to separate each partially snow covered image into foreground values (black, representing visible pavement) and background values (white, representing snow and/or ice), then the proportion of background values in each image can be used to adjust the RSI values set in Table 3. For example, the adjusted RSI of a partially snow covered pavement with 60% background values can be obtained by $0.8 - (0.8 - 0.5) * 60\% = 0.62$. In simpler terms, it is a form of pixel proportion analysis. In addition, to eliminate biases caused by unnecessary features within the image (e.g., vehicle headlight, roadside vegetation, etc.), prior to the image thresholding process, each image is cropped to focus on the pavement conditions in the primary driving lane. An example of these two processes can be found in Figure 8. The other two RSC categories (i.e., bare pavement and fully snow covered) are uniformly converted to the mean values of each RSI ranges (i.e., 0.9 and 0.35 respectively), as the snow status in these two categories are nearly identical. All RSC images are converted to RSI following the same procedures. As previously mentioned, there were 34 qualified hourly observations of RSI. Among these observations, the minimum, maximum, mean value and standard deviation of the converted RSI values were 0.35, 0.90, 0.59 and 0.19, respectively.



Figure 8 An example of image cropping and thresholding process

b) Road Surface Index Interpolation

With a reliable DL model at hand, RSI can be obtained from each image following the image thresholding procedures described previously. Then, to spatially interpolate the RSI, RK is integrated following the previously described procedures. Table 4 shows the fitted trend model results from the selected four examples in the study area. *p*-value was adopted to confirm the statistical significance between RSI, geographical, and topographical features at a 5% significance level. Looking at the signs of the coefficients, the calibrate results of the significant variables are intuitive. For instance, RSI is likely to be lower (i.e., the road surface is slipperier) when its location is in a higher altitude.

Table 4 Examples of trend removal results involved in RK for RSI

Date	Hour	Significant variables	Sign of Coefficients	R ²
2018-11-25	10 a.m.	Altitude ²	(-)	56%
2019-01-12	1 a.m.	Altitude/Altitude ²	(+)/(-)	32%
2019-01-12	8 a.m.	Altitude	(-)	21%
2019-01-18	8 p.m.	Altitude ²	(-)	37%

However, by comparing the R² value for each regression model, it was found that the predictive ability of the regression models varies drastically. One possible reason for this is that the random fluctuation in RSI is caused by different weather events. The shortfalls in RSI estimations suggest that using trend model alone may be insufficient, and that kriging should be incorporated to further refine the model to improve its accuracy.

After constructing all hourly semivariogram models, RK is then incorporated to estimate RSI for the remaining locations by incrementally adding point measurements (from one to three) at/near the RWIS stations in the study area. Examples of the cross validation results are shown below in Table 5.

Table 5 Examples of the fitted semivariogram models for RSI

Date	Hour	Nugget	Range (m)	Sill	#inputs point measurements	RMSE	$\overline{\sigma^2}$
2018-11-25	10 a.m.	0.019	52330.546	0.019	1	0.151	0.037
					2	0.152	0.027
					3	0.146	0.023
2019-01-12	1 a.m.	0.010	30676.037	0.018	1	0.160	0.033
					2	0.158	0.023
					3	0.179	0.019
2019-01-12	8 a.m.	0.007	35581.762	0.015	1	0.140	0.026
					2	0.128	0.018
					3	0.116	0.015
2019-01-18	8 p.m.	0.009	55272.873	0.009	1	0.100	0.017
					2	0.154	0.013
					3	0.132	0.011

Looking at the Table 5, RMSE value should decrease as the number of input point measurements increases, as seen in 2019-01-12_8a.m. However, this does not consistently occur for all hourly events. One possible explanation is the sampling location of the point measurements (i.e., the existing RWIS station's location). Some locations provide inputs values that may not be as representative of distant areas as other locations, and thus local attributes of the data points will affect the overall estimation performance, especially if the locations have significantly different local attributes. Different hourly weather events can be another reason for this phenomenon, as the temporal variation of meteorological factors tends to affect spatial structures of RSC. However, the kriging estimation variance does drop with every additional station, meaning by adding more stations, we account for more variance in the RSI estimations, making them more reliable. Ultimately, this shows that adding stations require a placement strategy that incorporates local attributes

and weather events; a potential future research topic worth investigating.

The comparisons between estimated and actual observed RSI profiles (depicted in Figure 9) provide visual evidence that RK model well captures the general spatial variation pattern seen in the observed RSI values. The findings further attest to the proposed method's ability to use as little as a single point measurement as input to fill in the large spatial gap that exists between RWIS stations.

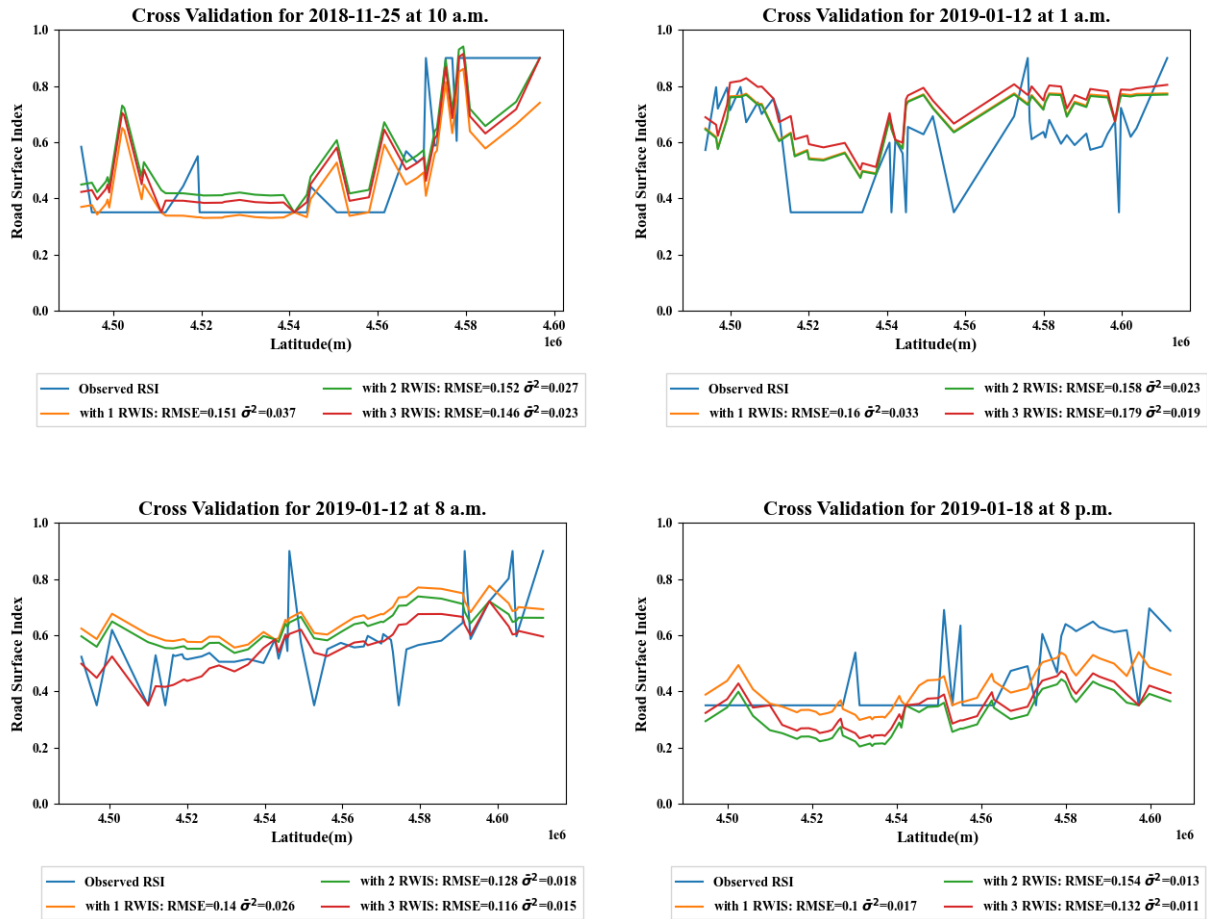


Figure 9 Examples of the RK estimated RSI versus observed RSI

4. Conclusions and Future Recommendations

Monitoring and estimating winter RSC have long been recognized as a challenging task; however, it is critical for optimizing WRM operations and supporting advanced traveler information systems (ATIS). This paper discussed current practices for monitoring winter RSC, introduced state-of-art techniques that can be used in estimating the RSC for unmeasured locations using input from RWIS, proposed an innovative methodological framework that incorporates parametric and non-parametric methods tailored for RSI continuous mapping, and used a case study to demonstrate the feasibility and robustness of the proposed framework.

By cross validating results from RK-estimated RSI, we provided evidence for the feasibility of the proposed method, as well as identified how estimation quality is dependent on the density of the RWIS network. With as little as one point measurement as input, the RK can well mimic the general pattern of the RSI

along the highway stretch. Generally, the estimation accuracy can be improved when the number of point measurements increases. Some hourly events showed that estimation errors (i.e., RMSE) did not decrease as number of input point measurements increased. This phenomenon may be due to weather events since it is not typically uniform over space nor time, ultimately affecting the RK interpolation accuracy. Contrary to this, kriging estimation variance showed that additional stations improved the estimation reliability, it was observed that variance decreased with increasing input point measurements. Overall, this implies that the placement strategy of RWIS stations is very important to account for local factors.

To automate the process of RSC image recognition, a DL (i.e., CNN) model was developed. The high training and validation accuracy out of the DL model confirmed its effectiveness in determining RSC from vehicle-mounted dash camera images. However, we are also aware that these results only indicate the performance of a simple CNN model used for our specific application, and that heavily dependent on image quality.

In terms of future research extensions, to better serve the decision-making process of WRM activities, the application of RWIS location optimization method (29,30) can be applied to determine the optimal number of new RWIS stations required. Their corresponding optimal locations should also be considered by running multiple simulations and incorporating various objectives (e.g., traffic monitoring), weather events, and local-specific attributes.

It is believed that with a larger spatial coverage of RSC monitoring and the techniques presented in this paper, WRM activities can be deployed more quickly, efficiently, and cost effectively. This will benefit the general public with improved traffic safety and mobility, municipalities with a tool to meet their obligations, and taxpayers with reduced costs.

References

1. Feng F. Winter road surface condition estimation and forecasting. University of Waterloo; 2013.
2. Kwon TJ, Fu L, Jiang C. Effect of Winter Weather and Road Surface Conditions on Macroscopic Traffic Parameters. *Transportation research record*. 2013 January; 2329(1): 54-62.
3. Gu L, Wu M, Kwon TJ. An Enhanced Spatial Statistical Method for Continuous Monitoring of Winter Road Surface Conditions. *Canadian Journal of Civil Engineering*. 2019 November; 47(10): 1154-1165.
4. Wu M, Kwon TJ, Fu L. Spatial Mapping of Winter Road Surface Conditions via Hybrid Geostatistical Techniques. ; 2021. Report No.: No. TRBAM-21-00464.
5. Wang X, Gu L, Kwon TJ, Qiu TZ. A geostatistical investigation into the effective spatiotemporal coverage of road weather information systems in Alberta, Canada. *Journal of Advanced Transportation*. 2018 January; 2018.
6. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. Tensorflow: A system for large-scale machine learning. In 12th
7. Baldwin S. Compute Canada: advancing computational research. In *Journal of Physics: Conference Series*; 2012: IOP Publishing. p. 012001.
8. Fu L, Thakali L, Kwon TJ, Usman T. A risk-based approach to winter road surface condition classification. *Canadian Journal of Civil Engineering*. 2017 March; 44(3): 182-191.
9. Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*. 2004; 13(1): 146-165.
10. Dawson-Howe K. *A practical introduction to computer vision with opencv*: John Wiley & Sons; 2014.
11. Wu M. *Evaluating the Safety Effects of Driver Feedback Signs and Citywide Implementation Strategies*. 2020.