A Geostatistical Approach to Winter Road Surface Condition Estimation using Mobile RWIS Data

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Civil Engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjce-2018-0341.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>31-Oct-2018</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Gu, Lian; University of Alberta, Civil &amp; Environmental Engineering Kwon, Tae J.; University of Alberta, Civil and Environmental Engineering Qiu, Zhi-Jun; University of Alberta, Civil and Environmental Engineering</td>
</tr>
<tr>
<td>Keyword:</td>
<td>GIS, road surface temperature (RST), kriging, RWIS, winter road maintenance</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue? :</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
A Geostatistical Approach to Winter Road Surface Condition Estimation using Mobile RWIS Data

Lian Gu*
MSc Candidate
Department of Civil & Environmental Engineering
University of Alberta, 6-362 ICE
Edmonton, AB, T6G 2R3, Canada
Tel: 1-780-695-9285
Email: lgu2@ualberta.ca

Tae J. Kwon
Assistant Professor
Department of Civil & Environmental Engineering
University of Alberta,
Edmonton, AB, T6G 2R3, Canada
Tel: 1-780-492-6121
Email: tjkwon@ualberta.ca

Tony Z. Qiu
Associate Professor
Department of Civil & Environmental Engineering
University of Alberta,
Edmonton, AB, T6G 2R3, Canada
Tel: 1-780-492-1906
Email: zhijunqiu@ualberta.ca

For publication in the Canadian Journal of Civil Engineering
Word count: Words (7145) + 8 Figures + 3 Tables
*Corresponding author
Abstract

In winter, it is critical for cold regions to have a full understanding of the spatial variation of road surface conditions such that hot spots (e.g., black-ice) can be identified for the effective mobilization of winter road maintenance operations. Acknowledging the limitations in present study, this paper proposes a systematic framework to estimate road surface temperature (RST) via the geographic information system (GIS). The proposed method uses a robust regression kriging method to take account for various geographical factors that may affect the variation of RST. A case study of highway segments in Alberta, Canada is used to demonstrate the feasibility and applicability of the method proposed herein. The findings of this study suggest that the geostatistical modelling framework proposed in this paper can accurately estimate RST with help of various covariates included in the model and further promote the possibility of continuous monitoring and visualization of road surface conditions.

Key words: GIS, road surface temperature (RST), kriging, RWIS, winter road maintenance

1. Introduction

During winter seasons, cold regions in countries like the U.S. and Canada often face frequent inclement weather events which cause detrimental effects to road safety and mobility of all traveling public. According to the study conducted by the U.S. Department of Transportation, nearly 22% of all crashes are caused by the adverse weather conditions due to the degradation of visibility and pavement friction (https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm). Therefore, acquiring road surface condition (RSC) information is of utmost importance for road administrations and transportation agencies to ensure that the roads are free of ice and snow, thereby enhancing overall quality of winter road maintenance (WRM) service. These operations, however, always demand substantial financial cost and resources - it is estimated that more than $2.3 billion are spent by North American transportation authorities annually on winter road
maintenance and $170 million by the Canadian government alone (https://ops.fhwa.dot.gov/weather/weather_events/snow_ice.htm) (Lysyk et al. 2015). Therefore, there has been a considerable amount of effort and time put forth to seek cost-effective ways to minimize these WRM expenditures while maintaining a high level of service.

In order to reduce the WRM expense and the subsequent chemical damage on road pavement surfaces, information capturing the spatial variation of road surface condition (RSC) is essential for transportation authorities to identify key locations that are in need of frequent monitoring (e.g., weather related accident prone areas) and decide the proper treatment. Road surface temperature (RST), which is one of the key RSC variables, plays an important role in WRM as it provides information that is essential to performing anti-icing and de-icing operations and can predict the potential for black-ice formation. Previous research on this topic shows that RST can vary by more than 10°C across a road network during the night in winter. These variations in RST indicate that a certain stretch of road may fall below the freezing point while some sections may still be above freezing (Shao et al. 1996). When the surface temperature is lower than the dew point temperature and continues to fall below freezing, there is a greater chance that water will condense and freeze, and black ice may begin to form on the surface. Hence, nocturnal RST is considered as one of the main factors contributing to the formation of black-ice, which can subsequently pose a danger to motorists as it is transparent and thus very difficult to see on a black asphalt pavement.

Acknowledging the need for an accurate prediction of roadway RST, several numerical models have previously been proposed in an attempt to understand the spatial variation of temperature and a general review has been provided by Hammond et al. (2010). Chapman et al. (2001b) developed a multiple regression model to demonstrate that up to 75% of the variation of residual RST can be affected by surrounding geographical parameters. Sass (1992) also proposed a prediction model
using the heat condition and the surface energy-balance models. In recent years, a purely statistical
method known as multivariate linear regression model was introduced by Kršmanč et al. (2013).
Separate models were developed for different locations and periods to deal with non-linear
relationships between parameters and target variable (RST). Sokol et al. (2017) applied an
ensemble technique to the METRo-CZ model for RST forecasting whereas the results were
underestimates. Even though these prior studies helped provide some insights on how temperature
varies over space, they suffer from one major limitation – the models were developed to provide
only site-specific condition information rather than an entire segment of road. Having continuous
RST information over a road network is critical to not only road users for improved safety but also
to winter maintenance authorities responsible for maintaining a good level of service.

In order to improve RST monitoring, many techniques were applied to collect the road surface
information. Thermal mapping (TM) is one of the techniques that can quantify the variation of
RST. It uses a vehicle-mounted infrared thermometer to collect data under different weather
conditions which can then be graphically visualized by creating thermal fingerprints (Shao 1990;
Chapman and Thornes 2005; Marchetti et al. 2011). It is customary that thermal surveys would be
done along a road network over multiple nights with varying atmospheric stability, thus the reliable
amplitude of RST variations can be obtained (Chapman and Thornes 2008). Among many,
Thornes (1991) studied the impacts of different weather conditions on the amplitude of thermal
fingerprints and classified them into four distinct categories (i.e., extreme, intermediate, damped
and other) based on the observed wind speed and cloud cover information (Thornes 1991).
Marchetti et al. (2014) used principle components analysis method to interpolate different thermal
surveys to generate a thermal map to forecast RST. Using these methods could allow road
maintenance personnel to identify the hotspots that may require more frequent monitoring and
additional treatment (Zwahlen et al. 2003). Nevertheless, thermal maps drawn by thermal
fingerprints are quite expensive, as the process of generating such is laborious and time consuming.
This becomes even more problematic for cities where a large-scaled implementation is of high
necessity. Compounding the problem is that this approach can only provide a static forecast of
minimum RST and the thermal maps are merely a snapshot which do not capture the temporal
thermal behavior of road surface conditions (Chapman and Thornes 2006).

Of the many data collection methods used by road maintenance authorities during the winter
months, the most preferred technique is the road weather information systems (RWIS). RWIS can
collect, transmit and disseminate road weather and surface condition information to help road
maintenance personnel effectively plan deicing and snow removal, and reduce material usage (salt
and sand). For this reason, RWIS are widely used in many places in the northern hemisphere, and
North America alone has more than 3,000 RWIS stations currently in operation, and is continuing
to expand to improve their existing WRM services and maximize the return on their investments
(Kwon and Gu 2017). Generally, there are two types of RWIS: stationary and mobile. A stationary
RWIS station is installed along the roadside and collects data at a fixed location, while a mobile
RWIS is a patrol vehicle equipped with advanced sensors and collects data as it travels along the
road. Due to their different data collection mechanisms, the stationary RWIS provides high
temporal but low spatial coverage, whereas the mobile RWIS is able to provide low temporal but
high spatial coverage (Strong and Shi 2008). Both RWIS types are applied by transportation
authorities to improve winter road maintenance.

During the last few decades, a considerable number of RWIS installations were deployed in
Canada as well. For instance, Alberta is one of the leading provinces to take advantage of RWIS
networks to acquire real-time and near future road surface condition (RSC) information. The
advanced sensors and cameras provide detailed and tailored weather forecasts, making WRM more efficient and cost-effective. While effective in providing valuable information, RWIS stations are expensive to install and operate and therefore, can be installed at a limited number of locations. Moreover, as mentioned above, RWIS stations only provide a ‘spot’ measurement of weather and road surface conditions and can be, at times, unrepresentative of distant surrounding areas. Considering the vast road network and the possible inclement winter weather conditions, it is important to accurately extrapolate the road weather and surface condition of locations between different pairs of RWIS stations to help maintain safer road conditions for travelers. To overcome this shortcoming, the focus must move away from just ‘measuring’ and go to ‘modeling’ to make the best use of the information obtained from mobile RWIS.

Therefore, the main objective of this research is to develop a methodological framework to estimate RST using a geographic information system (GIS). Since road weather and surface conditions are generally affected by multiple facets that are very difficult to explain as they exhibit significant randomness, our primary focus will be on developing models related to nocturnal events with different road weather types. To achieve this goal, an advanced geostatistical method known as regression kriging (RK), which has seldom been explored in the transportation field, is proposed to show the feasibility of better capturing the spatial variations of the variable of interest and improve the reliability of prediction results. In particular, advanced geomatics applications such as remote sensing and geographic information system (GIS) are implemented in this study to further improve the predictability of the proposed model.

The remainder of the paper proceeds as follows: Section 2 describes the materials and methods behind kriging. Section 3 discusses the geographical parameters that may impact on RST, while a...
case study of Alberta, Canada is presented in Section 4. The last section provides conclusions and recommendations for future research.

2. Theory of geostatistical analysis

2.1 Kriging

Kriging is a geostatistical interpolation method and its original idea was proposed by a mining engineer D.G. Krig. It provides interpolated values at unknown locations based on a set of available observations by characterizing and quantifying the spatial variability of the area of interest. Let \( x \) and \( x_k \) be location vectors for estimation point and a set of observations at known locations, respectively, with \( k = 1, \ldots, m \), and \( Z \) be a variable of interest (i.e., RST). The predictions are commonly calculated by the following general kriging model (Goovaerts 1997):

\[
\hat{Z}(x) = m(x) + \sum_{k=1}^{m} \lambda_k [Z(x_k) - m(x_k)]
\]

where \( \hat{Z}(x) \) is the predicted value of the target variable at an unknown location. The terms \( m(x) \) and \( m(x_k) \) are expected values (means) of the random variables \( Z(x) \) and \( Z(x_k) \), and \( \lambda_k \) is a kriging weight assigned to datum \( Z(x_k) \) for estimation location \( x \).

The results of kriging would vary by the model adopted for the random function \( Z(x) \) itself. But all kriging methods share the same goal that the weights \( \lambda_k \) are chosen when prediction error variance is minimized. The weights are calculated using the spatial autocorrelation structure, which is typically obtained from a semivariogram model. This general procedure of interpolation is the basis of all kriging methods.

However, some kriging approaches (i.e., simple/ordinary kriging) are limited by considering only the spatial location of measurements without taking into account the influence of other parameters (Olea 2003). Hence in recent years, hybrid interpolation techniques which combine...
two conceptually different methods to model and map spatial variability have received much
close attention among geostatisticians. These techniques generate interpolations not only based on point
observations of the target variable, but also use regression analysis on auxiliary variables (i.e.,
parameters derived from digital elevation models, satellite imagery). One of the most renowned
hybrid interpolation methods is named regression kriging which involves various combinations of
regressions on auxiliary environmental information and kriging (Hengl et al. 2007, Ligas and
Kulczycki 2010). Commonly, a multiple linear regression approach is applied into RK. The
predictions are made separately for the drift and residuals and then added back together as shown
in Equation (2):

\[
\hat{Z}(x) = \hat{m}(x) + \hat{b}(x) = \sum_{i=0}^{k} \beta_k \cdot q_i(x) + \sum_{k=1}^{p} \lambda_k \cdot e(x_k)
\]

where \( \hat{m}(x) \) is the fitted drift (estimation from regression model), \( \hat{b}(x) \) is the interpolated residual,
\( \beta_k \) are coefficients of the estimated drift model and \( \hat{b}_0 \) is the estimated intercept, \( p \) is the number
of auxiliary variables, \( \lambda_k \) are kriging weights and \( e(x_k) \) is the regression residual. Fig. 1 provides
a visual representation of the general concepts of RK that combines both deterministic and
stochastic components of spatial variations of the variable under investigation (i.e., RST). In this
figure, a linear regression model is used first to capture the trend of the target variable, followed
by kriging interpolation on the residuals by characterizing and quantifying the underlying spatial
structure of the observed measurements (to be further discussed in Section 2.2). The estimated
residuals are then added back to the regression results and generate the final predictions.

Specifically, the coefficients \( \beta_k \) of the regression model are estimated using ordinary least
squares (OLS) or, optimally, generalized least squares (GLS). The advantage of GLS is that it
accounts for the spatial correlation of the residuals obtained from the regression model. The equation of GLS is described as follows:

\[
\beta_{GLS} = (q^T \cdot C^T \cdot q)^{-1} \cdot q^T \cdot C^{-1} \cdot z
\]  

(3)

where \( \beta_{GLS} \) is the vector of estimated coefficients, \( C \) is the covariance matrix of residuals described below, and \( q \) is a matrix of predictors at measured locations.

\[
C = \begin{bmatrix}
C(x_1, x_1) & L & C(x_1, x_n) \\
M & O & M \\
C(x_n, x_1) & L & C(x_n, x_n)
\end{bmatrix}
\]  

(4)

After the trend has been estimated, the residual can be interpolated using kriging and added back to the estimated trend. The RK can be conveniently expressed in matrix notation:

\[
\hat{z}(x) = q_0^T \cdot \beta_{GLS} + \lambda_0^T \cdot (z - q \cdot \beta_{GLS})
\]  

(5)

where \( q_0 \) is the vector of \( p+1 \) predictors and \( \lambda_0 \) is the vector of \( n \) kriging weights used to interpolate the residuals at unknown locations.

Previous research indicated that these hybrid techniques tend to outperform the plain geostatistical methods, such as simple or ordinary kriging. They can yield more detailed and accurate predictions by incorporating various covariates in modeling the trend component (Hengl et al. 2004). Since RST is known to be influenced by many external factors, including geographical characteristics and meteorological elements (to be discussed in detail in Section 3), RK is a better option and thus used in this study.

2.2 Semivariogram for Building a Spatial Structure

To apply the regression kriging approach mentioned above, quantifying the spatial autocorrelation of the variable of interest, in our case, RST, is a prerequisite in geostatistics. The spatial variability can be measured by modeling a semivariogram, which depicts how the data are
correlated with its spatial distance based on the observations and location information (Journel and Heuvelink 1978). Due to scarce data points in reality, the points are typically grouped per distance vector \( h \) and the resulting semivariogram is expressed as follows:

\[
\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{k=1}^{m(h)} [z(x_k) - z(x_k + h)]^2
\]

(6)

where \( \hat{\gamma}(h) \) is the sample semivariogram, \( z(x_k) \) is a measurement taken at location \( x_k \), and \( m(h) \) is the number of pairs of observations separated by the lag \( |h| \) in the direction of the vector. The number of pairs to be estimated in this model should at least be equal to 30. Also, the lag distance for an experimental semivariogram should be constrained to half of the diameter in the sampling domain for all direction analysis (Journel and Heuvelink 1978). Note that there should be no trend of systematic variation, thus the estimated result is independent from the individual site \( x_k \).

Generally, three key parameters are used to describe a semivariogram model, including nugget effect, sill, and range as graphically illustrated in Fig. 2.

The nugget effect represents micro-scale variation and measurement errors, or refers to any spatial variability that exists at a distance smaller than the shortest distance of two measurements. The value of \( h \) means the lag distance, and the range indicates lag separation distance at which plateau is reached (i.e., uncorrelated). Sill represents the variance of the random field and magnitude of the plateau beyond the range.

Typically, an experimental semivariogram is smoothed by a mathematical model due to the fact that the estimated model is commonly irregular and the real spatial structure of the region is likely unknown (Oliver and Webster 1990). There are many negative definite functions that can be fitted to describe the semivariances of the sample data such that, negative values of variances can be avoided. The most commonly adopted models are exponential, Gaussian, and spherical models, and the detailed descriptions of these models can be found in Olea (2006).
Since it is critical to ensure that the model selected best captures the shape of the spatial variability of the observations, assessing the goodness of fit for each model is imperative. Crossvalidation is one possible approach that can be adopted to quantitatively analyze the performance of a predictive model using various statistical measures. It is a verification process in which every single observation would be removed with replacement to generate an estimate at the same site of the removal (Olea 2006). The difference between the “observed value” and the “estimated value” is regarded as error. Moreover, directional investigation, also known as anisotropic analysis, should be conducted when there is a significant difference among experimental semivariograms for different directions tested. In this case, any trend in the measurements could be removed to meet the requirement of the semivariogram modeling.

3. Geographical parameters affecting RST

Road surface temperature (RST) is affected by numerous interacting parameters, including those of the meteorological (i.e., solar radiation, wind speed, cloud cover), geographical parameters (i.e., altitude, topography, landuse), and road construction (i.e., traffic, diffusivity). Since the main focus of this study is nocturnal RST, the traffic volume at nights is expected to be low (thus negligible) and the road construction can be considered as constant. As such, geographical parameters that would affect the variation of night time RST are considered and scrutinized in this study, as summarized below.

Latitude

Latitude has effects on controlling the theoretical maximum incoming radiation, which poses a constraint on climate and RST. However, it has little control on the overall minimum night-time RST, as radiative cooling processes begin to dominate after sunset (Chapman et al. 2001a). For
example, countries at higher latitudes always have more snow and ice related problems than low latitude countries.

**Altitude**

Road surface Temperature decreases when altitude increases, typically 6.5 °C per 1000 m and up to a maximum of 9.8 °C per 1000 m (Tabony n.d.). A study conducted by Shao et al. (1997) outlines the impacts of altitude on RST in Nevada in the U.S. and found that the relationship between these two variables was sometimes non-linear. The effects of altitude would be most apparent when the atmospheric stability, which describes the tendency of air to rise or not, is low.

**Topography**

Topography is a main factor resulting in RST differences in extreme nights (Bogren Jorgen and Gustavsson 1991, Yang et al. 2015). Several studies show that katabatic wind flow can generate pools of freezing air in hollows and valley bottoms (Gustavsson 1990). Variables like slope describing the features of topography, can be extracted by a digital elevation model (DEM) to estimate the effects of topography.

**Landuse**

Landuse has a significant impact on RST, and can be observed by the difference between urban/rural temperatures (Faghih Mirzaei et al. 2015). The increased temperatures in urban environments can be accounted for by a phenomenon known as the urban heat island effect, where built up areas are commonly a little warmer than surrounding rural areas. In order to estimate the impacts associated with land use, Normalized Difference Vegetation Index (NDVI), a variable used to describe the spatial heterogeneity of vegetation cover can be adopted to represent the landuse (Garrigues et al. 2006). NDVI is calculated from the visible and near-infrared light reflected by vegetation, which represents the density or heathy state of vegetation since different
kinds of vegetation absorb visible light and reflect near-infrared light to different extents (Kumar and Shekhar 2015).

4. Case Study

4.1 Study Area

The study area of this research covers a total length about 150 km of Highway 16 starting from Edson. As depicted in Fig. 3, it is a major east-west highway located north of Edmonton in Alberta, Canada. The figure also shows the location of existing RWIS stations currently in operation. There are three stationary RWIS stations (red triangle) sited along the survey route. The reason for selecting this route is the varied topography and the diversified surrounding areas including residential areas, Grande Prairie, and woodland, make it possible to collect the data required to conduct the analysis proposed herein.

Additionally, to minimize any possible effects of the temporal variation of RST, the whole study route was divided into three smaller routes, with each of the routes being approximately 60km near RWIS stations. Note that the use of stationary RWIS data (e.g., relative humidity, wind speed, air and dew point temperatures, etc.) is out of scope in this study due data quality issues and is reserved for the next phase of the research to further improve the estimation performance of geostatistical models. A schematic diagram of the steps involved in data collection, integration and aggregation on a GIS platform, and the methodology to develop a regression kriging model is shown in Fig. 4 and the details of each step will be discussed in the following sections.

4.2 Data Description and Processing

A primary technique to collect data used for model calibration and validation in this research is a mobile RWIS unit. It is a newer method for collecting road weather and surface information, using patrol vehicles equipped with innovative technologies such as non-intrusive spectral sensors.
that provide accurate road surface temperature. The data are constantly logged every three seconds when traveling along the road, and it can provide spatially continuous measurements, providing the unique opportunity to calibrate robust statistical models. The mobile RWIS provides observations of various parameters that could possibly make contributions to explaining the variations in the RST, including geographical parameters (i.e., latitude, longitude, altitude), meteorological parameters (i.e., air temperature, dew-point) and road surface conditions (i.e., friction which represents road slipperiness, snow cover situation).

Data sets used in this study are from twelve surveys that were carried out on four winter nights. The dates and the descriptive statistics of the twelve-surveyed RST data are summarized in Table 1.

From this table, it is found that the largest difference between the maximum and minimum road surface temperature can reach to approximately 9°C in segment A on December 18th in 2014, which further proves that the nocturnal RST would vary a lot along the roadway.

As mentioned previously, the use of RK requires auxiliary information to obtain the best prediction results of RST. Hence, a geographic information system (GIS) was used to process and extract the required data in an efficient manner. GIS are computer software packages that integrate user-friendly interfaces for storing, retrieving, analyzing, and visualizing all types of geographically referable data (Kwon and Gu 2017). Not only is GIS capable of dealing with vector and raster data, but it can also effectively process substantial numbers of geospatial datasets. For this reason and beyond, GIS has been widely applied in many studies including transportation engineering over the last decade.

Using Esri’s ArcGIS 10.3, the RST datasets with spatial reference from mobile RWIS were converted into shapefiles and imported with other geographical parameters. The RST data points
were joined to vector road data; in this case, each point will have road attribute data appended to it. Furthermore, to reduce the mathematical complexity of the proposed method and obtain representative geographical data, a uniform buffer zone of 500m was created as a minimum spatial grid to represent the data. Measurements of variables that fall within each equal-length cell were averaged and assigned to the centroid by a geoprocessing tool available in ArcGIS.

The variable slope was derived from DEM using standard surface analysis functions embedded in ArcGIS. For studying the influence of spatial vegetation cover, Landsat satellite images were used to calculate NDVI. The collected images contain reflected light bands in the spectrum of blue, green, red, near-infrared, etc. Since the data comes in a raster format, it can be conveniently integrated and calculated from RED and NIR reflectance on a GIS platform with the following equation:

\[
\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}
\]  

(7)

The result of NDVI varies from -1 to +1; a value close to zero refers to barren areas of rock, or sand and a value close to +1 indicates the high density of green vegetation might be quite high. Consequently, five geographical factors, namely latitude, longitude, altitude, slope, NDVI are considered in this study.

4.3 Modeling of Road Surface Temperature (RST)

The modeling for RST estimation via RK can generally be divided into three parts. Firstly, a multiple linear regression (MLR) analysis is performed to see how much the variance in RST could be explained by these geographical parameters, including latitude, longitude, altitude, slope and NDVI. Secondly, the kriging interpolation approach is used to modify the predicted RST obtained from MLR. The third part of the process is the validation of the calibrated models.
The stepwise MLR analysis was performed using SPSS software by fitting a first-order polynomial to each set of the target variable RST and ensuring the variable is free of trend. Note that a 95% confidence interval was adopted to test the significance of each parameter and the t-value was used to confirm if the independent variable was statistically significant at a significance level of 5%. Table 2 shows the summary results of MLR analysis.

All regression coefficients of geographical parameters make intuitive sense. For example, the surface temperature increases when the value of NDVI increases because of the high density of green vegetations. Furthermore, the study area is a west-east highway, implying that the longitude varies greatly along the road. In this case, the RST decreases with the increase of longitude. The results are also true for latitude and altitude as the RST would drop as it moves to north and higher elevation. Slope measuring the steepness or the degree of inclination of the horizontal plane also has a negative relationship with RST as expected.

The predictive ability of the regression models varies from 1% to 83%, which indicates that the regression model cannot always explain the covariates’ influence very well. The possible cause of these results could be different weather conditions. When the atmospheric stability is low, it is hard to accurately incorporate the impacts of topography into the model due to the more complex surface conditions, which could potentially decrease the accuracy. Weather type is believed to affect the modeling quality - for high wind speed and heavy cloud cover conditions, the modeling quality appears to be lower, and for low wind speed conditions, the quality appears to become higher. Overall, the shortfalls in RST estimations suggest that using MLR alone may not achieve desirable results and that the proposed kriging method be incorporated to further refine the model and improve the accuracy.
The next step is to use MLR results and develop a kriging model as described previously. For this, 70% of observed data for each event were chosen randomly as the training datasets, and the remaining 30% were used as the testing datasets. All the geostatistical analysis and calculation have been conducted in ArcGIS as detailed below.

Step 1: Quantify and model the covariance structure of the residuals derived from the target variable obtained by MLR. Semivariogram models are developed to model the spatial autocorrelation structure, following the criteria defined earlier.

Step 2: Determine kriging prediction map of residuals. This step involves interpolating the residuals from MLR with ordinary kriging at unknown locations by the calibrated semivariogram in Step (1). In addition, the crossvalidation is performed to ensure the accuracy of the predicted data using statistical measures such as root mean square error (RMSE), mean absolute error (MAE) and average standard error.

Step 3: Use the 30% testing datasets to validate the kriging prediction. Validation is a necessary step to further test the goodness of fit of the calibrated model using RMSE or other measures.

Step 4: Generate a final RST prediction map and add the interpolated residuals back to MLR results of the target variable at each prediction location. The outcome of this step is a series of final RST prediction maps.

Fig. 5 -- 7 show the crossvalidation results and the final RST prediction maps with examples for segment A on different dates.

4.4 Results and Discussions

Fig. 8 vividly shows the comparison between the estimation results and the observations for Segment A. A visual inspection confirms that the predicted RST models well capture the general
variation pattern seen in the observed dataset. The largest difference between the predicted and observed values is less than 0.5°C attesting the strong predictability of the RK models developed in this study. Note that similar results have also been obtained for Segments B and C. In addition, the three key parameters of semivariogram models developed in this study and the summary of the calibration and validation results are shown in Table 3.

As for the crossvalidation of training data sets, the RMSE values of all study segments were found to be relatively small, indicating a good prediction capability of the developed models. Additionally, the average RMSE value and MAE value from validation of the testing datasets are 0.323°C and 0.222°C, representing a good performance of the prediction models developed in this study. To see how close the data is to the fitted kriging models, the mean standardized error was calculated and the value was found to be -0.001, which is almost equal to zero. This shows that it greatly enhances the predictive power of the MLR models. All the results of this study indicate that RK has the potential to be adopted to improve the accuracy of model outputs by taking into account the geographical parameters and quantifying the autocorrelation structure of the variable under investigation.

5. Conclusions and recommendations

In this study, one of the renowned geostatistical methods called regression kriging (RK) was proposed to show its feasibility for developing a systematic framework for estimating a key road surface condition variable. Road surface temperature (RST) was used as a variable of interest, and its spatial structure was quantified and modeled using semivariograms. Since the geographical parameters including locational attributes (lat/long), altitude, and slope, could potentially contribute to the variation of nocturnal RSTs, those parameters were prepared and assimilated on a GIS platform for efficient data handling. Furthermore, LANDSAT satellite images were used to
calculate vegetation index (NDVI) to examine the effect of landuse on RST. Using these external parameters as input, MLR models were first calibrated, followed by ordinary kriging to generate highly accurate RST prediction maps. Highway segments in Alberta, Canada, were used as a case study to implement the methods proposed in this study. The main findings are summarized as follows:

- According to the MLR analysis using various geographical parameters, the key factors affecting the variation of RST were found to be longitude, altitude, slope, and NDVI. The sign of their coefficients made all intuitive sense – a negative sign for longitude (continentality), altitude (elevation), and slope (varied topography) and a positive sign for NDVI (vegetation index). The low $R^2$ values of the MLR models posed a strong need to further improve the model quality using the proposed kriging method.

- Using the residuals of the MLR outputs, ordinary kriging was employed to generate a series of highly accurate RST prediction maps (an average RMSE of 0.323 °C). These maps can potentially be used by the respective highway authorities to make more informed decisions on their various winter maintenance activities and prevent road users from getting involved in, for instance, black-ice related collisions during winter seasons.

- Using the models proposed and developed herein, it is anticipated that the continuous monitoring and visualization of road weather and surface conditions could be possible by winter road maintenance contractors for improving the overall quality of their maintenance services while reducing the cost of road patrolling.

The study presented in this paper can be extended in several directions:

- First, other external parameters that may affect the RST variation for instance, the distance from mountains and roadside features should be considered to improve the accuracy of the
models developed in this study. Use of another key RSC variable, namely road surface
index (RSI), should also be explored. RSI is a friction-like measurement and serves as a
WRM performance indicator since it can measure the effect of various winter maintenance
operations on road users. Therefore, further consideration of RSI will be essential in our
future study.

- Second, more case studies should be carried out to see if there is any relationship between
RST spatial structure and different weather events to further generalize estimation models,
as per different weather groups with information collected by stationary RWIS.

- Third, since the temporal variation of RSC might not be well captured via mobile RWIS
units, the methodology should be extended to account for both spatial and temporal
attributes of road weather and surface condition variables. One possible approach to tackle
this would be to fuse the two different types of RWIS data (i.e., mobile and stationary) and
integrate spatiotemporal variograms into the modelling process to develop a more robust
space-time geostatistical model.

- Lastly, since what is commonly available to highway maintenance personnel is stationary
RWIS data, further investigation would be required to infer the road surface conditions
using stationary RWIS data based on the findings generated and the extended knowledge
gained in this research.

Although more case studies and larger datasets should be incorporated to acquire definite and
more conclusive results, this study has reinforced the possibility that the spatial modelling of
geographical parameters coupled with RK can be implemented to improve the overall prediction
capability of road surface conditions variables, which will eminently benefit winter road
maintenance communities.
Acknowledgement

The authors would like to thank Dave Gary and Beata Bielkiewicz at Alberta Transportation for providing mobile RWIS data, which were essential to this study. We would also like to extend our gratitude to anonymous reviewers for their comments and suggestions, which have significantly contributed to improving the overall quality of this manuscript. This research was funded by Alberta Transportation, and partially by Natural Sciences and Engineering Research Council of Canada (NSERC).
Reference


Kwon, T.J., and Gu, L. 2017. Modelling of winter road surface temperature (RST) - A GIS-based


Shao, J. 1990. A winter road surface temperature prediction model with comparison to others. University of Birmingham.


Yang, J., Wong, M.S., Menenti, M., and Nichol, J. 2015. Study of the geometry effect on land
surface temperature retrieval in urban environment. ISPRS Journal of Photogrammetry and

Systems for Snow and Ice Removal Operations- Part I: RWIS. Ohio Research Institute for
Transportation and the Environment, FHWA/OH-20.
LIST OF TABLE CAPTIONS

Table 1. Descriptive analysis of RST data (°C).

Table 2. Summary of MLR results.

Table 3. Summary of semivariogram models and kriging models’ performance.
LIST OF FIGURE CAPTIONS

Fig. 1. The general idea of regression kriging.

Fig. 2. An example of generic semivariogram with key parameters.

Fig. 3. The three study segments in Highway 16.

Fig. 4. An overview of the data collection, processing and proposed methodology.

Fig. 5. Crossvalidation results of 70% training data sets.

Fig. 6. Final RST prediction maps – 20141218 and 20150203.

Fig. 7. Final RST prediction maps – 20150209 and 20161201.

Fig. 8. Predicted RST vs observed RST for segment A on four dates.