

Research Article

A Geostatistical Investigation into the Effective Spatiotemporal Coverage of Road Weather Information Systems in Alberta, Canada

Xu Wang (),^{1,2} Lian Gu (),² Tae J. Kwon,² and Tony Z. Qiu ()²

¹School of Qilu Transportation, Shandong University, Jinan 250061, China
 ²Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada T6G2W2

Correspondence should be addressed to Xu Wang; xuwang@sdu.edu.cn

Received 2 August 2018; Revised 24 October 2018; Accepted 7 November 2018; Published 2 December 2018

Guest Editor: Darius Andriukaitis

Copyright © 2018 Xu Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Inclement weather acutely affects road surface and driving conditions and can negatively impact traffic mobility and safety. Highway authorities have long been using road weather information systems (RWISs) to mitigate the risk of adverse weather on traffic. The data gathered, processed, and disseminated by such systems can improve both the safety of the traveling public as well as the effectiveness of winter road maintenance operations. As the road authorities continue to invest in expanding their existing RWIS networks, there is a growing need to determine the optimal deployment strategies for RWISs. To meet such demand, this study presents an innovative geostatistical approach to quantitatively analyze the spatiotemporal variations of the road weather and surface conditions. With help of constructed semivariograms, this study quantifies and examines both the spatial and temporal coverage of RWIS data. A case study of Alberta, which is one of the leaders in Canada in the use of RWISs, was conducted to indicate the reliability and applicability of the method proposed herein. The findings of this research offer insight for constructing a detailed spatiotemporal RWIS database to manage and deploy different types of RWISs, optimize winter road maintenance resources, and provide timely information on inclement road weather conditions for the traveling public.

1. Introduction

Inclement weather acutely affects road and driving conditions. Approximately 22% of vehicle crashes and 25% of total travel time delays are reported to be adverse weatherrelated in the USA [1, 2]. To mitigate the risk of adverse weather on traffic, proactive strategies for transportation management and road maintenance operations have been developed based on road weather data [3–5]. As the predominant sources of road weather data, stationary road weather information systems (RWISs) [6] provide high temporal but limited spatial data coverage. In contrast, a mobile RWIS [7], which has vehicles collecting road and atmospheric condition information, provides spatially continuous but temporally discrete measurements. To reap the benefits of both systems, road weather data from stationary and mobile RWIS can be integrated to construct an RWIS database with both high temporal and spatial coverage. However, installing and operating both stationary and mobile RWISs is costly. In practice, how to deploy stationary and mobile RWIS to achieve complete spatial and temporal coverage in a costeffective way is yet unresolved. Spatiotemporal coverage is the maximum distance and time lag, beyond which the measurements are no longer representative. It is apparent that an RWIS network with better spatiotemporal coverage will generate more timely and reliable estimations. Timely and reliable road surface condition information is critical for transportation authorities to perform road maintenance and provide weather impact warnings to travelers [8–10]. Adverse road weather impact warnings, such as driving risk alert, advisory speed, and alternative route recommendations, can only be designed when real-time road weather condition data is of high spatiotemporal coverage and resolution. Thus, a cost-effective RWIS network design and deployment require an understanding of spatiotemporal variability in road weather information. To realize such goals, the spatiotemporal coverage of RWIS must be quantified. This study quantifies that data.

Traditionally, RWISs measure atmospheric and pavement conditions in the field and send data to the traffic management centre (TMC) for road weather impact analysis. Tomás et al. [5] proposed an autonomous system that monitors, detects, and forecasts weather incidents with data from stationary RWISs. Gu et al. [11] proposed a regression Kriging method to interpolate road weather conditions using mobile RWIS data. Integrated with other data sources, RWISs can support advanced road weather management. Mahoney and Myers [12] introduced a winter road maintenance decisionsupport system, which integrates weather and road condition data along with maintenance operation rules of practice. Similarly, Cluett et al. [13] developed a framework concept that integrates weather and traffic operation data. In addition, the emerging connected vehicle (CV) technologies enable convenient communication between TMC and road users. In the state of Wyoming, USA, the road condition information is broadcasted from roadside units to CVs for spotting weather impact warnings [14].

Vehicles equipped with road weather sensors can also serve as mobile RWISs and enrich road weather-related data sources. Dey et al. [10] provided a comprehensive practice review and found that CVs can be used as mobile RWISs to enhance route-specific road weather data collection, condition estimation, and traffic management. Petty and Mahoney [15] revealed the potential of CVs in road weather data collection. In addition to the basic atmospheric observations, the data from onboard vehicle sensors, such as wiper state and antilock braking systems, can also be used in road weather condition estimation. With CV-enabled data in hand, Drobot et al. [16, 17] designed a quality and accuracy check process, which includes crosschecks with sensor specifications, climatological ranges, neighboring vehicle, and station measurements, for weather-related data from probe vehicles. Other data analysis shows that the quality of temperature and pressure data is affected by many factors, such as vehicle type, speed, and precipitation occurrence [16], and the temperature data from probe vehicles closely resembles data from stationary weather stations [17]. Most recently, Boyce et al. [18] developed a road weather condition assessment and forecast system, called Pikalert System. It integrates observations from connected vehicles with those from stationary RWIS, radar, and weather model analysis fields. Then it recommends snow and ice removal to road maintenance personnel and provides road weather and condition information to the public. The prototype of the tool has been tested in the Kansas City, USA, since the fall of 2017 [19].

RWISs, the stationary ones in particular, collect road weather condition information effectively but expensively. Additionally, a single RWIS data source can only take measurements that are either rich in space or rich in time—not both. Constructing a database that has both high spatial and temporal coverage data calls for an effective distribution of RWIS Environment Sensor Stations (ESSs) and vehicles. To plan and distribute a cost-effective RWIS network, researchers have attempted to measure the monitoring capability of RWISs by experience-based and modelbased methods. Manfredi et al. [20] suggested that an area with stable road weather and surface conditions can be regarded as regional representativeness in RWIS ESS allocation. Using experience-based methods, Mackinnon and Lo [21] conducted RWIS network expansion design by considering traffic loads, collision rates, climatic zones, and availability of meteorological information. For model-based methods, Singh et al. [22] assumed that the RWIS spatial coverage is a decreasing function of distance. Zhao et al. [23, 24] determined the spatial coverage based on spatial variability. They computed the standard deviation of weather severity in microzones. The spatial coverage of a station was then defined as the size of the critical buffers where the standard deviation changes most quickly with the buffer size. Kwon et al. [25, 26] introduced a spatial variogram approach to determine the autocorrelation range that measures the spatial variability of road weather conditions. The quantified monitoring capability of a stationary RWIS was then used as an input to determine optimal stationary RWIS density [25] and allocate RWIS in a road network [26].

All the studies above attempted to analyze the patterns of spatial dependence in road weather conditions. While time series analysis is a method commonly used to quantify correlation over time, to the authors' knowledge, there still lacks research that investigates spatiotemporal variability of road weather conditions. Incorporating spatial and temporal analysis, spatiotemporal analysis benefits from complete spatiotemporal information and the interactive effects of combining spatial and temporal data. The spatiotemporal variogram commonly used in geostatistics is a useful way to model spatial and temporal dependency and their interactions. It allows researchers to visualize the spatiotemporal variability and estimate the spatiotemporal autocorrelation. A variety of studies have applied the spatiotemporal variogram model to estimate spatiotemporal environmental, meteorological or climatological distributions [27–30].

To bridge the spatiotemporal research gap, this study applied the spatiotemporal variogram to measure the spatiotemporal coverage of RWISs. A case study of Alberta, Canada, was conducted to test the proposed method and look into seasonal trends of spatiotemporal data representativeness. There were three objectives of this study: (1) propose a method to quantify the spatiotemporal coverage of RWIS data; (2) evaluate the proposed method with realworld RWIS data; and (3) investigate the seasonal variations of spatiotemporal coverage. Determining the spatiotemporal coverage of RWIS data can provide a guideline that helps decision makers deploy stationary and mobile RWISs. Also, the spatiotemporal coverage works as an essential parameter in data integration and fusion among multiple data sources.

The remainder of this paper is organized into sections: The next section looks into the theory behind the spatiotemporal variogram; the Methodology section details the proposed method of this study, including the research procedure, data quality diagnostics, spatiotemporal variogram modelling, and cross validation; the Case Study section deploys the proposed method in the study site to investigate the monthly spatiotemporal correlations; and, finally, the last section discusses the concluding remarks and suggests future work.

2. Spatiotemporal Variogram

Geostatistics is known as a class of numerical techniques that characterize spatial attributes and infers variability from the autocorrelation of insufficient data and measure the uncertainty associated with estimation [31]. However, data variability sometimes develops simultaneously in both space and time [32]. As such, the traditional pure spatial analysis has been incorporated with temporal analysis, namely, spatiotemporal geostatistics analysis.

Generally, the key element in spatiotemporal geostatistics is finding a valid and reliable spatiotemporal variogram that quantifies the data structure in space and time domains. A set of variables $z = \{z(s, t) \mid s \in S, t \in T\}$ varies within a spatial domain *S* and a temporal domain *T*. The spatiotemporal variogram is estimated as half of the mean squared difference between data separated by a given spatial and temporal lag (h_s, h_t) [33, 34]:

$$\gamma(h_{s}, h_{t}) = \frac{1}{2n(h_{s}, h_{t})} \sum_{i=1}^{n(h_{s}, h_{t})} \left[z(s_{i}, t_{i}) - z(s_{i} + h_{s}, t_{i} + h_{t}) \right]^{2}$$
(1)

where $\gamma(h_s, h_t)$ is the semivariance of the empirical variogram; $z(s_k, t_k)$ is the measurement at the spatial location s_i and temporal location t_i ; and $n(h_s, h_t)$ is the number of pairs in the observation. Note that h_s is the Euclidean distance, which is a two- or three-dimensional spatial distance vector. Additionally, there should be no trend of systematic variation in the dataset and thus the estimated variogram is independent of the individual location. Figure 1 shows a threedimensional spatiotemporal variogram (Figure 1(a)) together with its projection on the front or side planes (Figure 1(b)). The variogram in Figure 1(b) can be regarded as a pure spatial or temporal variogram.

A variogram model can be described by three parameters: the nugget effect, range, and sill. The nugget effect is the variance at the distance of zero, representing the microscale variation or measurement error. The range is the distance at which the data is no longer autocorrelated and the semivariance starts to level off. That is to say, the spatial and temporal ranges correspond to the spatially and temporally correlated portions in the variogram. Beyond the range, the semivariance becomes a constant value. The sill is the semivariance where the variogram levels off [35].

Typically, the empirical variogram is then smoothed by a mathematical model as the estimated model is commonly irregular, and the real data structure is likely unknown [36]. An appropriate model ensures the positive definiteness of the variogram and sufficient flexibility for the data autocorrelation structure. Thus, many models, including the separable covariance model, product-sum covariance model, and metric covariance model, have been proposed.

The separable covariance model assumes that the spacetime covariance function is the product of separate spatial temporal components (i.e., $C(h_s, h_t) = C_s(h_s)C_t(h_t)$). This model separates the dependence on the two domains and simplifies the variogram modelling. In contrast, the productsum covariance model (i.e., $C(h_s, h_t) = kC_s(h_s)C_t(h_t) +$ $C_{s}(h_{s}) + C_{t}(h_{t})$, with k > 0) was proposed to analyze the interaction between spatial and temporal domains. Also, the metric covariance model includes a space-time anisotropy ratio κ to incorporate distance in time and space (i.e., $C(h_s, h_t) = C_{joint} \sqrt{h_s^2 + (\kappa \cdot h_t)^2}$. Evaluation criteria, such as root-mean-square-error (RMSE) and mean squared error (MSE), can be applied to assess the goodness-of-fit. By measuring the difference between the observed values and the estimated values, the model that best reproduces the variability of the dataset can be selected.

3. Methodology

3.1. Research Procedure. This study analyzed the spatiotemporal patterns of RWIS data in three steps (see Figure 2). First, the raw RWIS data was extracted from the database and input into the data quality diagnostics module. The data quality diagnostics module included the data completeness test, reasonable range test, and model analysis test. Specifically, the data completeness test identified missing records; the reasonable range test recognized erroneous data that are out of a reasonable range; and the model analysis test searched for erroneous data using acceptable regions defined by models. Based on these tests, the data quality diagnostics module excluded invalid and erroneous data in the following steps. Then, with the processed data in hand, this study included the construction and fit of the variogram models by different types of spatiotemporal variograms. The fitted model parameters and the goodness-of-fit were obtained during this stage. Next, by comparing the goodness-of-fit, the best fitted variogram was selected. In this way, the spatiotemporal coverage of RWIS data was measured from the fitted spatiotemporal variogram.

3.2. Data Quality Diagnostics. In this study, missing and erroneous data was diagnosed by applying a diagnostics algorithm. RWISs typically take both road and weatherrelated measurements from various sensors. For data from a specific sensor, the diagnostics algorithm screened the raw data to check data completeness and quality in three tests: the data completeness test, reasonable range test, and model analysis test. First, if a data value from one sensor was void in one record, it was diagnosed as a missing measurement. Next, data that fell outside of the reasonable ranges was recognized as erroneous. The reasonable ranges were determined based on sensor specifications, locationspecific climatological ranges, and historical data ranges [16, 17]. Then, in the model analysis test, data that fell beyond acceptable regions was reported as invalid. The algorithm predetermined acceptable regions according to the relationship among data from different sensors in one record. For example, the feasible road surface temperature (RST)



FIGURE 1: Spatiotemporal variogram. (a) Three-dimensional spatiotemporal variogram. (b) Two-dimensional variogram.



FIGURE 2: Spatiotemporal coverage quantification research procedure.

regions were defined by the regression of the generalized additive model (GAM) from historical data, which is detailed below. GAM was applied to build the relationship between RST and other data. Once RST measurements exceeded the model estimation by a certain percentage (e.g., $\pm 20\%$), the measurements were determined to be erroneous data. In this way, the diagnostics algorithm identified and removed missing and erroneous measurements from the dataset.

In the data quality diagnostics process, GAM became a generalized linear model with linear predictors. The linear predictors were linearly correlated with their unknown smooth functions [37]. The smooth functions expressed the nonparametric nonlinear associations with predictors. The following equations show the GAM formulation.

$$V = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_I(x_I)$$
 (2)

$$f_i(x_i) = \sum_{j=1}^{J} s\left(x_j\right) \tag{3}$$

In (2), *V* is the variable of interest; β_0 is the intercept; and $f_i(x_i)$ represents the smooth function of the predictor x_i , which is formulated as (3). The predictors can include spatial, temporal, and spatiotemporal parameters.

3.3. Spatiotemporal Variogram Modelling. The processed dataset was then applied to construct spatiotemporal variogram models. The basic steps to construct a spatiotemporal variogram are as follows.

First, the systematic trend was removed from the original data. The variogram estimator $\gamma(h_s, h_t)$ was valid only when there was no systematic variation. The measurements $z(s_k, t_k)$ were decomposed into a mean component $m(s_k, t_k)$ and a residual component $r(s_k, t_k)$ (i.e., $z(s_k, t_k) = m(s_k, t_k) + r(s_k, t_k)$). $m(s_k, t_k)$ represents the trend while $r(s_k, t_k)$ represents the fluctuations to be estimated by spatiotemporal variograms. The mean component can be estimated by various theoretical or numerical models. In this step, the residuals were extracted from the original measurements.



FIGURE 3: Data source: stationary RWIS ESSs.

Next, the variogram estimator $\gamma(h_s, h_t)$ was calculated as shown in (1). Note that the residuals $r(s_k, t_k)$ from the last step were regarded as measurements $z(s_k, t_k)$ in (1). The empirical variogram was obtained in this step.

Then, variogram models were fitted. The typical variograms of separable, product-sum, and metric covariance models are given by the following equations, respectively [38].

$$\gamma_{st}(h_s, h_t) = \operatorname{sill} \cdot \left(\gamma_s(h_s) + \gamma_t(h_t) - \gamma_s(h_s) \cdot \gamma_t(h_t)\right) \quad (4)$$

where γ_{st} is the modelled spatiotemporal variogram, γ_s is the spatial variogram, γ_t is the temporal variogram, and *sill* is the overall sill.

$$\gamma_{st} (h_s, h_t) = (\mathbf{k} \cdot \operatorname{sill}_t + 1) \cdot \gamma_s (h_s) + (\mathbf{k} \cdot \operatorname{sill}_s + 1)$$

$$\cdot \gamma_t (h_t) - \mathbf{k} \cdot \gamma_s (h_s) \cdot \gamma_t (h_t))$$
(5)

where k is positive and sill, and sill, are sills in space and time.

$$\gamma_{st}(h_s, h_t) = \gamma_j \left(\sqrt{h_s^2 + \left(\kappa \cdot h_t \right)^2} \right)$$
(6)

where γ_j is any known variogram including a nugget effect and κ is the space-time anisotropy ratio.

When spatiotemporal variogram models were constructed, the model parameters, e.g., spatial and temporal ranges, were extracted from the models. In addition, for interpolation using Kriging methods, the residual component was then interpolated on a fine grid and added to the mean component to generate a fine grid map. *3.4. Cross Validation.* To quantify the goodness-of-fit between the fitted models and the empirical variogram, the authors chose MSE as the measure of effectiveness. MSE allowed the authors to evaluate the overall difference between surfaces. MSE was calculated by the following equation. After comparing MSE among different models, the model with the best goodness-of-fit was selected.

$$MSE = \frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\gamma_{st} \left(h_{s,i}, h_{t,j} \right) - \gamma \left(h_{s,i}, h_{t,j} \right) \right)^{2}$$
(7)

where $\gamma_{st}(h_{s,i}, h_{t,j})$ and $\gamma(h_{s,i}, h_{t,j})$ are modelled and actual spatiotemporal variogram at spatial distance $h_{s,i}$ and temporal distance $h_{t,j}$, respectively.

4. Case Study

4.1. Study Site. The study area was located in the province of Alberta in Canada. Alberta has a humid continental climate, which often produces extremely cold winters [39]. The traffic problems brought by snowfalls in severe winters are always one of the major concerns of road authorities [21]. To acquire real-time road weather and surface condition data, Alberta is one of the leading provinces in Canada building an RWIS network. With innovative sensors and cameras, the stationary RWISs collect detailed road weather and surface conditions, which can be inputs for winter maintenance operations and weather impact warning applications. In this study, data from stationary RWIS ESSs in the vicinity of Edmonton were selected. Figure 3 schematically shows the locations of the 48 stationary RWIS ESSs selected. Stationary RWISs can record



FIGURE 4: Nonlinear associations between RST and predictors. (a) Daily average air temperature. (b) Time of day. (c) Latitude. (d) Longitude.

both road- and weather-related data at prespecified temporal resolutions. The data they measure includes latitude, longitude, elevation, air temperature, humidity, precipitation, surface temperature, and wind speed. RST is one of the most important parameters for road condition prediction (e.g., black ice). Road conditions are sensitive to RST, so this research took it as a measure for estimation. RWIS datasets were available from 2014 to 2016. During this period, data from winter seasons (from October to March) was chosen.

4.2. Data Quality Diagnostics. In the dataset used for this study, the data missing ratio was 1.76% and the erroneous ratio was 6.59%. To identify erroneous RST measurements in the model analysis test, specific GAM model predictors were selected: daily average air temperature, time of day and the latitude and longitude of the stations. The "mgcv" package in the R software (version 3.4.3) [40] was used to fit the GAM model. The nonlinear relationships obtained from the GAM model to identify the acceptable regions are shown in Figure 4. The RST measurements that exceeded $\pm 20\%$ of the model estimation were defined as erroneous data.

The monthly sample size after data quality check is exhibited in Figure 5(a). Each month has 86,498 RST measurements on average. The descriptive statistics for monthly RST data are plotted in Figure 5(b). In Figure 5(b), the central mark on each box is the median value and the edges are the 25th and 75th percentiles. The individual data points beyond the whiskers are outliers. It can be observed that there is an obvious seasonal trend between shoulder months (October and March) and winter months (from November to February).

4.3. Spatiotemporal Coverage. The spatiotemporal variogram modelling was coded in the R software program [40]. The "gstat" package was used to construct and fit the spatiotemporal variogram models. The original data was detrended to obtain residuals by removing the local mean values. The empirical spatiotemporal variogram for each month was built from the variance of difference between any two spatiotemporal residual points. To fit the empirical spatiotemporal variogram, some initial guess of the model parameters was taken from pure spatial variograms and time



FIGURE 5: Data summary after data quality check. (a) Monthly sample size. (b) Boxplot of the processed dataset.

series analysis. Then the spatiotemporal variograms were fitted using the "fit.StVariogram" function. An algorithm in the family of quasi-Newton methods [41], called L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno-B) optimization algorithm, estimated the model parameters.

For the monthly empirical variogram, three spatiotemporal variograms were fitted: separable, product-sum, and metric variograms. The fitted model parameters are listed in Table 1. Among the models, the product-sum model gave the smallest MSE compared to others. Therefore, only the product-sum model was used to fit the variability structure of RST in the following analysis. Figure 6 shows the empirical and fitted product-sum variograms for November 2014 and January 2015. The empirical variograms visualize the spatiotemporal variability of RST data. The parameters from the fitted variograms quantify the spatiotemporal patterns and help to estimate RST in any unobserved space-time point.

The variograms in Figure 6 illustrate the correlation between any two pairs of spatiotemporal data points: the higher the variogram γ_{st} , the higher the correlation. The values for the range model parameter were extracted from the fitted variograms. The data was uncorrelated when the spatial distance or time lag between data points went beyond the ranges. Thus, the ranges were related to data representativeness and can be interpreted as the spatiotemporal coverage. In Figure 6, the lines on the front and side planes are the pure spatial variogram at the zero time lag and temporal variograms at the zero spatial distance, respectively.

As observed from Table 1 and Figure 6, the model parameters differ between months. Thus, the model parameters were extracted from monthly fitted spatiotemporal variograms to examine whether seasonal trends exist. Figure 7 illustrates the monthly variations of spatial and temporal ranges around the winter seasons from 2014 to 2016. The ranges indicate that RST measurements were correlated within a distance on the spatial scale and within a time lag on the temporal scale. The maximum and minimum ranges in space were 7.37 km (March 2015) and 35.55 km (January 2016), and those in time were 6.81h (February 2016) and 16.90h (January 2015). Although ranges varied across months and years, it is still obvious that the data from the months suffering from low temperatures had higher spatial and temporal coverage. RST measurements from RWIS in the winter months (December, January, and February) generally represented more in space and time than the shoulder months (October and March). This phenomenon is reasonable and related with meteorological variations. The severe winter months in Alberta, Canada, experience extremely cold temperatures during which the RST stays in a low level steadily. Thus, each RST measurement in winter months can cover more in space and time. On the contrary, shoulder months witness the transition between seasons. The meteorological variations in either space or time are larger than in other months. As a result, one RST measurement point in shoulder months (October and March) represents road surface conditions in a small coverage of space and time.



FIGURE 6: Spatiotemporal variograms. (a) Empirical variogram of Nov 2014. (b) Fitted variogram of Nov 2014. (c) Empirical variogram of Jan 2015. (d) Fitted variogram of Jan 2015.

The spatiotemporal coverage results provide a reference for decision makers to plan an RWIS network and hence undertake necessary road maintenance and traffic management. The spatiotemporal coverage from variograms is the maximum distance and time lag, beyond which the measurements are no longer representative. In practice, apart from the coverage values from variograms, lower coverage values can be defined to guarantee higher acceptable estimation accuracy.

As previously discussed, RST was chosen without loss of generality, but other road surface and weather indicators can also follow the same procedure to determine their respective spatiotemporal coverage. Once the spatiotemporal coverage is finalized, the density of stationary RWIS and the frequency of mobile RWIS can be specified. On the one hand, as data from stationary RWIS ESSs are discrete in the space domain, the spatial coverage is an essential factor that needs to be considered in planning a stationary RWIS network. Since the spatiotemporal variogram gives a spatial range (SR), measurements from a stationary RWIS ESS can be regarded representative in the distance of SR. Then, along the target roadway, at least one RWIS is required to be allocated every $SR \times 2$. On the other hand, mobile RWIS provides discrete data in time, so the temporal coverage is a prerequisite in scheduling mobile RWIS vehicles. Using the temporal range (TR) from the variogram, measurements from a mobile RWIS can be regarded as representative in a time interval of TR. To construct a database with complete temporal information, one mobile RWIS vehicle should be scheduled to collect road weather data at least every $TR \times 2$. Moreover, to predict temporal changes in the near future, the time headway between two vehicles should be less than TR. TABLE 1: Fitted model parameters for the spatiotemporal variogram. (a) Separable covariance model. (b) Product-sum covariance model. (c) Metric covariance model.



FIGURE 7: Monthly range variations from spatiotemporal variogram. (a) Spatial. (b) Temporal.

Similarly, in a CV environment, only when the penetration rate of CV-enabled vehicles fulfills a full temporal coverage, the CV data can then be interpolated to generate a fine map that shows spatiotemporal variability of road surface and weather conditions.

5. Conclusions and Future Research

Reliable road maintenance and traffic management call for road surface and weather data collection with high coverage; however, RWISs collect road and weather data at a high cost. To help in planning a cost-effective RWIS network, a quantification method was developed to determine spatiotemporal coverage of RWIS data. The proposed method was assessed using authentic RWIS data. There are several key results from this research:

(a) Due to technical problems, the measurements that the studied RWIS ESSs collected contained a certain percentage of missing and erroneous data. Data quality diagnostics and imputation should be given great consideration.

(b) Among the three spatiotemporal variogram models, the product-sum model outperforms the others. The spatial and temporal range values from the model identify the spatiotemporal coverage of RWIS data. The spatiotemporal coverage can be used to decide the RWIS distributions to achieve a fine understanding of the spatiotemporal variability.

(c) A seasonal trend was found behind the monthly spatiotemporal coverage. Data from winter months covers more in space and time than data from shoulder months. The seasonal trend should be considered in the RWIS network design and operation.

(d) Analysis of spatiotemporal semivariograms suggested a stationary RWIS ESS should be located at least every 71.1 km, while a mobile RWIS vehicle should be deployed at least every 33.8 h in the Alberta case.

Based on the observations from this study, the proposed method can be applied in assigning an influence radius of RWIS ESS observations in various road weather-related applications. The current study can be extended in several directions. First, spatial stratified heterogeneity of the demand surface will be tested and attributed accordingly [42, 43]. In our present study, semivariograms are constructed by assuming a stationary process, which means the underlying spatiotemporal structures are translation-invariant over space and time. This rather strong assumption might not hold true when used in large regions with high spatial (i.e., landscape) and temporal (i.e., season) heterogeneity. Second, since semivariogram parameters are sensitive to the property of population, the way of sampling, and the method of estimation, further investigation is required on the spatial sampling trinity to examine the conclusiveness of the findings presented in our study. This will be particularly important when the study is extended to the stage of optimal RWIS network planning. Third, the stationary and mobile RWIS data will be integrated and fused to construct a database with high coverage in both space and time. Data from different sources may not match even in the same time and location. It is challenging to fuse data among multiple data sources and extend their temporal and spatial coverage. A practical, accurate, and time-efficient data integration and fusion method is still required for efficient and reliable road surface and weather condition estimation. Furthermore, the spatiotemporal coverage defined in this study could possibly vary depending upon several external factors, such as geographical and topological characteristics of the site under investigation. Hence, more case studies are essential to obtain more conclusive results.

Data Availability

The RWIS data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research work was jointly supported by the National Natural Science Foundation of China (61703236), Shandong Provincial Natural Science Foundation, China (ZR2017QF014, ZR2018MF027), and China Postdoctoral Science Foundation Funded Project (2017M612275).

References

- [1] Federal Highway Administration, "How do weather events impact roads?" https://ops.fhwa.dot.gov/weather/q1_roadimpact .htm.
- [2] R. M. Alfelor and C. Y. D. Yang, *Managing Traffic Operations during Adverse Weather Events*, Federal Highway Administration, Washington, Wash, USA, 2011, https://www.fhwa.dot .gov/publications/publicroads/11janfeb/01.cfm.
- [3] G. Liu, Y. Ge, T. Z. Qiu, and H. R. Soleymani, "Optimization of snow plowing cost and time in an urban environment: A case study for the city of Edmonton," *Canadian Journal of Civil Engineering*, vol. 41, no. 7, pp. 667–675, 2014.
- [4] S. Choi and C. Oh, "Proactive strategy for variable speed limit operations on freeways under foggy weather conditions," *Transportation Research Record*, vol. 2551, pp. 29–36, 2016.
- [5] V. R. Tomas, M. Pla-Castells, J. J. Martinez, and J. Martinez, "Forecasting adverse weather situations in the road network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2334–2343, 2016.
- [6] P. A. Pisano and G. G. Nelson, "Advanced decision support for winter road maintenance: FHWA documentation of requirements for intelligent transportation systems," *Transportation Research Record*, no. 1741, pp. 129–136, 2001.
- [7] S. Drobot, M. Chapman, P. A. Pisano, and B. B. McKeever, "Using vehicles as mobile weather platforms," *Advances in Intelligent and Soft Computing*, vol. 81, pp. 203–213, 2010.
- [8] T. Kwon, L. Fu, and C. Jiang, "Effect of winter weather and road surface conditions on macroscopic traffic parameters," *Transportation Research Record*, no. 2329, pp. 54–62, 2013.
- [9] M. A. Linton and L. Fu, "Connected vehicle solution for winter road surface condition monitoring," *Transportation Research Record*, vol. 2551, pp. 62–72, 2016.
- [10] K. C. Dey, A. Mishra, and M. Chowdhury, "Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1107–1119, 2015.
- [11] L. Gu, T. J. Kwon, and T. Z. Qiu, "A geostatistical approach to winter road surface condition estimation using mobile RWIS data," in *Proceedings of the 97th Annual Meeting of Transportation Research Board*, 2018.
- [12] W. P. Mahoney and W. L. Myers, "Predicting weather and road conditions: integrated decision-support tool for winter roadmaintenance operations," *Transportation Research Record*, vol. 1824, pp. 98–105, 2003.
- [13] C. Cluett, F. M. Kitchener, L. Osborne, and S. Conger, "Weather integration in transportation management centers," *Transportation Research Record*, no. 1978, pp. 201–208, 2006.
- [14] US Department of Transportation, "Connected Vehicle Pilot Deployment Program: Wyoming," 2017, https://wydotcvp .wyoroad.info/assets/promotion/WyomingCVPilot_Factsheet_ v2_020817.pdf.

- [15] K. R. Petty and W. P. Mahoney, "Enhancing road weather information through vehicle infrastructure integration," *Transportation Research Record*, vol. 2015, pp. 132–140, 2007.
- [16] S. Drobot, M. Chapman, E. Schuler et al., "Improving road weather hazard products with vehicle probe data," *Transportation Research Record*, vol. 2169, no. 1, pp. 128–140, 2010.
- [17] M. Chapman, S. Drobot, T. Jensen et al., "Diagnosing road weather conditions with vehicle probe data," *Transportation Research Record*, vol. 2169, no. 1, pp. 116–127, 2010.
- [18] B. Boyce, G. Weiner, A. Anderson, and S. Linden, *Pikalert® system vehicle data translator (VDT) utilizing integrated mobile observations*, U.S. Department of Transportation, Washington, Wash, USA, 2017.
- [19] Federal Highway Administration, "FHWA road weather management program update," 2018, https://sicop.transportation .org/wp-content/uploads/sites/36/2018/07/RWMP_ProgramUpdate_ July2018.pdf.
- [20] J. Manfredi et al., Road weather information system environmental sensor station siting guidelines, U.S. Department of Transportation, Washington, Wash, USA, 2005.
- [21] D. Mackinnon and A. Lo, Alberta transportation road weather information system (RWIS) expansion study, Alberta Transportation, Edmonton, Alberta, Canada, 2009.
- [22] A. K. Singh, J. Li, M. Murphy, and C. M. Walton, "Optimizing environmental sensor station locations for road weather management: overview and a generalized modeling framework," *Transportation Research Record*, vol. 2551, no. 1, pp. 18–28, 2018.
- [23] L. Zhao, S. I. Chien, X. Liu, and W. Liu, "Planning a road weather information system with GIS," *Journal of Modern Transportation*, vol. 23, no. 3, pp. 176–188, 2015.
- [24] L. Zhao, S. Chien, J. Meegoda, Z. Luo, and X. Liu, "Costbenefit analysis and microclimate-based optimization of a RWIS network," *Journal of Infrastructure Systems*, vol. 22, no. 2, 2016.
- [25] T. J. Kwon and L. Fu, "Spatiotemporal variability of road weather conditions and optimal RWIS density — an empirical investigation," *Canadian Journal of Civil Engineering*, vol. 44, no. 9, pp. 691–699, 2017.
- [26] T. J. Kwon, L. Fu, and S. J. Melles, "Location optimization of road weather information system (RWIS) network considering the needs of winter road maintenance and the traveling public," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 1, pp. 57–71, 2017.
- [27] Y. Yang, J. Wu, and G. Christakos, "Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model," *Ecological Indicators*, vol. 56, pp. 125–133, 2015.
- [28] L. Li, J. Zhang, W. Qiu, J. Wang, and Y. Fang, "An ensemble spatiotemporal model for predicting PM2.5 concentrations," *International Journal of Environmental Research and Public Health*, vol. 14, no. 5, article 549, 2017.
- [29] D. Hu, H. Shu, H. Hu, and J. Xu, "Spatiotemporal regression Kriging to predict precipitation using time-series MODIS data," *Cluster Computing*, vol. 20, no. 1, pp. 347–357, 2017.
- [30] S. O. Ahmed, R. Mazloum, and H. Abou-Ali, "Spatiotemporal interpolation of air pollutants in the Greater Cairo and the Delta, Egypt," *Environmental Research*, vol. 160, pp. 27–34, 2018.
- [31] R. A. Olea, *Geostatistics for Engineers and Earth Scientists*, Springer, Boston, Massachusetts, Mass, USA, 1999.
- [32] P. C. Kyriakidis and A. G. Journel, "Geostatistical space-time models: a review," *Mathematical Geology*, vol. 31, no. 6, pp. 651– 684, 1999.

- [33] P. W. Gething, P. M. Atkinson, A. M. Noor, P. W. Gikandi, S. I. Hay, and M. S. Nixon, "A local space-time kriging approach applied to a national outpatient malaria data set," *Computers & Geosciences*, vol. 33, no. 10, pp. 1337–1350, 2007.
- [34] S. Shekhar, H. Xiong, and X. Zhou, *Encyclopedia of GIS*, Springer, New York, NY, USA, 2008.
- [35] K. T. Chang, Introduction to Geographic Information Systems, McGraw-Hill, New York, NY, USA, 5th edition, 2010.
- [36] M. A. Oliver and R. Webster, "Kriging: A method of interpolation for geographical information systems," *International Journal of Geographical Information Science*, vol. 4, no. 3, pp. 313–332, 1990.
- [37] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*, vol. 43, Chapman and Hall, London, UK, 1990.
- [38] E. Pebesma and G. Heuvelink, "Spatio-temporal interpolation using gstat," *RFID Journal*, vol. 8, no. 1, pp. 204–218, 2016.
- [39] Government of Alberta, Agroclimatic Atlas of Alberta: Climate of Alberta, 2003, https://wwwl.agric.gov.ab.ca/\$department/deptdocs .nsf/all/sag6299.
- [40] R Development Core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [41] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: fortran subroutines for large-scale bound-constrained optimization," ACM Transactions on Mathematical Software, vol. 23, no. 4, pp. 550–560, 1997.
- [42] J. Wang, A. Stein, B. Gao, and Y. Ge, "A review of spatial sampling," *Spatial Statistics*, vol. 2, no. 1, pp. 1–14, 2012.
- [43] J. F. Wang, T. L. Zhang, and B. J. Fu, "A measure of spatial stratified heterogeneity," *Ecological Indicators*, vol. 67, pp. 250– 256, 2016.

