Application and Integration of Lattice Data Analysis, Network *K*-Functions, and Geographic Information System Software to Study Ice-Related Crashes

Ghazan Khan, Kelvin R. Santiago-Chaparro, Xiao Qin, and David A. Noyce

Advances in geographic information system (GIS) software and exploratory spatial data analysis (ESDA) techniques give transportation safety engineers tools to observe and analyze safety-related data from a new perspective. This research takes the use of GIS software and ESDA techniques one step further by incorporating advanced statistical techniques for a more thorough and complex analysis of safety data. This is achieved by implementing a network-constrained cross *K*-function to analyze the relationship between bridges and the occurrences of ice-related crashes within a county. The counties in Wisconsin included in the analysis were selected through the use of the local Moran's *I* statistic; this statistic allows for the selection of counties within the same geographical area, which have similar parameters (in this case, ice-related crash rates). The objective of this research is to explore the relationship between ice-related crashes and bridges in counties that display similar ice-related crash rates, to compare and analyze winter maintenance techniques. The results identify clustering of ice-related crashes around bridges in four counties with similar ice-related crash rates in southeast Wisconsin. Similarly, two of four counties show clustering of ice-related crashes around bridges in northwest Wisconsin. These results make a strong case to suggest that counties in these regions should focus additional winter maintenance efforts at bridge locations. In addition, this research shows how the use of advanced spatial statistical techniques, particularly network-based statistics applied within a GIS environment, can be used as a unique and innovative approach toward safety data analysis.

Advances in geographic information system (GIS) software have provided transportation safety engineers with tools to observe and analyze safety-related data from a new perspective. Recently, there has been a boom in the use of exploratory spatial data analysis (ESDA) techniques for safety data analysis. This research takes advantage of expanding GIS capabilities and incorporates ESDA with advanced statistical techniques to take the analysis of safety data one step further.

G. Khan and K. R. Santiago-Chaparro, 2205 Engineering Hall; X. Qin, 1214 Engineering Hall; and D. A. Noyce, 1204 Engineering Hall, Traffic Operations and Safety (TOPS) Laboratory, Department of Civil and Environmental Engineering, University of Wisconsin–Madison, 1415 Engineering Drive, Madison, WI 53706. Corresponding author: G. Khan, gkhan@wisc.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2136, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 67–76.
DOI: 10.3141/2136-08

Winter maintenance costs can consume a large percentage of the budgets allocated to departments of transportation (DOTs) in northern states, depending on the intensity and severity of weather conditions. Winter maintenance decision making is one of the most complex tasks that DOTs in northern states have when assessing where to implement effective maintenance activities. Decision makers are faced with the challenge of how to optimize the use of continually decreasing resources. As problems become more complex, so have solutions. Fortunately, the use of GIS tools provides a powerful platform to perform the complex analyses that are needed to optimize resource use.

One variable that can be analyzed by the use of GIS and spatial statistical analyses is the location that should be the primary target of winter maintenance activities such as deicing and anti-icing, to reduce ice-related crashes. One of the most common approaches to identify locations for safety treatments is hotspot identification. Traditional hotspot identification techniques, however, do not have any statistical grounds. Specifically, the technologies used to identify hotspots are based on the locations of crashes, but the methods do not take into consideration whether the crashes are random events or the result of some underlying factors.

The objective of this research is to incorporate the use of advanced spatial statistical methods with GIS software to evaluate an innovative approach to safety data analysis. Moreover, the research is intended to provide winter maintenance personnel with a means of evaluating their activities in relation to specific locations on the system, through the results of spatial data analysis techniques coupled with safety (i.e., crash) data.

Through spatial pattern analysis of lattice data, counties with both statistically significant similar and statistically significant dissimilar ice-related crashes were identified. The analysis was performed for all counties in Wisconsin using the local Moran's *I* statistic, which identified eight counties demonstrating similar ice-related crash rates. This finding was the basis for comparing and contrasting local- and microscopic-level patterns of ice-related crashes in these counties. Spatial pattern analysis on a local level was performed using a network cross *K*-function, which identified the clustering of crashes around specific locations (i.e., bridges). This technique not only identified areas of hotspots for ice-related crashes but also enhanced understanding of the factors affecting these crashes, such as their proximity to the geometric features. Unlike a planar *K*-function, which analyzes patterns for data distributed in a planar space, a network cross *K*-function brings added accuracy to data analysis (e.g., for road crashes) that is inherently network-constrained.

LITERATURE REVIEW

The literature review looked at the state-of-the-practice of GIS in analyzing traffic safety data, advanced spatial statistical methods, and pattern analysis techniques for safety data. Recently, there have been tremendous advancements in GIS software capabilities and increases in the availability of spatial data sets, especially those that give point locations for each crash. GIS use has been most effective in analyzing point-based crash data, because it identifies spatial patterns of safety trends and issues that are otherwise difficult to observe from tabular data sets. Several studies have established spatial patterns in vehicle or pedestrian crashes for identification of critical locations (*1*). Kim and Yamashita analyzed spatial patterns of pedestrian crashes in Honolulu, Hawaii, by using *K*-means clustering techniques (*2*). These spatial patterns identified areas of high levels of pedestrian-involved crashes, which were present in light of various demographic features such as population or land use. Similarly, Levine et al. (*3*) conducted a spatial analysis of Honolulu crashes in the context of varying conditions and noted the limitations of "blackspot" analysis in describing the location and cause of different types of crashes. Thomas carried out a study for black zones (i.e., locations with high crash frequencies) and found several advantages in defining black zones using spatial autocorrelation and kernel methods on road segments (*4*). Abdel-Aty and Wang (*5*) studied the spatial effects of crashes at intersections along corridors.

The aforementioned research applied various methodologies to evaluate the spatial patterns of crashes alone, identifying potential hotspots or high crash locations at various scales. The aim of this research was to extend the spatial pattern analysis of crashes in conjunction with geometric feature locations to study the interactions between two point patterns. The idea was to determine the underlying factors and relationships between crashes and geometric features that lead to crashes.

Spatial statistical tools have been used for many years, especially in the fields of epidemiology and social sciences, to study the spatial variation and geographic dependencies in relevant data sets. Such data sets can occur anywhere in planar space and, hence, the methodologies have been developed accordingly. In the case of crash data analysis however, the assumption of planar space is no longer valid because distances are only relevant on a network. Therefore, spatial statistical techniques have to be modified to address the issue of network dependencies. Okabe and Yamada derived the *K*-functions and cross *K*-functions for a network in their groundbreaking research in 2001 (*6*). Yamada and Thill compared network and planar *K*-functions by analyzing crash data from New York to show how the assumptions of planar space are unsuitable for crash data analysis (*7*). In subsequent research, Yamada and Thill described another network-based *K*-function to identify local spatial patterns for crashes in the New York area (*8*).

This research advances the traditional hotspot analysis by making use of the cross *K*-function on a network to analyze the relationships between two point patterns: crashes and geometric features. The literature suggests that most research designed to analyze spatial point patterns has been focused on a particular scale: for example, at the citywide or statewide level. This research is unique in that it considers a statewide level through lattice data analysis to prioritize locations that were in turn further analyzed on an individual point scale for each county. This method provides the most comprehensive analysis on varying scales starting from the statewide level and scaling down to individual crash locations, a scope of research that is absent from the literature.

OBJECTIVE AND HYPOTHESIS

The objective of this research is to combine advanced spatial statistical methods with GIS functionality to analyze spatial patterns of safety data. The aim is to show the usefulness of these techniques by analyzing county-level data and identifying spatial patterns of ice-related crashes that have occurred in the selected counties. Researchers focused on the identification of specific features to better understand the underlying factors affecting those crashes, therefore providing the grounds to improve safety.

Ice-related crashes are selected for analysis because bridge decks and nearby locations are widely known to be prone to ice formation during the winter season. In fact, almost all the counties in Wisconsin do significant anti-icing or deicing at bridges and nearby locations (*9*). As a result, bridge locations can be considered one of the factors that can be evaluated using statistical methodologies that are able to detect clustering.

The clustering of ice-related crash patterns against bridge locations would provide evidence that these crashes are related to the location of the bridges. It would help prioritize locations for winter maintenance personnel and help them focus their activities at these locations, in turn leading to more effective, efficient, and proactive winter maintenance activities. Moreover, the identification of relationships between ice-related crashes and bridge locations for counties with similar ice-related crash rates would present a suitable basis for comparing these counties' ice-related crash patterns against bridge locations. This would help winter maintenance personnel to compare and contrast their activities across different jurisdictions and to make suitable improvements.

This research also expands the state-of-the-practice by using lattice-based pattern analysis and network-based point-pattern analysis together with GIS using safety data to identify areas where winter maintenance should be focused and to provide useful insight to winter maintenance personnel on how to prioritize and modify their winter maintenance activities. The procedures identified and streamlined in this research can easily be incorporated to study other types of safety data at any location.

DATA COLLECTION AND PROCESSING

The first step in the data collection process was to assemble the various data elements required for a state-level analysis. These state-level data were further segregated by county to provide a well-defined jurisdictional-based global picture of the Wisconsin data being analyzed. It also provides a global overview of the safety issues to be studied by this research, to identify areas for more detailed analysis. County-level analysis was the first step by which areas of similar safety performance in terms of ice-related crashes could be scrutinized for further microscopic analysis. Figure 1 shows a flowchart of each data set collection, processing using GIS analyses and subsequent statistical analyses.

Four years of Wisconsin crash data (2003 through 2006) were obtained for this research. The crash data were reduced to November 1 through April 30 of the following year, which is the typical wintertime period in Wisconsin. This period is also used by the Wisconsin Department of Transportation (WisDOT) for winter maintenance purposes. Crash data covering the winter seasons of 2003 to 2004, 2004 to 2005, and 2005 to 2006 were considered. Wisconsin crash data contains two sections pertinent to weather conditions at the time of the crash: weather conditions and road conditions. Ice-related

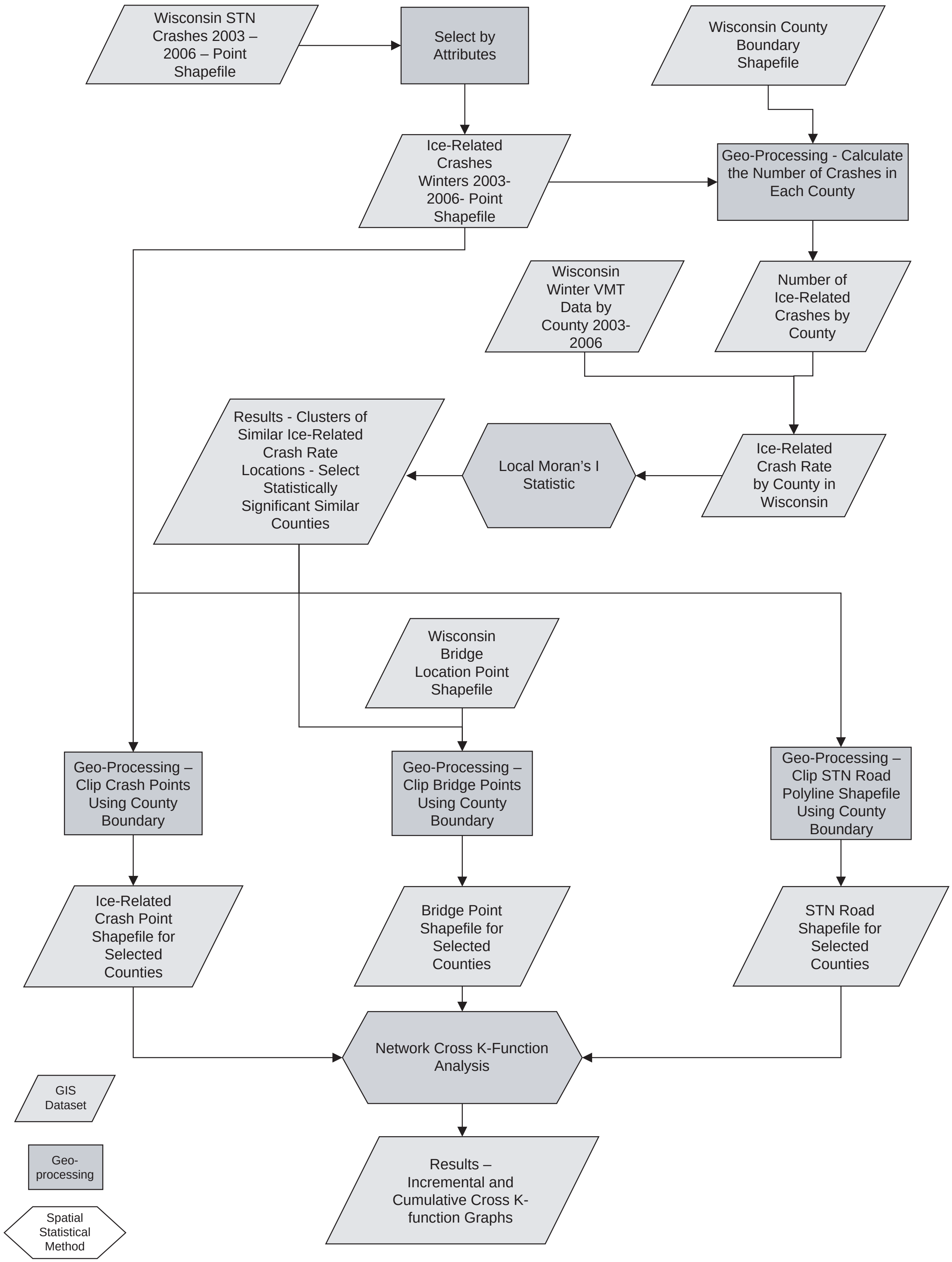

FIGURE 1 Flowchart of data collection and processing for Wisconsin ice-related crash analysis, 2003–2006 (STN = State Trunk Network).

crashes were identified for the three winter seasons as those crashes which occurred with ice on the pavement, sleet falling, or both. Along with the crash data, data on wintertime vehicle miles traveled (VMT) were obtained from the Wisconsin winter maintenance reports for the aforementioned three wintertime seasons. Ice-related crash rates were calculated to normalize crashes by some level of exposure to facilitate the comparisons between counties. An ice-related crash rate was defined as follows:

$$x_i = \frac{IC_i}{WV_i} \tag{1}$$

where

x_i = ice-related crash rate for county i,
IC_i = total number of ice-related crashes in county i, and
WV_i = exposure represented as 100 million VMT in entire winter season in county i.

Data on bridge locations and on roads comprising the Wisconsin State Trunk Network (STN) system were also obtained with the use of advanced geo-processing techniques in ArcGIS software, only for those counties that were analyzed in this research. Details of the data sets and processing procedures are presented in Figure 1. There were 66 bridges in Barron County, 38 in Bayfield County, 29 in Rusk County, 20 in Washburn County, 56 in Kenosha County, 51 in Ozaukee County, 51 in Racine County, and 160 in Waukesha County. Crash data were collected only for roads on the STN system because traffic-volume information was not available for local roads. The Wisconsin STN system consists of Interstate, U.S., and state highways. Moreover, exact geographic locations of crashes were required to conduct some of the point-pattern analysis (to be described further in this research), and these data were not available for all local roads. A pilot research project has recently been completed to fill this data gap with local roads (*10*). For the purpose of this research, data analysis was confined to crashes only on roads in Wisconsin's STN system. A shapefile of crash locations was generated by WisDOT by using intersection or milepost location, distance of the crash from that intersection or milepost location (in increments of one-hundredth of a mile), and STN-specific reference point tables that identify specific locations on the STN system, thus affording researchers accurate positions of the crashes.

METHODOLOGY

The first step in this research was to plot the ice-crash rates and then analyze the subsequent patterns on a statewide level, to identify counties that were part of a wider region displaying similar safety trends. These counties were then selected for microscopic-level analysis. Analyses of locations displaying similar global safety trends could then be compared, to identify potential differences in winter maintenance activities and procedures.

To analyze the patterns of spatially distributed features, overall patterns can be visually interpreted through frequency, mean, or proportion measurements on GIS-generated maps. Although ice-related crashes can easily be plotted statewide, the ability to visually discern spatial patterns and to identify areas with similar or dissimilar performance is limited. There is a need to identify statistical processes to quantifiably measure spatial patterns rather than a predefined ranking or number-based classification, because visual interpretations alone cannot provide conclusive results. Spatial patterns supported by statistically significant quantities that describe those patterns accurately, can resolve the issue by providing a quantifiable method of analysis.

Local Moran's *I* Statistic

There are several statistical techniques available for the analysis of spatial patterns of lattice data providing different answers based on desired results. Some techniques identify clusters of high or low attribute values (Getis-Ord Gi* statistic), while others identify clusters of similar or dissimilar values (Anselin's local Moran *I* statistic). Both of these statistics are part of local spatial autocorrelation statistics that can identify a local spatial clustering around an individual location, especially in cases where global statistics may fail to detect these patterns (*11*). The use of local spatial statistical techniques can discern spatial patterns that could be masked by global spatial autocorrelation statistics, and this method adds depth and significance to the results that could otherwise be a chance occurrence. With these requirements in mind, Anselin's local Moran's *I* statistic (I_i) was selected to analyze patterns of ice-related crashes on a statewide level for Wisconsin (*11*).

I_i identifies clusters of areas that have statistically significant similar or dissimilar values (*11*). Its output consists of a statistic value *I* and associated *Z* score for each feature in the study area. The resulting index value *I* for a specific feature indicates that it is clustered with other features with similar attribute values. A negative value for a feature indicates that the feature is clustered by dissimilar values, hence it is an outlier. *Z* scores are measures of standard deviation associated with a standard normal distribution calculated using the ratio of differences between observed and expected (mean) values, representing statistical significance of the index value. Anselin's research provides additional details on calculating expected values and *Z* scores (*11*). *Z* scores indicate whether the similarity or dissimilarity in attribute values between the feature and its neighbors is greater than one would expect simply by chance. The *Z* score can be interpreted similar to the index value. A low *Z* score indicates clustering of dissimilar values, while a high *Z* score indicates clustering of similar values. The more positive or negative the *Z* score, the more significant are the results. The I_i statistic can be presented as follows:

$$I_i = \frac{x_i - \bar{x}}{S^2} \sum_{j=1}^{N} w_{ij}\left(x_j - \bar{x}\right) \tag{2}$$

where

x_i = ice-related crash rate of site i,
x_j = ice-related crash rate of neighboring locations to site i,
w_{ij} = spatial weight matrix for all sites j, and
N = number of weighted points, each representing ice-related crash rate for each county

and

$$S^2 = \frac{\sum_{j=1, j\neq i}^{N} x_j^2}{N-1} - \bar{x}^2 \tag{3}$$

In Equation 2, i is the site with an attribute value x_i where the I_i statistic is being calculated and x_j are the neighboring locations with similar or dissimilar attribute values. For analysis, the attribute values were multiplied by the spatial weight matrix w_{ij} that defines which locations were included in the analysis and the corresponding weight. Locations i and j in the above equations are depicted by the geometric centroids of individual counties because the data were aggregated at a county level. The attribute values used at these sites were ice-related crash rates, which have already been defined.

In any type of clustering analysis, one of the most important questions is the conceptualization of spatial association among the features, or the construction of the spatial weight matrix W. For the purposes of this research, it was decided that the proposed choice of weight matrix would be based on an inverse distance relationship, which means that the influence of spatial relationships decreases as an inverse function of increasing distance. The choice was based on the premise that features close to one another are more similar than features further away, although further research will be required to bring some objectivity to this subjective selection.

Network Cross *K*-Function

The second step of analysis was based on the analysis of two point patterns and their interrelationship. The idea was to analyze the spatial patterns of ice-related crashes for counties displaying similar safety trends. The patterns were analyzed for each county identified as belonging to a statistically significant cluster of counties displaying similar safety trends. This would enable the comparison of the distribution of ice-related crashes against bridge locations in each county. As mentioned in the literature review, there are a number of point-pattern analysis methodologies that have been developed for use in the field of epidemiology and in the social sciences. The *K*-function method is one such procedure, which has been most widely used (*12*). However, this method is based on the assumption that data are distributed in planar space. This assumption is violated for the purposes of crash data analysis, hence the cross *K*-function for network was selected as the appropriate method for this research (*13*).

The network cross *K*-function describes the relationship between the patterns of two sets of points, for example $A = \{a_1, a_2, \ldots, a_{n_a}\}$ and $B = \{b_1, b_2, \ldots, b_{n_b}\}$, placed on a finite planar network (L_T), and shows whether the set of points B affects the location of the set of points A (*6*). To examine this effect, the null hypothesis is that the set of points A is distributed randomly according to the binomial point process regardless of the location of the set of points B. If this hypothesis is rejected, it can be reasoned that the location of the set of points B affects the distribution of the set of points A. No assumption is made with respect to the distribution of points B (*6*). The cross *K*-function can now be defined as follows:

$$K^{ba}(t) = \frac{1}{\rho_a} E\begin{pmatrix}\text{number of points of A within} \\ \text{network distance '}t\text{' of a point } b_i \text{ in } B\end{pmatrix} \tag{4}$$

where

A = set of point locations of ice-related crashes on the STN roads in each county,
B = set of point locations of bridges on STN roads in each county,
$E(\ )$ = expected value of A following binomial point process, with respect to $b_1, \ldots, b_{n_b}$ ($b_i \in B$),
ρ_a = density of points of A, which is equal to $n_a/|L_T|$,
n_a = total number of ice-related crashes,
L_T = finite planar network of STN roads in each county, and
$K^{ba}(t)$ = network cross *K*-function of A relative to B, for the binomial point process.

The results of the observed network cross *K*-function can be plotted on a graph that shows the clustering or dispersion of points at various distance scales. The expected value can also be plotted on the graph to show the upper and lower 5% bounds and show the statistical significance of the observed network cross *K*-function at the 95% confidence level. If the line of observed values lies above the upper 5% line, the pattern is said to be statistically significant clustered. If the observed line lies below the lower 5% line, the pattern is statistically significant dispersion. If the observed line lies within the upper and lower bound lines, there is no significant relationship between the two point patterns and the points are distributed independently of each other. A more thorough discussion of the network cross *K*-function can be found in the literature (*13*).

RESULTS AND DISCUSSIONS

Local Moran's *I* Analysis Results

The first step analyzed the safety performance of counties in terms of ice-related crashes. The goal was to identify counties with similar safety performances so that the results of local-level analysis conducted for those counties could be compared. Ice-related crash rates as defined in previous sections were calculated for each county in Wisconsin on the basis of crash and winter VMT data for the three winter seasons between 2003 and 2006. Table 1 presents winter VMT, the

TABLE 1 Results of Local Moran's *I* Analysis for Selected Statistically Significant Counties in Northwest and Southeast Regions in Wisconsin

County	Winter VMT (in 100 millions, for 3 years)	Ice Crashes (IC_i, for 3 years)	Percentage of Ice Crashes (%)	Crash Rate (x_i, by 100 million VMT)	Local Moran's I (I_i)	Moran's Z-Score
Northwest						
Barron	7.774	122	26.64	15.69	0.16	2.02
Bayfield	3.069	56	23.93	18.24	0.19	2.22
Rusk	2.307	41	23.70	17.77	0.14	1.97
Washburn	3.715	70	19.83	18.84	0.22	2.60
Southeast						
Kenosha	20.700	157	6.90	7.58	0.35	2.63
Ozaukee	12.958	37	4.12	2.86	0.64	5.92
Racine	22.953	152	5.62	6.62	0.48	3.50
Waukesha	57.315	258	5.36	4.50	0.602	5.51

number of ice-related crashes, crash rates, the percentage of ice-related crashes, and the results of local Moran's *I* analysis for the selected counties. The ice-related crash rates (per 100 million VMT) were plotted on a map for visual interpretation, which is presented in Figure 2*a.* Although the figure presents a fair picture of how the crash rates were distributed among the counties, it is difficult to discern any consistent spatial patterns from this figure alone. Moreover, the mapping of crash rates alone does not contain any statistical significance as to which counties are more similar than others. Any area clusters which display similar or dissimilar crash rates cannot be visually discerned in the absence of any statistical evidence.

To increase the statistical sensitivity of the selection of counties with similar ice-related crash rates, the local Moran's *I* statistic was used. The results of the local Moran's *I* statistic are displayed in Figure 2*b.* Counties with a *Z* score of greater than +1.96 represent locations that are part of statistically significant clusters of similar ice-related crash rates at a 95% confidence level, and vice versa. Figure 2*b* also includes the actual value of ice-related crash rates for each county as represented by dots the size of which is proportional to that county's ice-related crash rate, with larger dots representing higher rate values. Results identified four clusters in different regions of Wisconsin that display statistically significant similar ice-related crash rate values. These regions are located roughly in the northwest, southeast, north-central, and far-western regions of Wisconsin. Although there are some counties located next to each other that display similar safety trends, they are not part of a statistically significant cluster due to variability in the ice-related crash rates in the overall proximity of those counties.

The results of the local Moran's *I* analysis were used to select counties for which the network cross *K*-function analysis was conducted, and to identify the relationship between ice-related crashes and bridge locations. Although clusters in four different regions (each consisting of between two and eight counties) were identified, two regions with four counties each were selected for further analysis because they were part of the biggest clusters yielding the greater number of counties. Moreover, the geographic location of the clusters would be a good representation of how winter weather varies between the northern and southern parts of Wisconsin. The two regions representing a total of eight counties were in northwest and southeast Wisconsin (four counties in each region). The four northwest counties are shown in Figure 2*c,* and the four southeast counties are shown in Figure 2*d.* It can be seen from Table 1 that although the counties display varying ice-related crash and winter VMT trends, their ice-related crash rates are quite similar. This provided the basis for comparing counties' winter maintenance strategies for countering ice-related crashes, especially in the form of proactive anti-icing winter maintenance activities.

Network Cross *K*-Function Analysis Results

The network cross *K*-function analysis was conducted using the SANET tool developed by Okabe et al. (*14*). The aim was to identify clusters of ice-related crashes as well as to study the relationship between ice-related crashes and the location of bridges in each county. Figure 3 and Figure 4 show the results of network cross *K*-function analysis for counties selected from, respectively, the northwest and southeast region. For each county, there are two graphs showing the results of incremental and cumulative cross *K*-function values up to a distance of 1 km from either side of a bridge. On all of the 16 graphs in Figures 3 and 4, for all bridges in a particular county, the *x*-axis shows the increasing distance away from a bridge location (in m). On the eight left-side graphs in Figures 3 and 4, the y-axis shows the number of crashes observed within each distance increment; on the eight right-side graphs in Figures 3 and 4, the *y*-axis shows the cumulative number of crashes. The graphs indicate the relationship among bridge locations in individual counties and whether ice-related crashes cluster significantly within 1 km of either side of the bridges.

Results from Northwest Counties

Figure 3 shows the results of the network cross *K*-function analysis conducted for the four counties selected from the northwest region of Wisconsin. Figures 3*a* and 3*b* (for Barron County) and Figures 3*c* and 3*d* (for Washburn County) display statistically significant clustering of ice-related crashes around bridge locations in those counties. In particular, Barron County shows a high clustering as depicted by the large spike in Figure 3*a* of the observed cross *K*-function line within the first 100 meters. Conversely, Rusk County (as shown in Figures 3*e* and 3*f*) and Bayfield County (as shown in Figures 3*g* and 3*h*), show no significant clustering of ice-related crashes around bridge locations. For these two counties, although the observed *K*-function line is above the mean line at certain distance increments (which suggests a clustering tendency), the results are inconclusive at a 95% confidence level.

The results of the *K*-function analysis for the northwest counties suggest that bridge locations in Washburn and Barron Counties are more prone to the occurrence of ice-related crashes than are the locations in Rusk and Bayfield Counties. Given the similar safety performance of these counties in the rates of ice-related crashes, as seen in the crash-rate data listed in Table 1 for these four counties, the differences in the occurrence of ice-related crashes at bridge locations are clear. The results provide conclusive evidence that Washburn and Barron Counties should focus additional maintenance attention on bridge locations. Moreover, the results also provide an opportunity for the counties' personnel to compare and contrast their winter maintenance activities in relation to bridge locations, to improve their individual county's results. Several reasons could account for the differences in patterns, including differences in winter maintenance techniques and priorities, specifically anti-icing versus deicing strategies.

Results from Southeast Counties

Figure 4 shows the results of network cross *K*-function analysis conducted for the four counties selected from the southeast region of Wisconsin. Figures 4*a* through 4*h* show statistically significant clustering of ice-related crashes around bridge locations in all four counties selected in the southeast region. Clusters are close to the bridges' location: almost within the first 50 m on either side of the bridges. Moreover, the clustering tends to become insignificant quickly as distance from the bridges increases. These results are consistent among all four counties, similar to the ice-related crash rates.

The results of the network cross *K*-function analysis for these southeast counties suggest a significant relationship between the occurrence of ice-related crashes and bridge locations. The results provide conclusive evidence that Ozaukee, Waukesha, Racine, and Kenosha Counties should focus additional maintenance efforts on bridge locations to reduce the occurrence of ice-related crashes.

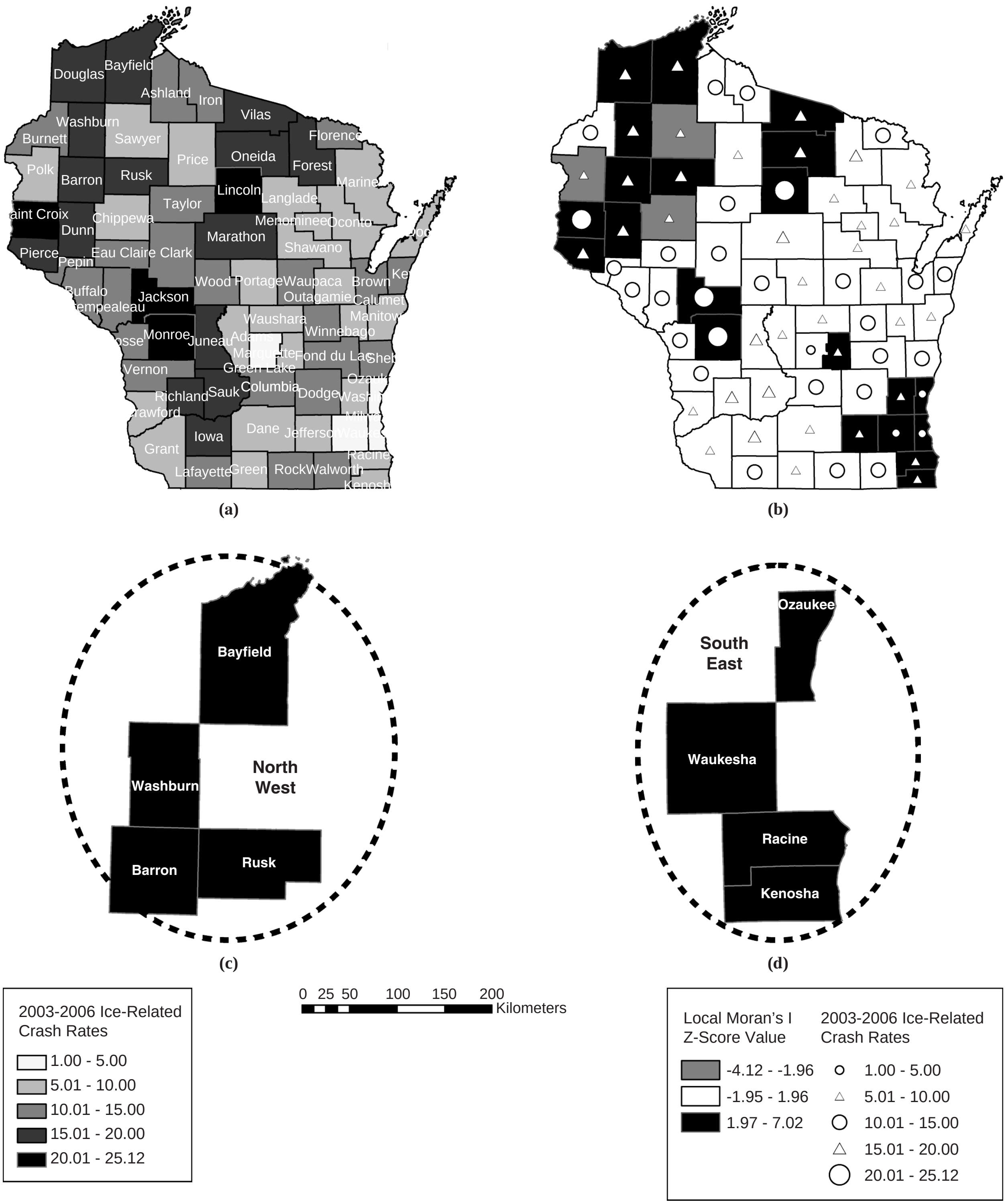

FIGURE 2 Local Moran's *I* analysis results: (*a*) ice-related crash rates for Wisconsin counties, 2003–2006 winter crash data; (*b*) local Moran's *I* analysis *Z*-score values for Wisconsin counties; (*c*) selected counties from cluster of statistically significant counties of similar ice-related crash rates in northwest region; and (*d*) selected counties from cluster of statistically significant counties of similar ice-related crash rates in southeast region.

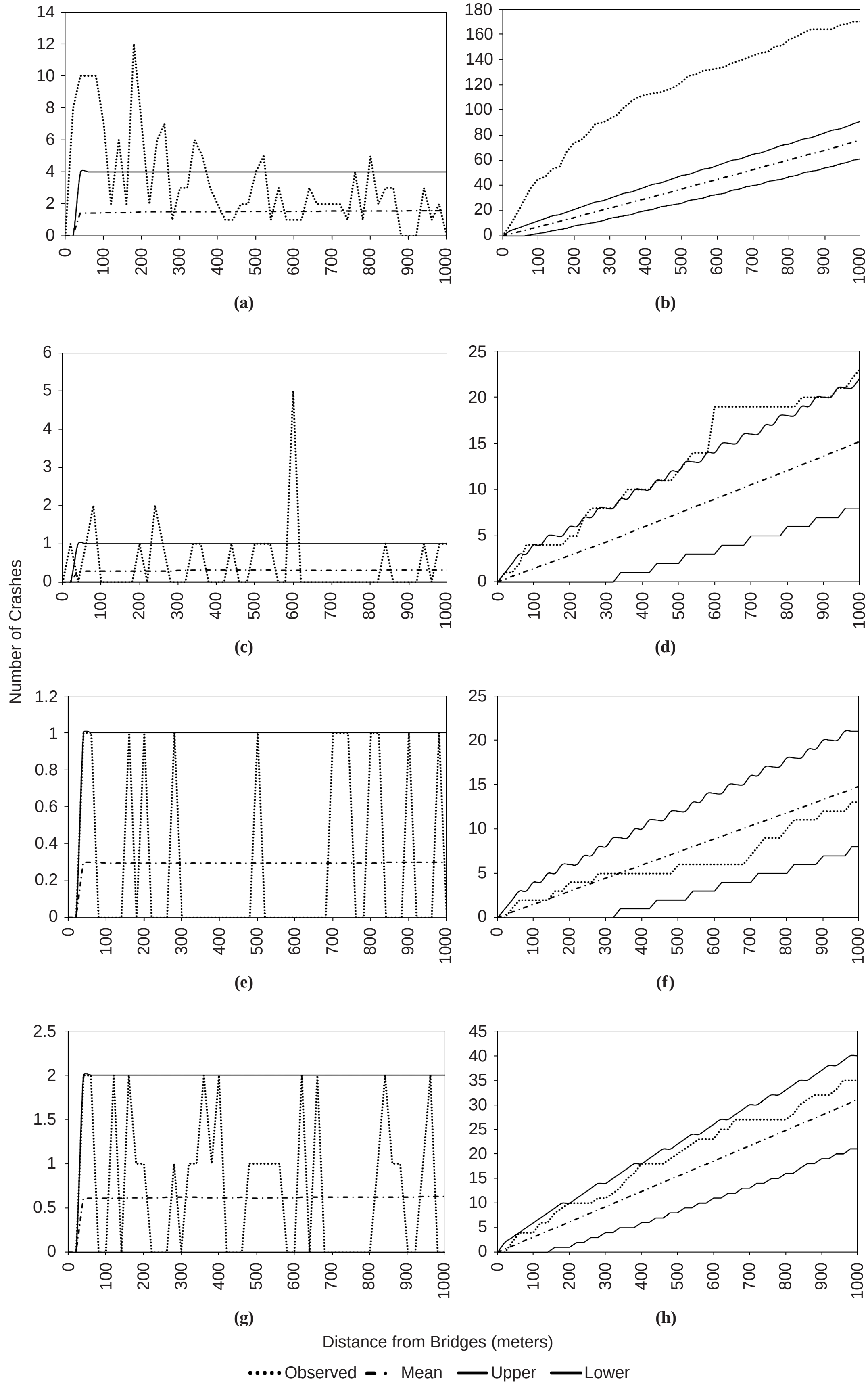

FIGURE 3 Network Cross *K*-function results of statistically significant counties selected from the northwest region of Wisconsin: (*a*, *b*) Barron, (*c*, *d*) Washburn, (*e*, *f*) Rusk, and (*g*, *h*) Bayfield.

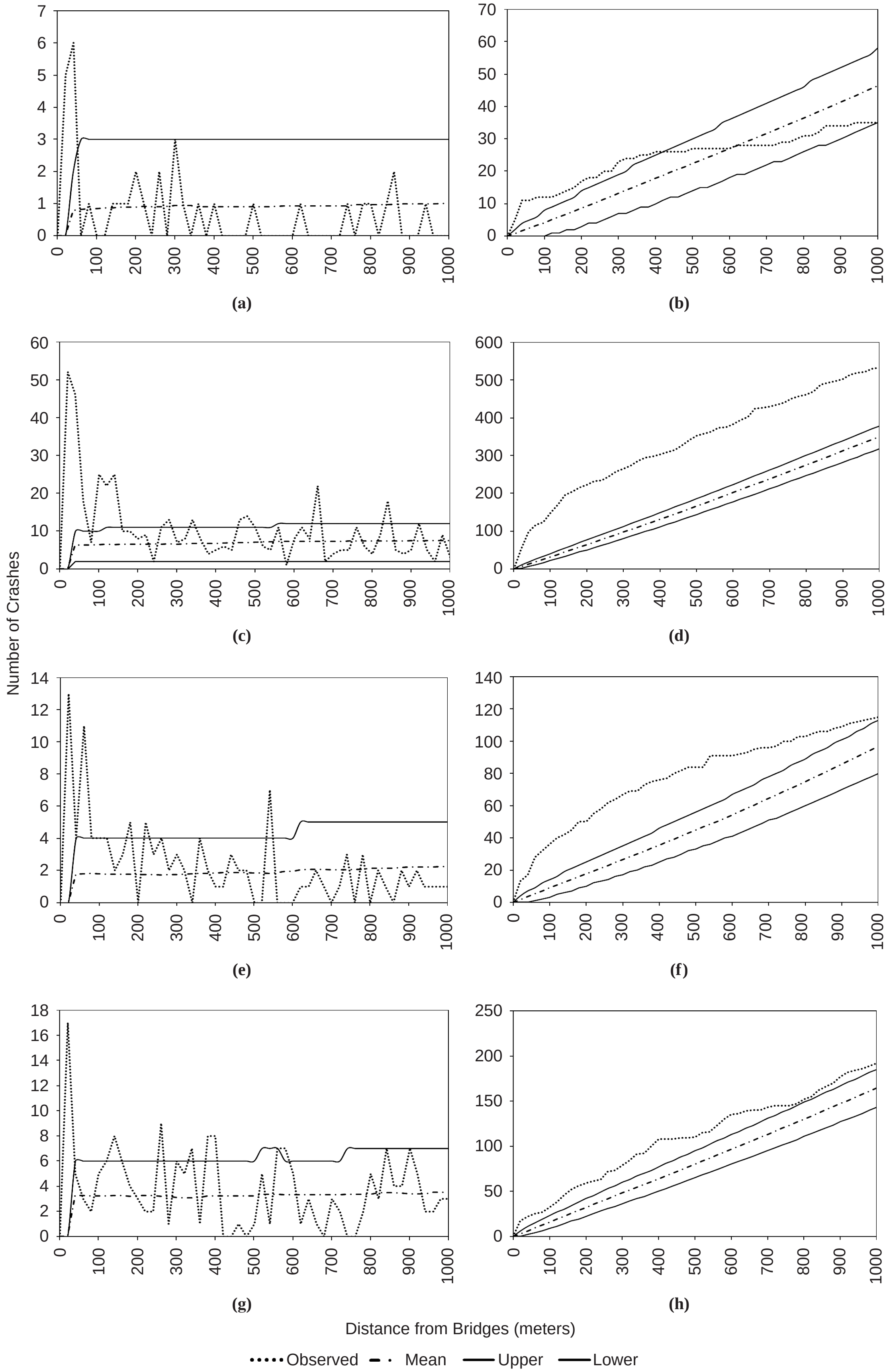

FIGURE 4 Network Cross *K*-function results of statistically significant counties selected from the southeast region of Wisconsin: (*a*, *b*) Ozaukee, (*c*, *d*) Waukesha, (*e*, *f*) Racine, and (*g*, *h*) Kenosha.

CONCLUSIONS

This research demonstrates the innovative application, as well as the usefulness, of integrating GIS-based data with advanced spatial statistical techniques for the analysis of safety data for winter maintenance purposes. The use of network-based statistical methods such as the cross *K*-function was the most significant improvement in this research. As previously mentioned, most spatial statistical methodologies have been developed for data sets distributed in planar space. The assumption of planar space is violated with crash data; hence, the use of network-based methods takes on added significance, especially when analyzing the microscopic-level components of a large-scale system, as is the case when looking at individual features over several counties.

The aforementioned procedures bring an added dimension of accuracy to the analysis of safety data (which had been previously missing) and can be particularly useful to support winter maintenance decisions. The use of safety data when evaluating winter maintenance alternatives provides a different perspective to decision making on issues of winter maintenance, one that has not been previously explored. The results of this research were not only based on visual maps and interpretations, but they also incorporate the use of advanced statistical methodologies. Analysis at the levels shown in this research is rarely found in the literature. In fact, most studies available focus their analysis at larger scales such as streets or county levels due to data or computational limitations.

Research findings show patterns of ice-related crashes in relation to bridges that are considered ice-prone locations and that are the focus of counties' winter maintenance activities such as anti-icing and de-icing. This research adds to that knowledge by providing statistical measurements that suggest that ice-related crashes cluster around bridges at several locations. Thus, in those counties where ice-related crashes cluster around bridge locations, winter maintenance activities should be focused on bridges to improve safety. Results that indicate bridges where ice-related crashes cluster as hotspots, will enable stakeholders to focus their maintenance and safety improvements at locations with those features that cause the clustering of crashes. County winter maintenance personnel can use these results to improve current winter maintenance policies and implement proactive measures such as anti-icing at bridges. Moreover, the fact that patterns of ice-related crashes were analyzed for counties in the same region with similar ice-related crash rates provides a basis for winter maintenance personnel to compare and assess the differences in winter maintenance techniques.

Although this research was focused on applying methodologies to identify crash clustering around bridges, the set and sequence of procedures used in this research are not limited to analysis of weather-related crashes. This methodology can be easily applied to other types of crash data either by individual type or severity, against different geometric features such as intersection locations and segment midpoints. Tools used in this research are readily available online and require only basic GIS knowledge and the use of ArcGIS software.

ACKNOWLEDGMENTS

The network cross *K*-function analysis was conducted by using the SANET tool developed by Atsuyuki Okabe, Kei-ichi Okunuki, and Shino Shiode at the Center for Spatial Information Science at the University of Tokyo. The authors thank them for providing the tool and for all their help and necessary guidance in implementing the use of the tool.

REFERENCES

1. Jones, A. P., I. H. Langford, and G. Bentham. The Application of K-Function Analysis to the Geographical Distribution of Road Traffic Accident Outcomes in Norfolk, England. *Social Science and Medicine,* Vol. 42, Issue 6, 1996, pp. 879–885.
2. Kim, K. E., and E. Y. Yamashita. Using K-Means Clustering Algorithm to Examine Patterns of Pedestrian Involved Crashes in Honolulu, Hawaii. *Journal of Advanced Transportation,* Vol. 41, No. 1, 2007, pp. 681–689.
3. Levine, N., K. E. Kim, and L. H. Nitz. Spatial Analysis of Honolulu Motor Vehicle Crashes: I. Spatial Patterns. *Accident Analysis and Prevention,* Vol. 27, Issue 5, October 1995, pp. 663–674.
4. Thomas, I. Spatial Data Aggregation: Exploratory Analysis of Road Accidents. *Accident Analysis and Prevention,* Vol. 28, Issue 2, 1996, pp. 251–264.
5. Abdel-Aty, M., and X. Wang. Crash Estimation at Signalized Intersections Along Corridors: Analyzing Spatial Effect and Identifying Significant Factors. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1953,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 98–111.
6. Okabe, A., and I. Yamada. The K-Function Method on a Network and Its Computational Implementation. *Geographical Analysis,* Vol. 33, No. 3, 2001, pp. 271–290.
7. Yamada, I., and J. Thill. Comparison of Planar and Network K-Functions in Traffic Accident Analysis. *Journal of Transport Geography,* Vol. 12, 2004, pp. 149–158.
8. Yamada, I., and J. Thill. Local Indicators of Network Constrained Clusters in Spatial Point Patterns. *Geographical Analysis,* Vol. 39, 2007, pp. 268–292.
9. Martinelli, J. T. *Annual Winter Maintenance Report 2002–2003 Season.* Bureau of Highway Operations, Winter Operations Unit, July 2003. https://trust.dot.state.wi.us/extntgtwy/dtid_bho/extranet/winter/reports/reports.sht. Accessed June 23, 2008.
10. Dutta, A., S. T. Parker, X. Qin, Z. Qiu, and D. A. Noyce. System for Digitizing Information on Wisconsin's Crash Locations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019,* Transportation Research Board of the National Academies, Washington, D.C. 2007, pp. 256–264.
11. Anselin, L. Local Indicators of Spatial Association (LISA). *Geographical Analysis,* Vol. 27, 1995, pp. 93–115.
12. Ripley, B. D. Modeling Spatial Patterns. *Journal of Royal Statistics,* Vol. 39, 1977, pp. 172–212.
13. Okabe, A., and I. Yamada. *The K-Function Method on a Network and Its Computational Implementation: Center for Spatial Information Science (CSIS) Discussion Paper 26.* Center for Spatial Information Science, University of Tokyo, April 20, 2000.
14. Okabe, A., K. Okunuki, and S. Shiode. SANET: A Toolbox for Spatial Analysis on a Network. *Geographical Analysis,* Vol. 38, No. 1, 2006, pp. 57–66.

The Statistical Methodology and Statistical Computer Software in Transportation Research Committee sponsored publication of this paper.