Development of a Novel Road Weather Information System Location Allocation Model Considering Multiple Road Weather Variables over Space and Time

Simita Biswas and Tae J. Kwon

Abstract
This study represents an advanced approach to road weather information system (RWIS) network planning. Here, a methodological framework is developed to determine optimal RWIS locations by integrating two analysis domains: space and time. Using a case study, the application of the proposed method is demonstrated using three critical RWIS variables: air temperature, road surface temperature, and dew point temperature. With these three variables, a series of geostatistical semivariogram analyses are performed to construct a single spatiotemporal model named joint semivariogram, which is able to preserve both spatial and temporal aspects. The constructed joint semivariogram is then used to find the optimal RWIS locations for a randomly generated study area using a popular heuristic algorithm—spatial simulated annealing. The proposed method enhances the previously developed RWIS location allocation model by considering both spatial and temporal components of multiple variables. The finding from this analysis reveals that optimal RWIS location strongly depends on the spatiotemporal autocorrelation structure of the variable of interest. Consequently, location solutions generated using the three variables are found to be different from each other. The variation among the RWIS location solutions is then further quantified by developing a spatial similarity index that is used to measure the degree of spatial similarities between different variables. Overall, the findings documented in this study will provide RWIS planners with a more complete and conclusive location allocation strategy and can act as a decision support tool for long-term RWIS network planning.

Keywords
infrastructure, highway maintenance, road weather

Adverse climatic conditions and an increasing number of weather-related road incidents persist as vital and challenging issues in cold regions. As an important part of modern transportation engineering, intelligent transportation systems (ITS) play an essential role in our everyday life by improving transportation safety and mobility through minimizing weather-related road crashes. Road weather information systems (RWIS)—a critical part of ITS infrastructure—contain advanced sensors that collect, process, and distribute road weather and condition information like air and road surface temperatures, dew point temperature, atmospheric pressure, windspeed, and many others. RWIS information is often disseminated to transportation agencies to assist them in providing a timely and proactive response to severe winter weather events. Additionally, RWIS information is also used for generating dynamic message signs, websites, and services to provide road users with road condition information (1, 2). Because of the numerous benefits associated with RWIS stations, many North American transportation agencies have invested millions of dollars to deploy additional RWIS stations on their road network to improve winter road monitoring coverage (1, 3). However, because of the significant costs associated with installation and maintenance, it is essential to determine
the RWIS sitting locations that can maximize the monitoring coverage of the entire network.

Several studies were conducted in the past to establish proper guidelines for RWIS installations (3–8). The most notable of these was an initiative taken by the U.S. Federal Highway Administration (FHWA), which aimed at providing a standard for RWIS network planning that is dependent on the knowledge and experience of field operators (4). Besides installation guidelines, several studies were performed to build an innovative methodological framework for identifying optimal RWIS locations by generating a more objective way of quantifying the spatial coverage of RWIS data. A GIS-based framework was proposed by Kwon and Fu (5) to evaluate RWIS locations in Ontario, Canada by modeling local road weather conditions (i.e., variability of road surface temperature, mean surface temperature, and snow water equivalent) and topographic location attributes. An approach based on cost–benefit was used by Zhao et al. (8) to determine the optimal locations of RWIS by maximizing spatial coverage considering the standard deviation of weather severity. Another attempt was made by Jin et al. (6), who proposed a location optimization method that maximized spatial coverage of existing RWIS sensors based on weather-related crash data, which was converted into a safety concern index. In a more recent study, Kwon et al. used advanced geostatistical analysis to generate a methodological framework for determining the underlying spatial structure of RWIS measurements. In this research, an efficient combinatorial optimization algorithm, spatial simulated annealing (SSA) was used to perform an RWIS network location allocation analysis. The optimization was formulated as a nonlinear integer programming (NIP) problem to maximize monitoring capability while minimizing kriging errors (3).

While the above discussed literature has contributed to advancing RWIS location models, the studies so far suffer from several limitations. First, optimal RWIS locations were generated by investigating the spatial characteristics of one single RWIS variable, road surface temperature (RST). Although RST is an important measurement, RWIS provide many other weather variables that need to be considered alongside. Measurements like AT (air temperature) and DPT (dew point temperature) are just as critical in forecasting road weather conditions (for example, generation of black ice). Second, previous studies dealt solely with spatial domain of a single RWIS variable, which does not account for inherent temporal correlation of weather and road surface conditions. And since the weather parameters vary over space and time, it is essential to investigate both the spatial and the temporal variability of the variables of interest.

Spatiotemporal semivariogram analysis is an advanced geostatistical method that is used to evaluate spatial and temporal dependency of parameters that tend to fluctuate over space and time (e.g., road weather variables). This method works by combining spatial and time-series analysis to preserve the interactive effect of time variation on spatial domain and vice versa, allowing for the visualization of the spatiotemporal variability in the variable of interest and the determination of autocorrelation range over space and time. For this reason, spatiotemporal analysis has been established as more accurate than spatial analysis alone (9). Recently, a small handful of studies used the above-mentioned technique in their research. Several pieces of research used the underlying techniques to model air pollutants by measuring the space–time variability of certain particles’ concentrations (9–12). Hu et al. (13) applied spatiotemporal regression kriging to predict precipitation using moderate-resolution imaging spectroradiometer (MODIS) and normalized difference vegetation index (NDVI) data. In the field of transportation, the usage of spatiotemporal data has been noticed recently for evaluation of spatiotemporal outliers and the identification of erroneous sites (14); traffic accident prediction using a deep learning approach (15); and several others. One notable study related to the topic of interest is the investigation of spatiotemporal variability of road weather and surface conditions using RWIS data from Alberta, Canada (16). The output of this study provided both spatial and temporal features of the RWIS database. Although the findings were well received, there still a need for research to combine spatial and temporal variability to generate a single spatiotemporal model by preserving both space and time attributes. Moreover, spatiotemporal analysis has never been taken into consideration in the location allocation optimization of RWIS.

Therefore, the main objective of this study is to develop a methodological framework to optimize RWIS location by considering both spatial and temporal variability of multiple weather variables. The problem was formulated on the basic premise that data from RWIS stations within a region should be collectively used to maximize their overall monitoring quality. By considering multiple variables, this study will evaluate the spatiotemporal autocorrelation structure of these variables and examine their effects in generating optimal location solutions and their spatial similarities.

Methodology

Overview of Research Procedures

The first phase of this study involves analyzing spatiotemporal correlation of RWIS measurements, followed by RWIS location optimization using semivariogram
parameters. Three different variables of interest are used in this study; namely, Air Temperature (AT), Road Surface Temperature (RST) and Dew Point Temperature (DPT). The raw RWIS measurements are extracted from the RWIS database, and are then checked using the following steps: data completeness test, reasonable range test, and a neighborhood value comparison. Once the data has been checked, the RWIS data is detrended with respect to time using a generalized additive model (GAM), followed by statistical analysis using descriptive statistics and correlation analysis. In the second stage, one of the most comprehensive spatial sampling techniques, called geostatistical analysis, is employed in conjunction with geographically distributed data. This technique maximizes the probability of capturing the spatial and temporal variations of the RWIS variables and minimizes the potential bias associated with input data. More specifically, spatiotemporal analysis is performed by constructing empirical variograms from processed data, which optimizes parameter estimations for unsampled locations and captures the possible autocorrelation associated with the RWIS variables. Joint semivariogram models are then developed by combining spatial and temporal semivariograms. Finally, a model-based approach via kriging is utilized in the final stage to obtain unbiased estimates with the lowest variance (i.e., uncertainty) to determine the optimal RWIS locations. The study is concluded by generating three sets of optimal RWIS locations and performing a similarity analysis. Research procedures for this study are summarized in Figure 1.

Data Quality Diagnostics

Downloaded RWIS data is processed to remove missing and erroneous data. All three types of data are cross checked with one another to search for outliers. After that, data completeness is checked by identifying the total amount of missing data for each sensor—any stations with more than 15% of missing data are not used in this study. Following this, reasonable ranges of each variable under investigation are checked based on historical data ranges for the associated region and month. Filtered data is then used for pattern analysis by plotting the day of the month versus average daily temperature for all selected sensors, all weather parameters, and all months. All selected sensors for a specific weather parameter are expected to show a similar pattern throughout the month. If any unusual pattern is noticed for a specific station, the associated data is further investigated. In total, six sets of data (two months for three weather parameters) are analyzed in this study. At the end of data processing, AT, RST, and DPT data is de-trended with respect to time using a GAM to incorporate shorter scale variation in the temporal domain. Here, GAM works as a generalized linear model with linear predictors. The GAM function is

---

**Figure 1.** An overview of the research procedure.

Note: RWIS = road weather information system; AT = air temperature; RST = road surface temperature; DPT = dew point temperature; W.R.T. = with respect to.
formulated as \( m = \beta_0 + f_1(x_1) + f_2(x_2) + \ldots + f_i(x_i) \), where \( m \) = variable of interest, \( \beta_0 \) = intercept, \( f_i(x_i) \) = smooth function of predictor \( x_i \). The smooth function can be expressed as \( f_i(x_i) = \sum_{n=1}^{m} s(x_n) \) (16, 17). Once the data has been processed, statistical analysis is performed on the processed data set to produce descriptive statistics and correlation analysis among the weather parameters.

**Spatiotemporal Semivariogram Analysis**

Spatiotemporal analysis is generally conducted for variables that vary over space and time. In this study, spatial and temporal continuity analysis for weather variables is performed using geostatistical spatiotemporal semivariogram modeling. The traditional spatial analysis conducted in our previous efforts is incorporated with temporal analysis to consider spatiotemporal effects.

Spatiotemporal semivariogram is a graph of semivariance versus lag separation distance between pairs of measured points. Semivariance can be defined as a measure of dissimilarity between two measurements as a function of separation distance. The semivariogram has three basic parameters: nugget, sill, and range. Nugget represents measurement or sampling error; it can be defined as dissimilarity with zero separation distance, which is theoretically zero. Sill is the semivariance value at which the semivariogram levels off. The difference between sill and nugget is called partial sill and this value is often encountered during semivariogram analysis. The distance at which the semivariogram reaches the sill value is defined as the spatial range of autocorrelation, where there is no correlation among the measurements beyond this spatial range (18–21). A typical semivariogram with its parameters is presented in Figure 2a.

Spatiotemporal semivariogram modeling, on the other hand, is conducted by integrating both spatial and temporal effects of regionalized random variables (e.g., road weather). Generally, a set of variables in a spatiotemporal field can be defined as \( z = \{ z(s,t) | s \in S, t \in T \} \), where \( S \) = spatial domain and \( T \) = temporal domain. Thus, random field can be specified as \( Z \) and by \( z_i = Z(s_i,t_i), i = 1, 2, 3, \ldots, n \times T \) where \( n \) = number of stations and \( T \) = number of time points. The random fields \( Z(s,t) \) can be modeled as \( Z(s,t) = \mu(s,t) + \epsilon(s,t) \) where \( \mu(s,t) \) = the deterministic part and \( \epsilon(s,t) \) = the stochastic part. Deterministic part refers to the spatiotemporal trend while stochastic part is the zero-mean second-order stationary spatiotemporal random field (12). Spatiotemporal semivariogram modeling is conducted using the stochastic part. Spatial and temporal variance are estimated as half of the mean squared difference between pairs of data separated by a user defined spatial \( (h_s) \) and temporal \( (h_t) \) lag. The general equation of semivariance is presented in Equation 1.

\[
\gamma(h_s, h_t) = \frac{1}{2n(h_s, h_t)} \sum_{k=1}^{n(h_s, h_t)} [z(s_k, t_k) - z(s_k + h_s, t_k + h_t)]^2,
\]

where \( \gamma(h_s, h_t) \) = estimated semivariance value, \( n(h_s, h_t) \) = total number of pairs in analysis domain, and \( z(s_k, t_k) \) = measurement at spatial location \( s_k \) and temporal location \( t_k \) (22, 23). A three-dimensional spatiotemporal semivariogram is presented in Figure 2b.

After constructing the empirical variogram, a mathematical model is used to smoothen the graph by resolving the irregular pattern. There are several covariance models used for spatiotemporal semivariogram modeling (9, 24–26). The most popular and widely used covariance models are: (a) separable covariance model, (b) product-sum covariance model, (c) metric covariance model, (d) sum-metric covariance model, and (e) simple sum-metric

---

Figure 2. A typical: (a) semivariogram and (b) spatiotemporal semivariogram.
covariance model. According to previous studies, the sum-metric model performs better in fitting the spatiotemporal variogram using environmental parameters (i.e., smallest mean squared errors) \((11, 13)\). For this reason, the sum-metric covariance model was selected in this study. This model is a combination of spatial, temporal, and metric models. The equation for a covariance function of a sum-metric model is presented in Equation 2.

\[
C_{sm}(h, u) = C_s(h) + C_t(u) + C_{joint}\left(\sqrt{h^2 + (k_u)^2}\right). \quad (2)
\]

Here, \(h\) and \(u\) represent the spatial and temporal components respectively; \(k\) is a spatiotemporal anisotropy parameter. Similarly, the equation for a sum-metric variogram is given in Equation 3.

\[
\gamma_{sm}(h, u) = \gamma_s(h) + \gamma_t(u) + \gamma_{joint}\left(\sqrt{h^2 + (k_u)^2}\right). \quad (3)
\]

Spatial, temporal, and joint nugget are estimated separately in this model. A spatiotemporal anisotropy parameter (StAni) is used to create the joint semivariogram by combining the spatial and temporal semivariance. The number of space units equivalent to one time unit is defined as StAni. RMSE (root mean square error) is used in this study to measure the goodness-of-fit of the resultant model.

**Location Allocation via Spatial Simulated Annealing**

An innovative RWIS location modeling framework was developed in our previous efforts where the problem was formulated as an integer programming problem with the objective of minimizing the spatially averaged kriging variance (in other words, maximizing spatial coverage) across the road network \((3, 7, 27, 28)\). For the first time in the literature, the method developed there provided decision makers with a tool needed to optimize and expand their existing RWIS networks.

In this study, a more refined location optimization model is proposed by integrating joint semivariogram parameters generated for different weather variables to represent their distinctive spatiotemporal characteristics. Similar to our previous studies, the objective function is again formulated to minimize the sum of mean ordinary kriging (OK) estimation variance. The equations of the objective function and its related computation process are shown below.

\[
G = \begin{bmatrix}
\gamma_{x_1, x_1} & \gamma_{x_2, x_1} & \ldots & \gamma_{x_k, x_1} \\
\gamma_{x_1, x_2} & \gamma_{x_2, x_2} & \ldots & \gamma_{x_k, x_2} \\
\ldots & \ldots & \ldots & \ldots \\
\gamma_{x_1, x_k} & \gamma_{x_2, x_k} & \ldots & \gamma_{x_k, x_k} \\
1 & 1 & \ldots & 1 & 0
\end{bmatrix} \quad (4)
\]

\[g = [\gamma_{x_0, x_1} \gamma_{x_0, x_2} \ldots \gamma_{x_0, x_k}]^T \quad (5)\]

where \(x_i\) \((i = 1, 2, \ldots, k)\) is the sampling site of a sample subset of size \(k\), and in this case, \(k\) is equal to the number of RWIS stations; and \(\gamma_{x_i, x_j}\) is the semivariance between sampling site \(i\) and \(j\).

\[
\sigma^2_{OK1}(x_0) = g'G^{-1}g. \quad (6)
\]

Based on the above three equations, the objective function of this work can be formulated as

\[
f(w) = \frac{\sum_{i=1}^{n-k} \sigma^2_{OK1}(x_0)}{n} \quad (7)
\]

subject to

\[
n = \text{Total number of candidate RWIS station locations in the study area}. \quad (8)
\]

The RWIS location modeling being tackled here requires mathematical and computational methods to find optimal solutions for an objective function, which is usually performed under some form of constraints. For a larger-sized optimization problem, a heuristic algorithm is an effective method for finding the solutions \((29)\). The optimization method implemented in this study is SSA, which is a spatial counterpart to simulated annealing (SA) \((7, 30)\). SSA is a popular heuristic algorithm used to solve spatial optimization problems \((7, 31–34)\). SSA works by slightly perturbing previous sampling designs using random search techniques. As the optimization process continues, it is necessary to avoid local minima, and thus SSA accepts not only improving solutions, but also worsening solutions based on a certain probability. The probability of accepting worsening solutions is typically set initially at 0.2, and this probability decreases exponentially to zero as a function of the number of iterations \((30, 33, 34)\). The workflow of SSA for a certain number of RWIS stations is displayed in Figure 3.

The optimization follows an iterative process where stations are added incrementally into the study; and locations are selected based on heuristic attempts to minimize the objective function. When adding new stations, the placement area is limited to a square region within the study area. This is done to reduce computational complexity and algorithm runtime. The number of RWIS stations to be located in this square region is arbitrarily limited to 10, which is equal to the existing number of RWIS stations in the square region. For the optimization process, two top criteria are implemented. If the number
of iterations exceeds 100,000, the optimization process will stop. And if no improvements are made in the objective function after 200 iterations, the algorithm is set to automatically stop (7).

![Workflow of spatial simulated annealing.](image)

**Case Study—Iowa, United States**

**Study Area and RWIS Network**

The study area is in the state of Iowa, U.S.A. Iowa is generally a flatland area consisting of rolling plain lands and flat prairies. The altitude of this state is 146 m to 509 m (35). The total count of RWIS stations for Iowa is 88. The study period selected for this analysis includes a shoulder season month (October 2016) and a mid-winter month (January 2017) to best capture challenging winter driving conditions. Distribution of RWIS stations for the study area is presented in Figure 4. A randomly generated 10,000 km² region within the study area is used as the experimental boundary for location optimization, presented as the red box below, in which there are 10 RWIS stations.

![Distribution of road weather information system (RWIS) stations for Iowa.](image)

**Data Description and Quality Diagnostics**

RWIS data for Iowa was downloaded from Iowa State University website (http://mesonet.agron.iastate.edu/RWIS/) as an Excel file. Measurements from a typical RWIS station include, but are not limited to AT, RST,
DPT, visibility, wind speed, and road surface conditions. RWIS measurements were collected every 15 to 20 min, totaling 1488 h of data to be used in this study. As discussed previously (Figure 1), these data were processed to remove “noise” via four steps: (a) data completeness test to identify missing data, (b) reasonable range test to find erroneous data, (c) comparison with neighboring observations, and (d) detrending processed data with respect to time using GAM.

Statistical analysis was then performed using descriptive statistics of the processed data and correlation analysis among weather parameters. Descriptive statistics of AT, RST and DPT are presented in Table 1. On closer inspection of Table 1, it was revealed that there was relatively less variation in monthly temperatures in the mid-winter month than in the shoulder month for AT and RST, whereas DPT showed the opposite trend. AT varied from $-26.33^\circ C$ to $33.10^\circ C$ over the month of October 2016, while the RST and DPT varied from $-21.10^\circ C$ to $41.50^\circ C$ and from $-27.83^\circ C$ to $23.60^\circ C$, respectively. For the month of January 2017, temperature varied from $-22.00^\circ C$ to $11.20^\circ C$.

Figure 5 shows the maximum, minimum, average, and standard deviation of weather data for the study area. Correlation analysis was conducted among the weather variables to find how strongly two variables were correlated to each other using values between $-1$ and $+1$. A correlation coefficient of positive 1 indicates an ideal positive correlation, whereas a correlation coefficient of negative 1 indicates an ideal negative correlation. A correlation coefficient near zero indicates no correlation at all ($\approx 0$). Correlation analysis in this study was performed using the Analysis ToolPak in Excel. For each month, 10 RWIS stations were selected randomly, and the correlation coefficient values were generated between each pair of weather variables. These 10 values were then averaged to determine the resultant correlation coefficient. Table 2 shows the results from this analysis for October 2016 and January 2017.

According to Table 2, the weather variables were more correlated during the mid-winter month than the shoulder month. Correlation between AT and RST was higher than for any other variable pairs, having a correlation value of over 0.9 in both study months. In contrast, the lowest level of correlation was observed between RST and DPT, thereby further attesting to the need for taking into consideration the distinctive road weather characteristics later in the location optimization.
phase. Figure 6 presents a plot comparing the correlation coefficient among AT, RST, and DPT of Iowa, and clearly shows that the three variables considered have different degrees of similarities over different months.

Spatiotemporal Semivariogram Modeling

As evidenced in the previous section, correlation coefficients varied from one variable to another, and over different months. Spatiotemporal semivariogram modeling was subsequently conducted to gain a deeper understanding of their spatial and temporal variability, and to use as input to the location optimization model. For this purpose, RWIS data for those two select months were further processed and aggregated using 1 h interval for the time domain. A space–time matrix was then formulated as an input for the spatiotemporal analysis. Separate analysis was conducted for each month and each weather parameter, that is, AT, RST, and DPT. According to previous studies, 30 or more sampling points are needed to construct a reliable semivariogram model (19), which we exceeded by employing 88 stations from a robust RWIS network. Semivariogram modeling was performed using the R statistical package—version 3.2.5 (37, 38). Table 3 represents the spatial, temporal, and joint semivariogram parameters, respectively. According to Table 3, higher spatial ranges were obtained for the mid-winter month than for the shoulder month. The spatial range is close to 20 km for all weather variables in January 2017, and a 10–15 km spatial range was obtained for October 2016. The temporal range was found to be 8.5–12 h for the shoulder month, and 14–21.5 h for the mid-winter month. Such findings make intuitive sense since road weather and surface conditions tend to change more abruptly during shoulder months than they do during mid-winter months when the variability of weather is relatively low (39). To consider these unique characteristics and implement them in the optimization phase, spatial and temporal semivariograms were combined using spatiotemporal anisotropy to estimate the joint semivariogram. The resultant joint semivariogram parameters are presented in Table 3 with the associated StAni and RMSE. Spatiotemporal range for January 2017 was found to be 17 km for all three weather variables, whereas seasonal trend is well captured in the joint semivariogram parameters where a longer range was generated for the mid-winter month than for the shoulder month. In addition, the spatiotemporal range was found to be identical for the month of January, indicative of high correlation among three weather variables, whereas

Table 3. Spatiotemporal Semivariogram Analysis Result

| Month of analysis | October 2016 | January 2017 |
|-------------------|-------------|-------------|
| Weather parameter | AT          | RST         | DPT         | AT          | RST         | DPT         |
| Spatial variogram |             |             |             |             |             |             |
| Nugget (°C²)      | 0.00        | 2.20        | 0.71        | 0.00        | 0.00        | 1.10        |
| Range (m)         | 9505.06     | 9435.23     | 14939.82    | 19489.51    | 19461.90    | 19768.43    |
| Sill (°C²)        | 5946.45     | 6514.13     | 8986.01     | 20295.02    | 21321.83    | 9556.25     |
| Temporal variogram|             |             |             |             |             |             |
| Nugget (°C²)      | 1.29        | 5.86        | 0.00        | 0.00        | 1.34        | 0.00        |
| Range (min)       | 502.45      | 716.63      | 544.97      | 1246.01     | 1292.53     | 832.74      |
| Sill (°C²)        | 173.58      | 258.83      | 177.66      | 123.08      | 367.23      | 319.27      |
| Joint variogram   |             |             |             |             |             |             |
| Nugget (°C²)      | 0.00        | 1.67        | 0.19        | 0.00        | 0.22        | 0.00        |
| Range (m)         | 7613.03     | 7872.96     | 9970.49     | 16863.75    | 16839.31    | 16897.95    |
| Sill (°C²)        | 6682.49     | 1966.42     | 3052.39     | 4982.34     | 5817.95     | 3884.62     |
| StAni (model)     | 0.126       | 0.028       | 0.138       | 0.905       | 0.033       | 0.146       |
| RMSE              | 0.336       | 5.823       | 0.318       | 0.069       | 0.272       | 0.353       |
| Model Spherical   |             |             |             |             |             |             |

Note: AT = air temperature; RST = road surface temperature; DPT = dew point temperature; StAni = spatiotemporal anisotropy parameter; RMSE = root mean square error.

Figure 6. Correlation coefficients comparison of multiple weather variables.

Note: AT = air temperature; RST = road surface temperature; DPT = dew point temperature; IA = Iowa.
different degrees of correlations were observed for the month of October, possibly because of the presence of highly variable weather patterns. Based on the recorded results, the parameters of the joint semivariogram are used for RWIS location optimization.

Effect of Road Weather Variables on Optimal RWIS Locations

Optimal RWIS locations were determined by relying on spatiotemporal semivariogram analysis results that were obtained in the previous section. To quantitatively appreciate the effect each road weather variable has on the resulting RWIS locations, the optimization was performed separately for each set of weather variables and month of data using the R statistical package—version 3.2.5 (37, 38). To reduce the computational complexity associated with the location optimization process and shorten the algorithm runtime, a randomly selected square region (100 × 100 km) within the study area was used as the experimental boundary for location optimization. A 5 × 5 km prediction grid was generated along the square area in ArcGIS to create the candidate sites for RWIS station placement. The computer used to run the optimization was equipped with a 3.60 GHz CPU and 16 GB memory. An initial seed value was used for the optimization process to generate comparable results. The main goal of this location optimization is to prove the hypothesis that location solutions generated by individual weather variables may not lead to the same results, therefore revealing the need for generating multi-criteria location allocation optimization. The objective function value for each of the three sets of optimizations is plotted in Figure 7, a and b, for October 2016 and January 2017, respectively. The decreasing trend of the objective function value with respect to number of iterations demonstrates the superiority of the SSA algorithm implemented and its ability to achieve convergence near the 35,000th iteration. The average algorithm running time for generating each solution set was approximately 5 to 8 h.

As a result, the RWIS network optimization output under different weather variables is presented in Figure 8, a and b, for October 2016 and January 2017, respectively. A 10 km buffer was created around the stations arbitrarily to help better recognize the distribution of stations within the rectangle study area. It is evident that the RWIS location solution is substantially different depending on the variable of interest and over the month of analysis. These findings reveal that the spatiotemporal autocorrelations of the three weather variables have strong effect on generating optimal RWIS locations, and that there is a resurgent need to consider these effects collectively in the optimization framework so that locations selected are most representative of their uniquely different spatiotemporal characteristics.

Spatial Similarity Analysis

It is obvious from the location allocation output (Figure 8) that the optimized location solutions are visually different from one another. To quantitatively evaluate the closeness of the optimal RWIS station distributions, a spatial similarity index was developed. For this, a sensitivity analysis was performed to objectively measure the similarity of optimal locations using ArcGIS (40). Initially, buffer polygons were created around the stations up to a maximum of 20 km in linear units. To perform this analysis, 10 sets of buffer polygons were created for each solution set, which totals up to 60 sets of buffer polygons. The intersecting areas were then determined for each pair of solution sets, each linear unit combination, and each month of analysis. The percentage of intersecting area with respect to the study area represents the spatial similarity index; the higher the percentage of intersected area, the more similar the solution sets. Figure 9 presents the results from this analysis.

According to Figure 9, the percentage of intersecting area with respect to different buffer sizes follows an exponential function. It is clear from the figure that buffer size does not affect the results significantly. In most cases,
location solutions generated using the mid-winter month (January 2017) data set have been found to be closer than those of the shoulder-month (October 2016) data set. The reason behind this outcome is that the daily fluctuation in weather data is greater during shoulder months than during mid-winter months, which generates relatively higher spatiotemporal autocorrelation of weather data for January than for October. For a specific buffer size, percentage of intersecting area between AT and RST solution sets was found to be the highest, followed by, in decreasing order, AT and DPT solution sets, RST and DPT solution sets, and AT, RST, and DPT solution sets. This phenomenon is quite similar to the correlation analysis results among the weather variables, where the correlation between AT and RST was the highest among all, followed by AT and DPT, and RST and DPT. The findings of this analysis indicate that spatial similarity of location solutions can be examined by creating buffer polygons regardless of the size implemented.

**RWIS Location Allocation for the Entire Iowa State**

In the previous section, it was confirmed using the hypothetical square region that the generated locations were dependent on weather variables. Based on this finding, the proposed location optimization method was expanded to cover the entire Iowa state. To achieve a comparable result to that for the existing RWIS locations, a constrained optimization was performed with respect to the road network using a shoulder-month data set of weather variables. Although the existing number of RWIS stations is 88, the total number of all new station has been set to 61 for the location optimization. The number 61 was chosen based on our previous study on optimal RWIS density guidelines (28), where the number of RWIS stations per unit area was generated given the topographic variation and weather severity of regions. In computational efficiency, the average running time for each set of optimization algorithms was approximately 78 h. Figure 10 represents the plot of the objective function value as a function of iteration; and Figure 11 shows existing and optimized RWIS locations with multiple weather variables. The spatial distribution of location solutions proves the previously generated hypothesis. It is evident that the optimal locations are different and largely depend on the underlaying spatiotemporal autocorrelation structure considered during the optimization process. The spatial similarity analysis was also conducted to objectively assess the difference between three solution sets generated and suggested that the spatial

---

**Figure 8.** Spatial distribution of optimized road weather information system (RWIS) locations with respect to three variables: (a) October 2016 and (b) January 2017. 
*Note: AT = air temperature; RST = road surface temperature; DPT = dew point temperature.*
The correlation between RST and DPT was lower than that of AT and RST, which is consistent with our hypothetical test results. These findings, undoubtedly, provide a stepping stone to the development of a multi-objective RWIS location allocation model by considering the varying degrees of representativeness of multiple weather variables.

**Conclusion and Recommendations**

This study aimed at developing an advanced location optimization model by using spatiotemporal semivariogram parameters combining both spatial and temporal effects of three key RWIS variables—AT, RST, and DPT. In the proposed method, a joint spatiotemporal semivariogram model was generated using three RWIS measurements; and the parameters of the model were used for determining optimal locations for RWIS stations. The location allocation problem was solved using a popular mathematical programming approach—an SSA algorithm that has been proven effective in and has gained recognition for RWIS location optimization problems. The methodological framework developed here builds on our previous effort by integrating the temporal

---

**Figure 9.** Similarity among optimal road weather information system (RWIS) station placements.

*Note:* AT = air temperature; RST = road surface temperature; DPT = dew point temperature.

**Figure 10.** Plot of objective function with respect to number of iterations for location optimization of Iowa.

*Note:* AT = air temperature; RST = road surface temperature; DPT = dew point temperature.
domain in the location allocation problem. Additionally, RWIS data for multiple weather variables was used to generate three sets of optimal location solutions. The similarity among the generated location solutions was also quantified in this study using a spatial similarity index. The key findings of this research are listed below:

- An advanced RWIS location allocation model was developed based on the premise that monitoring capabilities can be increased by minimizing the spatially averaged kriging variance in both space and time. The framework developed represents the first in the existing literature that attempts to determine optimal regional RWIS sampling design by taking into account both spatial and temporal attributes of multiple road weather variables.
- This study investigated and confirmed that the location solutions generated using AT, RST, and DPT were significantly different from one another. Their closeness was further quantified and objectively validated using a spatial similarity index.
- RWIS measurements of AT, RST, and DPT were processed and analyzed, and their correlations quantified, revealing that variables were more correlated with each other during mid-winter months than during shoulder months. Most notably, similarity between AT and RST was found to be significantly high.
- The effective coverage of RWIS measurements was determined by analyzing multiple critical weather variables. Comparatively, a higher spatial and temporal continuity range of autocorrelation was observed for mid-winter months than for shoulder months. The spatiotemporal continuity range (from the joint semivariogram model) was determined to be 17 km for the mid-winter month, and 7.5–10 km for the shoulder month.

Recommendations for further research are given below:

- This study covered a flatland region as the study area. Therefore, more case studies including wider regions can be examined for better understanding of the effect of spatiotemporal structure on RWIS location optimization.
- The study period of this research was limited to 2 months: October 2016 and January 2017. Lengthening the study period can potentially improve the level of confidence in the output.
- Three different weather variables were used in this analysis: AT, RST, and DPT. Thus, other RWIS measured data, that is, subsurface temperature, road surface conditions, and so forth could be included in the analysis to strengthen the output.
- Microclimatic factors could be considered by combining RWIS data with other road weather and surface data extracted from entities such as automated surface observing system (ASOS) and airport weather observation system (AWOS). The combined spatially rich data set allows the RWIS location optimization algorithm to consider both the local and regional weather characteristics in the regions of interest.
- Finally, a framework for a multi-criteria location optimization model should be established incorporating the joint semivariograms generated from weather variables.

Acknowledgment
The authors would like to thank Iowa DOT for providing the data necessary to complete this research.

Author Contributions
The authors confirm contribution to the paper as follows: study conception and design: S. Biswas, T. J. Kwon; data collection: S. Biswas; analysis and interpretation of results: S. Biswas, T. J. Kwon; draft manuscript preparation: S. Biswas, T. J. Kwon. All authors reviewed the results and approved the manuscript.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this
article: This research is funded by Natural Sciences and Engineering Research Council (NSERC).

**ORCID iDs**

Simita Biswas [ID](https://orcid.org/0000-0003-1587-3296)
Tae J. Kwon [ID](https://orcid.org/0000-0002-1700-2956)

**References**

1. White, S. P., J. E. Thornes, and L. Chapman. *A Guide to Road Weather Information Systems*, Version 2. University of Birmingham, UK, 2006, p. 83.
2. Pilli-Sihvola, E., P. Leviakangas, and R. Hautala. Better Winter Road Weather Information Saves Money, Time, Lives and the Environment. *Proc., 19th Intelligent Transport Systems World Congress, ITS*, Vienna, Austria, 2012.
3. Kwon, T. J., 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, 2017, pp. 691–699.
4. Manfredi, J., T. Walters, G. Wilke, L. F. Osborne, R. Hart, T. Incrocci, and C. Schmitt. *Road Weather Information System Environmental Sensor Station Siting Guide*, Version 2.0. No. FHWA-HOP-05-026. United States Department of Transportation, Washington, DC, 2008.
5. Kwon, T. J., and L. Fu. Evaluation of Alternative Criteria for Determining the Optimal Location of RWIS Stations. *Journal of Modern Transportation*, Vol. 21, No. 1, 2013, pp. 17–27.
6. Jin, P., A. Walker, M. Cebelak, and C. Walton. Determining Strategic Locations for Environmental Sensor Stations with Weather-Related Crash Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2014, 2440: 34–42.
7. Kwon, T. J., 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, 2017, pp. 57–71.
8. Zhao, L., S. Chien, J. Meegoda, Z. Luo, and X. Liu. Cost-Benefit Analysis and Microclimate-Based Optimization of a RWIS Network. *Journal of Infrastructure Systems*, Vol. 22, No. 2, 2016, p. 04015021.
9. Gräler, B., E. Pebesma, and G. Heuvelink. Spatio-Temporal Interpolation Using gstat. *The R Journal*, Vol. 8, No. 1, 2016, pp. 204–218.
10. Li, L., 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, 2017, p. 549.
11. Ahmed, S. O., R. Mazloum, and H. Abou-Ali. Spatiotemporal Interpolation of Air Pollutants in the Greater Cairo and the Delta, Egypt. *Environmental Research*, Vol. 160, 2018, pp. 27–34.
12. RESSTE Network. Analyzing Spatio-Temporal Data with R: Everything You Always Wanted to Know—But Were Afraid to Ask. *Journal de la Société Française de Statistique*, Vol. 158, No. 3, 2017, pp. 124–158.
13. Hu, D., H. Shu, H. Hu, and J. Xu. Spatiotemporal Regression Kriging to Predict Precipitation Using Time-Series MODIS Data. *Cluster Computing*, Vol. 20, No. 1, 2017, pp. 347–357.
14. Galars, D., and R. Angryk. Beyond Accuracy—A SMART Approach to Site-Based Spatio-Temporal Data Quality Assessment. *Intelligent Data Analysis*, Vol. 22, No. 1, 2018, pp. 21–43.
15. Yuan, Z., X. Zhou, and T. Yang. Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. *Proc., 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 2018, pp. 984–992.
16. Wang, X., L. Gu, T. J. Kwon, and T. Z. Qiu. Determining the Spatiotemporal Coverage of Road Weather Information Systems—A Case Study in Alberta, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, (No. 19-00551), 2019.
17. Hastie, T. J., and R. J. Tibshirani. *Generalized Additive Models*. Chapman & Hall/CRC, New York, NY, 1990.
18. Olea, R. A. *Geostatistics for Engineers and Earth Scientists*. Springer Science & Business Media, New York, NY, 2012.
19. Olea, R. A. A Six-Step Practical Approach to Semivariogram Modeling. *Stochastic Environmental Research and Risk Assessment*, Vol. 20, No. 5, 2006, pp. 307–318.
20. Bohling, G. Introduction to Geostatistics and Variogram Analysis. *Kansai Geological Survey*, Vol. 1, 2005, pp. 1–20.
21. Solana-Gutiérrez, J., and S. Merino-de-Miguel. A Variogram Model Comparison for Predicting Forest Changes. *Procedia Environmental Sciences*, Vol. 7, 2011, pp. 383–388.
22. Gething, P. W., 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, 2007, pp. 1337–1350.
23. Shekhar, S., H. Xiong, and X. Zhou. *Encyclopedia of GIS*. Springer, New York, NY, 2008.
24. Pebesma, E., and B. Gräler. *Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation*. Wiley, New York, NY, 2012.
25. Pebesma, E., and B. Gräler. *Introduction to Spatio-Temporal Variography*. University of Münster, Germany, 2018.
26. Pebesma, E., B. Gräler, and M. E. Pebesma. Package “gstat.” The Comprehensive R Archive Network (CRAN), 2019.
27. Biswas, S., M. Wu, S. J. Melles, and T. J. Kwon. Use of Topography, Weather Zones, and Semivariogram Parameters to Optimize Road Weather Information System Station Density across Large Spatial Scales. *Transportation Research Record: Journal of the Transportation Research Board*, 2019, 2673: 301–311.
28. Biswas, S., and T. J. Kwon. Developing Statewide Optimal RWIS Density Guidelines Using Space-Time Semivariogram Models. *Journal of Sensors*, Vol. 2020, 2020, pp. 1–13.
29. Revelle, C. S., H. A. Eiselt, and M. S. Daskin. A Bibliography for Some Fundamental Problem Categories in Discrete Location Science. *European Journal of Operational Research*, Vol. 184, No. 3, 2008, pp. 817–848.
30. Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, Vol. 220, No. 4598, 1983, pp. 671–680.
31. Van Groenigen, J. W., W. Siderius, and A. Stein. Constrained Optimisation of Soil Sampling for Minimisation of the Kriging Variance. *Geoderma*, Vol. 87, No. 3–4, 1999, pp. 239–259.
32. Van Groenigen, J. W., and A. Stein. Constrained Optimization of Spatial Sampling Using Continuous Simulated Annealing. *Journal of Environmental Quality*, Vol. 27, No. 5, 1998, pp. 1078–1086.
33. Brus, D. J., and G. B. Heuvelink. Optimization of Sample Patterns for Universal Kriging of Environmental Variables. *Geoderma*, Vol. 138, No. 1–2, 2007, pp. 86–95.
34. Heuvelink, G. B., D. J. Brus, and J. J. de Gruijter. Optimization of Sample Configurations for Digital Mapping of Soil Properties with Universal Kriging. *Developments in Soil Science*, Vol. 31, 2006, pp. 137–151.
35. Kwon, T. J., and L. Fu. RWIS Network Planning: Optimal Density and Location. No. Aurora Project 2010-04. Iowa Department of Transportation, Ames, 2016.
36. Asuero, A. G., A. Sayago, and A. G. Gonzalez. The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry*, Vol. 36, No. 1, 2006, pp. 41–59.
37. Pebesma, E. J. Multivariable Geostatistics in S: The gstat Package. *Computers & Geosciences*, Vol. 30, 2004, pp. 683–691.
38. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. https://www.R-project.org/.
39. Kwon, T. J., and L. Fu. Performance Evaluation of Road Surface Temperature Forecasts. *Canadian Journal of Civil Engineering*, Vol. 48, No. 5, 2021, pp. 532–539.
40. ESRI. *ArcGIS*, Version 10.4.1. ESRI, Redlands, CA, 2015.