Application and Integration of Lattice Data Analysis, Network K-functions, and GIS to Study Ice-related Crashes

By

Ghazan Khan[1] and **Kelvin R. Santiago-Chaparro**
Graduate Researchers
Traffic Operations and Safety (TOPS) Laboratory
Department of Civil and Environmental Engineering
University of Wisconsin-Madison
2205 Engineering Hall
1415 Engineering Drive
Madison, WI 53706
Phone: (608) 698-8519
Fax: (608) 262-5199
Email: gkhan@wisc.edu

Xiao Qin, Ph.D., P.E.
Associate Researcher
Traffic Operations and Safety (TOPS) Laboratory
University of Wisconsin-Madison
1214 Engineering Hall
1415 Engineering Drive
Madison, WI 53706
Phone: (608) 262-3649
Fax: (608) 262-5199
Email: xqin@engr.wisc.edu

David A. Noyce, Ph.D., P.E.
Associate Professor
Director, Traffic Operations and Safety (TOPS) Laboratory
University of Wisconsin-Madison
Department of Civil and Environmental Engineering
1204 Engineering Hall
1415 Engineering Drive
Madison, WI 53706
Phone: (608) 265-1882
E-mail: noyce@engr.wisc.edu

Review Committee: ABJ80
Prepared for the 88th Annual meeting of the
Transportation Research Board, Washington, D.C. January, 2009
Length of Paper: 5561 words, 4 Figures + 1 table @ 250 words each
6811 equivalent words

(1) Corresponding author

ABSTRACT

Advances in GIS software and Exploratory Spatial Data Analysis Techniques (ESDA) have provided transportation safety engineers with tools to observe and analyze safety-related data from a new perspective. The presented research takes GIS and ESDA one step further and incorporates the use of advanced statistical techniques for a more thorough and complex analysis of safety data. This is achieved through the implementation of a network constrained cross K-function to analyze the relationship between occurrences of ice-related crashes and bridges within a county. The counties included in the analysis were selected through the use of Local Moran's I statistic, which allowed the selection of counties within the same geographical area that are similar in terms of a parameter, in this case the ice-related crash rates. The objective was to explore the relationship between ice-related crashes and bridges for counties displaying similar ice-related crash rates to compare and analyze winter maintenance techniques.

The results identified clustering of ice-related crashes around bridges in four counties of similar ice-related crash rates in the Southeast region. Similarly, two out of four counties showed clustering of ice-related crashes around bridges in the Northwest region. From these results, there is a strong case to suggest that counties in these regions should focus additional winter maintenance efforts at bridge locations. Furthermore, this research showed 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 towards safety data analysis.

INTRODUCTION

Advances in Geographical 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 techniques (ESDA) for safety data analysis. The presented research takes advantage of the expanding GIS capabilities and incorporates ESDA with advance statistical techniques to take the analysis of safety data one step further.

Winter maintenance costs can consume a large percentage of allocated budgets of Northern states DOTs, depending upon the intensity and severity of weather conditions. Winter maintenance decision making is one of the most complex tasks that a Northern state Department of Transportation (DOT) deals with in terms of where to implement effective maintenance activities. Decision makers are thus 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 needed to optimize resource use.

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

The objective of this research was to incorporate the use of advanced spatial statistical methods with GIS to evaluate an innovative approach of safety data analysis. Moreover, it was also intended to provide winter maintenance personnel with 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 the state of Wisconsin using Local Moran's I statistic, identifying eight counties showing similar ice-related crash rates. This provided the basis for comparing and contrasting the local/microscopic level patterns of ice-related crashes in these counties. Spatial pattern analysis on a local level was performed through the use of a network cross K-function which identified the clustering of crashes around specific locations, specifically bridges. This not only identified areas of hotspots for ice-related crashes but brought a measure of understanding towards the factors affecting these crashes, such as their proximity to the aforementioned geometric features. Furthermore, unlike the planar K-function which analyzes patterns for data distributed in a planar space, a network cross K-function brings added accuracy to data analysis that is inherently network-constrained, for example road crashes.

LITERATURE REVIEW

A literature review was completed which looked at the state-of-the-practice of GIS in traffic safety data analysis, advanced spatial statistical methods, and pattern analysis techniques for safety data. The recent past has seen tremendous advancements in GIS software capabilities and an increase in the availability of spatial datasets, especially point datasets representing a point location for each crash. GIS use has been most effective in analyzing point crash data as it identifies spatial patterns of safety trends and issues which are otherwise difficult to observe from tabular datasets. Several studies have been conducted to establish 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 using K-means clustering techniques (*2*). These spatial patterns identified areas of high pedestrian crashes that were present in light of various demographic features such as population or land-use. Similarly, Levine 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 causation of different types of crashes (*3*). Thomas carried out a study for black zones and found several advantages in defining black zones (i.e., high crash frequency locations) using spatial autocorrelation and kernel methods on road segments (*4*). Abdel-Aty studied the spatial effects of crashes at intersections along corridors (*5*).

The research previously cited applied various methodologies to evaluate the spatial patterns of crashes alone, indentifying 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 leading to the causation of the 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 datasets. Such datasets can occur anywhere in planar space hence the methodologies have been developed accordingly. However, in the case of crash data analysis, the assumption of planar space is no longer valid because distances are only relevant on a network. Therefore, spatial statistical techniques had to be modified to address the issue of network dependencies. Yamada and Okabe derived the K-functions and cross K-functions for a network in their ground breaking research in 2001 (*6*). Yamada and Thill provided a comparison of 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 indentify 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. Literature suggests that most research designed to analyze spatial point patterns was focused on a particular scale, for example at a city or statewide level. This research is unique in that it considers a statewide level through lattice data analysis to prioritize locations that were further analyzed on an individual point scale for each county. This method provided the most comprehensive analysis on varying scales starting from the statewide level and scaling down to individual crash locations, void in the literature.

OBJECTIVE AND HYPOTHESIS

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

Ice-related crashes were 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 the state of Wisconsin perform significant levels of anti-icing or de-icing treatment at bridges and/or nearby locations (*9*). As a result, bridge locations can be considered one of the factors that can be evaluated using the 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 winter maintenance activities at such locations leading to more effective, efficient, and pro-active 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 the patterns of these counties with each other. This would help winter maintenance personnel compare and contrast their activities across different jurisdictions and make suitable improvements.

This research also furthers 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 of winter maintenance focus and provide useful insight to winter maintenance personnel in terms of prioritizing and modifying 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 various data elements required for analysis on a state level. The state was further segregated by county because it provided a well-defined jurisdictional-based global picture of the state in terms of data being analyzed. It also provided a global overview of the safety issue studied here to identify focus 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 dataset collection, processing using GIS analyses, and subsequent statistical analyses.

Four years of Wisconsin crash data (2003 through 2006) were obtained for the purpose of this research. The crash data were reduced to November 1st through April 30th of the following year which is the typical time period of winter season in Wisconsin. This time frame is also used by the Wisconsin Department of Transportation (WisDOT) for winter maintenance purposes. Crash data covering three winter seasons of 2003-2004, 2004-2005, and 2005-2006 were considered. Wisconsin crash data contains two sections regarding weather conditions at the time of the crash, namely “weather conditions” and “road conditions”. Ice-related crashes were identified for the three winter seasons as those crashes which occurred with ice on the pavement, sleet falling, or both. Alongside the crash data, winter vehicle miles travelled (VMT) data were also obtained from the Wisconsin winter maintenance reports covering the previously mentioned three winter seasons. An ice-related crash rate was calculated to normalize crashes by some level of exposure to facilitate the comparisons between counties. An ice-related crash rate was defined as follows:

$$x_i = \frac{IC_i}{WV_i} \qquad (1)$$

Where:
x_i = Ice-related crash rate for county *i*,
ICi = 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*.

Bridge location data and State Trunk Network (STN) system roads were also obtained for the entire state of Wisconsin in the form of shapefiles using advanced geo-processing techniques in ArcGIS software, only for those counties which were analyzed in this research. Details of the datasets and processing procedures are presented in Figure 1. There were 66, 38, 29, 20, 56, 51, 51, and 160 bridges each for Barron, Bayfield, Rusk, Washburn, Kenosha, Ozaukee, Racine, and Waukesha counties, respectively. Note that crash data were only collected for the STN system because traffic volume information was not available for local roads. The Wisconsin STN system consists of the Interstate, US, and State highways. Moreover, exact geographic locations of crashes were required to conduct some of the point pattern analysis described further in this research, which were not available for all local roads. A pilot research project has recently been completed to overcome this local road limitation (*10*). For the purpose of this research, the data analysis was confined to crashes that occurred only on the Wisconsin STN roads. A shapefile of crash locations was generated by the Wisconsin Department of Transportation (WisDOT) using intersection or milepost location, distance of the crash from intersection/milepost in increments of one hundredth of a mile, and STN specific reference point tables identifying specific locations on the system, thus allowing the researchers with an accurate position of the crashes.

METHODOLOGY

The first step in this research was to plot the ice-crash rates and analyzing their 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. Moreover, analysis of locations displaying similar global safety trends could then be compared to identify potential differences in winter maintenance activities and procedures.

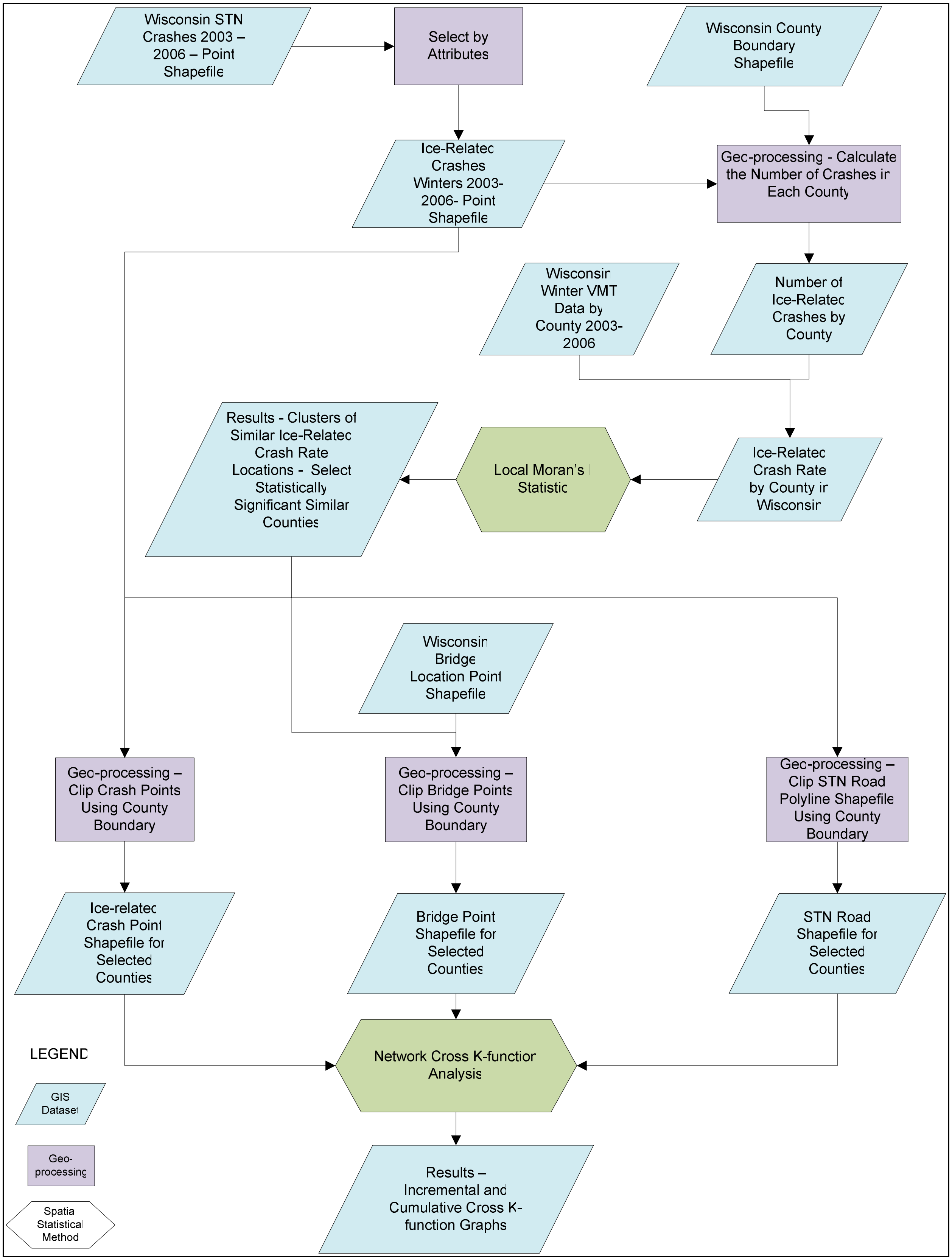

Figure 1 Flowchart of data collection and processing for Wisconsin ice-related crash analysis 2003-2006

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 could easily be plotted statewide, the ability to visually discern spatial patterns and to identify areas which have similar or dissimilar performance was limited. There was a need to identify statistical processes that can provide a quantifiable measure of spatial patterns rather than a pre-defined ranking or number based classification because visual interpretations alone cannot provide conclusive results. Spatial patterns supported by statistically significant quantities, which 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 analysis of spatial patterns of lattice data providing different answers based on desired results. Some 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 which can identify the 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 which could get masked by global spatial autocorrelation statistics and adds depth and significant to the results which could otherwise be a chance occurrence. With these requirements in mind, the 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*). 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 the results are. The I_i statistic can be presented as follows:

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

Where:

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

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.

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 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 equation are depicted by the geometric centroids of individual counties since the data were aggregated at a county level. Attribute values used at these sites were ice-related crash rates, which has already been defined.

In any type of clustering analysis, one of the most important questions is that of the conceptualization of spatial association among the features, or the construction of the spatial weight matrix *W*. It was decided that the proposed choice of weight matrix in this research 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 each other are more similar than features further away, although further research is 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 inter-relationship. The idea was to analyze the spatial patterns of ice-related crashes for counties that display similar safety trends. The patterns were analyzed for each county identified as belonging to a statistically significant cluster of counties that display 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 developed for use in the field of epidemiology and social sciences. The K-function method is one such procedure which has been most widely used (*12*). However, these methods are based on the assumption of data being 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 $A = \{a_1, a_2, ..., a_{n_a}\}$ and $B = \{b_1, b_2, ..., b_{n_b}\}$, placed on a finite planar network (L_T), and shows whether set of points *B* affects the location of 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 set of points *B* affects the distribution of the set of points *A*. Note that no assumption is made with respect to the distribution of points *B* (*6*). The cross K-function can now be defined as follows:

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

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 $b_1,...,b_{n_b}\left(b_i \in B\right)$,
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 which 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 five percent bounds and show the statistical significant of the observed network cross K-function at the 95 percent confidence level. If the line of observed values lies above the upper five percent line, the pattern is said to be statistically significant clustered. If the observed line lies below the lower five percent 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 of the analysis was to analyze 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 with each other. Ice-related crash rates as defined in previous sections were calculated for each county in Wisconsin based on crash and winter VMT data for three winter seasons between 2003 and 2006. Table 1 presents the number of ice-related crashes, VMT, crash rates, percentage of ice-related crashes, and the results of Local Moran's I analysis for the selected counties. The crash rates (per 100 million vehicle miles traveled) were plotted on a map for visual interpretation as shown in Figure 2 (a). Although the figure presents a fair picture of how the crash rates are distributed amongst 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 95 percent confidence level, and vice versa. Figure 2(b) is also overlaid with the actual value of ice-related crash rates for each county represented by dots the size of which is proportional to the value of ice-related crash rates, with larger dots representing higher rates. Results identified four clusters in different regions of the state that display statistically significant similar ice-related crash rate values. These regions are located roughly in the Northwest, Southeast, North central, and far West regions of the state. 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 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 were identified, consisting of between two and eight counties each, two regions with four counties each were selected for further analysis. These two regions were selected because they were part of the biggest clusters yielding greater number of counties for analysis. Moreover, the geographic proximity would well represent the varying nature of winter weather between the North and South of the state. The two regions were located in the Northwest and Southeast areas of the state and the eight selected counties are shown in Figure 2(c) and Figure 2(d). It is also observed 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 comparisons between the counties in terms of their winter maintenance strategies to counter ice-related crashes, especially in the form of pro-active anti-icing winter maintenance activities.

Table 1Results of Local Moran's I Analysis for Statistically Significant Selected 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

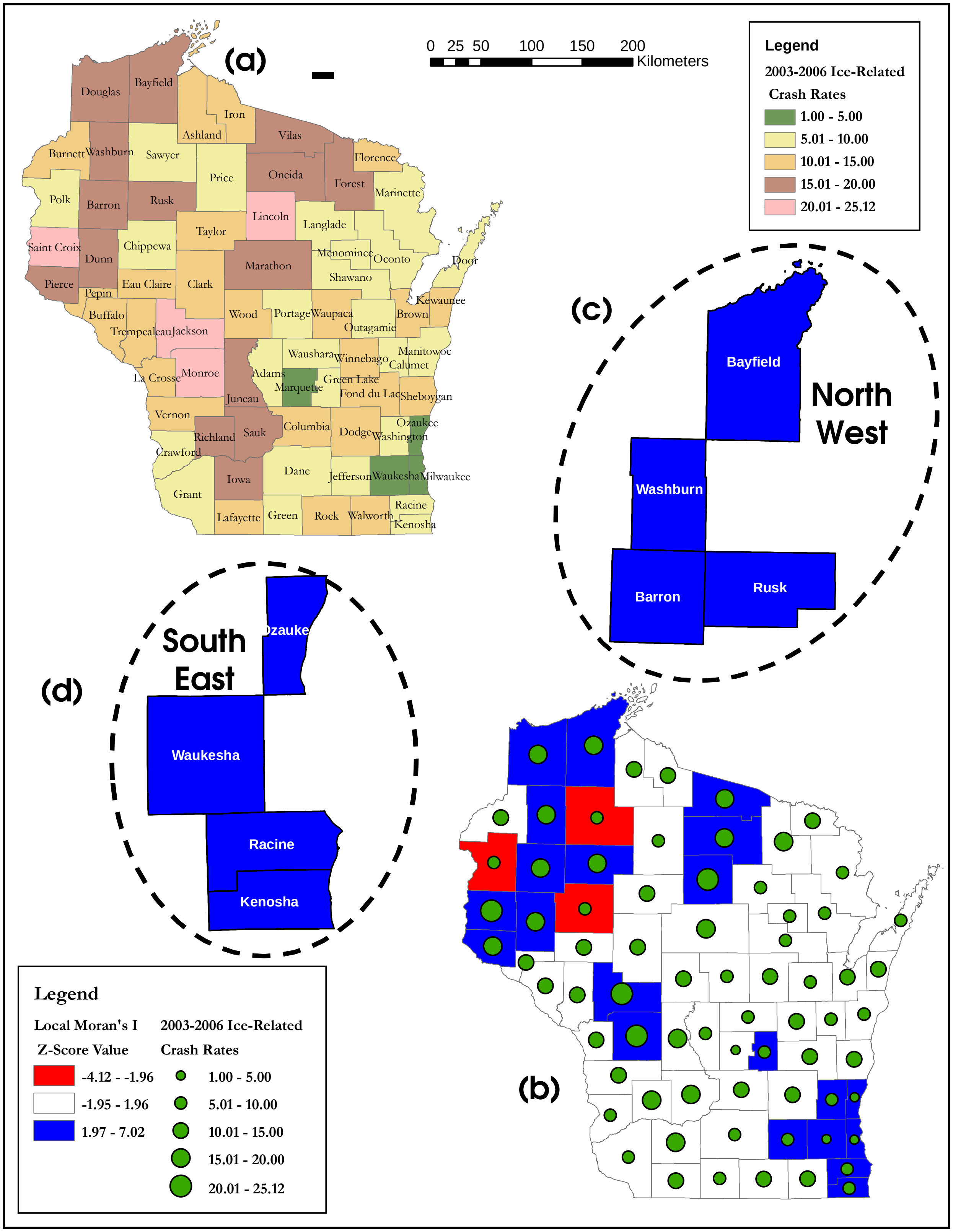

Figure 2 (a) Ice-related crash rates for Wisconsin counties 2003-2006 winter crash data. (b) Local Moran's I analysis Z score values for Wisconsin counties. (c) Selected counties from cluster of statistically significant counties of similar ice-related crash rates in Northwest region. (d) Selected counties from cluster of statistically significant counties of similar ice-related crash rates in Southeast region

Network Cross K-function Analysis Results

The network cross K-function analysis was conducted using the SANET tool developed by Okabe, Okunuki, and Shiode (*14*). The aim was to identify clusters of ice-related crashes as well as study the relationship between ice-related crashes and the location of bridges in each county. Figures 3 and 4 show the results of network cross K-function analysis for counties selected from the Northwest and Southeast regions, respectively. 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 either side of a bridge. The x-axis shows the increasing distance away from a bridge location and the y-axis shows the number of crashes observed within each distance increment for all bridges in a county. The graphs indicate the relationship between bridge locations in individual counties and whether ice-related crashes cluster significantly within 1 km of either side of the bridges.

Northwest Counties Results

Figure 3 shows the results of cross K-function analysis conducted for each county selected from the Northwest region of the state as shown in Figure 2(c). Figure 3(a,b) and Figure 3(c,d) display statistically significant clustering of ice-related crashes around bridge locations in Barron and Washburn counties, especially the county of Barron which shows very high clustering depicted by the large spike of the observed cross K-function line within the first 100 meters (Figure 3(a)). Conversely, Rusk and Bayfield counties show no significant clustering of ice-related crashes around bridge locations as shown in Figure 3(e,f) and Figure 3(g,h), respectively. Although the observed K-function line is above the mean line at certain distance increments, suggesting a clustering tendency, the results are inconclusive at 95 percent confidence level.

The results of the K-function analysis for counties in the Northwest regions suggests that bridge locations in Washburn and Barron counties are more prone to the occurrence of ice-related crashes than in the counties of Rusk and Bayfield. Given the similar safety performance of these counties for ice-related crashes, the differences in the occurrence of ice-related crashes at bridge locations are clear. The results provide conclusive evidence that the counties of Washburn and Barron should focus additional maintenance attention on bridge locations. Moreover, the results also provide an opportunity for the counties to compare and contrast their winter maintenance activities in relation to bridge locations to improve each county's results. Several reasons could cause differences in patterns, including difference in winter maintenance techniques and priorities, specifically in terms of anti-icing versus de-icing strategies.

Southeast Counties Results

Figure 4 shows the results of K-function analysis conducted for each county selected from the Southeast region of the state as shown in Figure 2(d). Figure 4(a-h) shows statistically significant clustering of ice-related crashes around bridge locations in all the selected counties in this region. Clustering is also very close to the location of the bridges, almost within the first 50 meters on either side of the bridge. Moreover, the clustering tends to become insignificant very quickly as the distance from the bridges increases. These results are consistent throughout the four counties similar to the ice-related crash rates.

The results of the K-function analysis for counties in the Southeast regions suggest there is a significant relationship between the occurrence of ice-related crashes and bridge locations. The results provide conclusive evidence that the counties of Ozaukee, Waukesha, Racine, and Kenosha should focus additional maintenance efforts on bridge locations to reduce the occurrence of ice-related crashes.

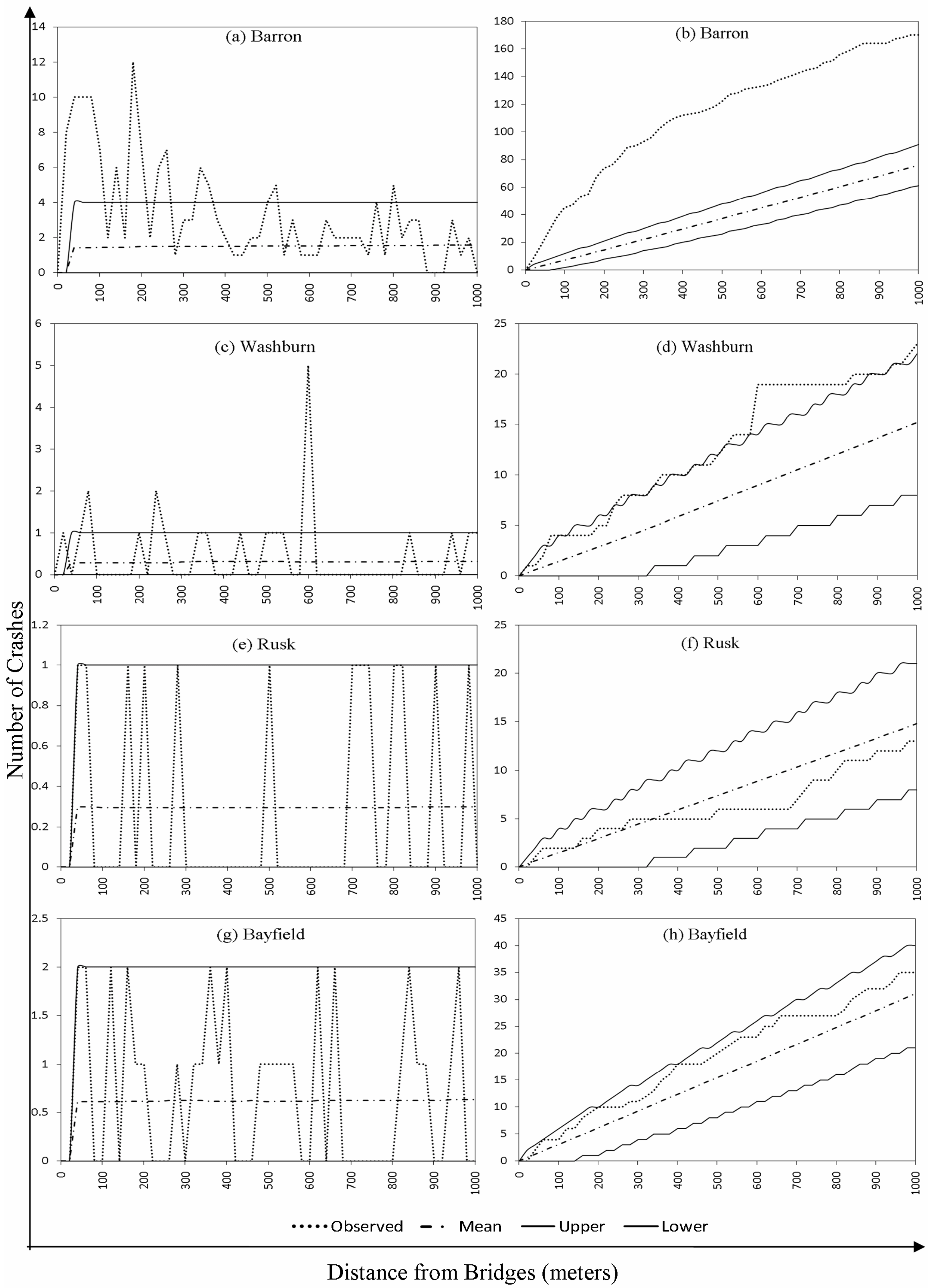

Figure 3 Network K-function results of statistically significant counties selected from the Northwest region of Wisconsin

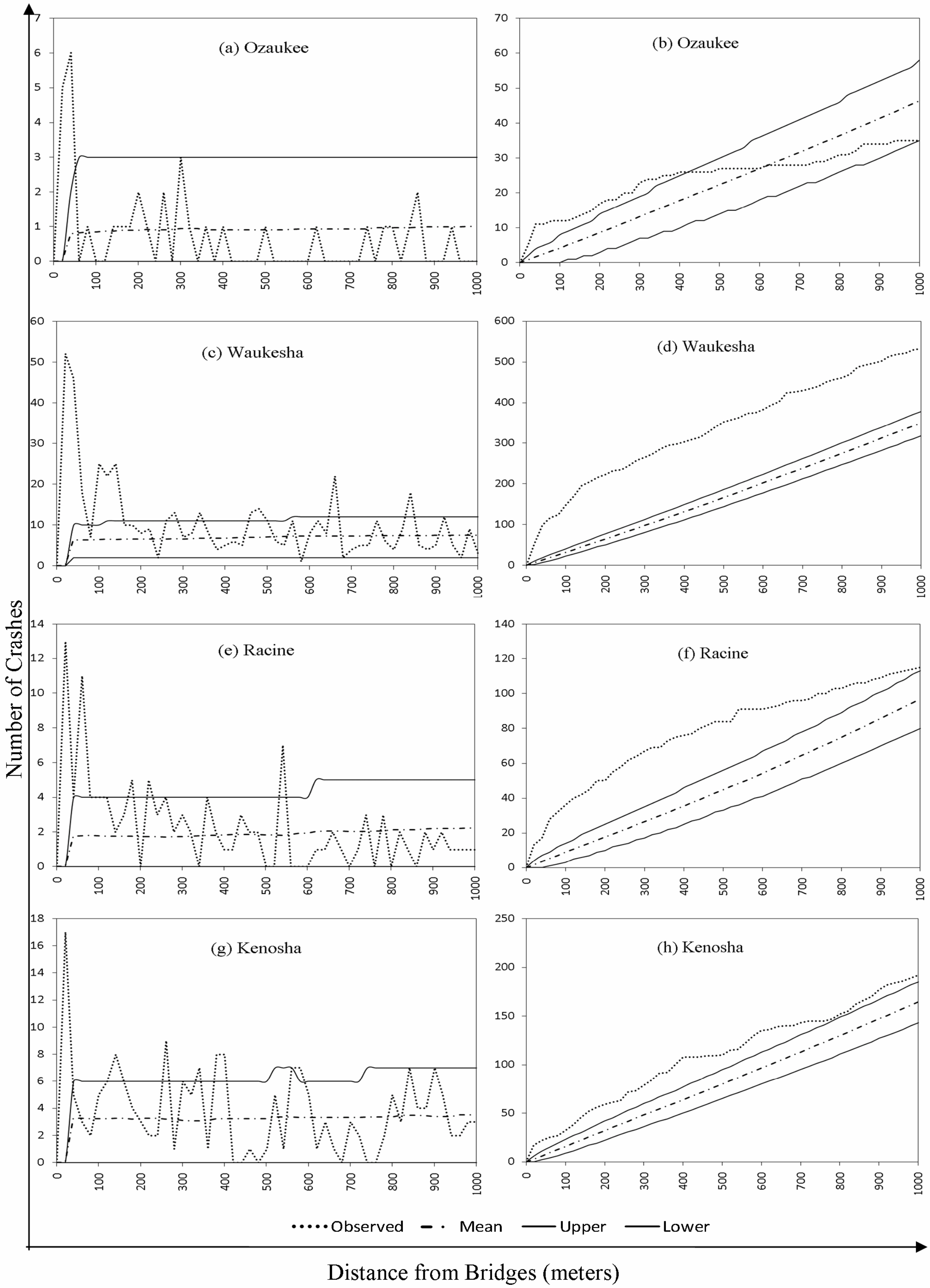

Figure 4 Network K-function results of statistically significant counties selected from the Southeast region of Wisconsin

CONCLUSIONS

This research demonstrated the innovative application and 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 were developed for datasets distributed in planar space. The assumption of planar space is violated for crash data; hence, the use of network based methods takes added significance, especially when analyzing the microscopic components of a large scale system such as the case of looking at individual features over several counties.

The aforementioned procedures bring an added accuracy dimension 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 for winter maintenance evaluations provides a different perspective to winter maintenance decision making which has not been previously explored. The results of this research were not only based on visual maps and interpretations but also incorporate the use of advanced statistical methodologies. Analysis at the levels shown in this research are rare in the literature. In fact, most studies available focus their analysis at smaller scales such as streets or county levels due to data or computational limitations.

Research findings showed patterns of ice-related crashes in relation to bridges which are considered to be ice-prone locations, and are the focus of winter maintenance activities conducted by counties such as anti-icing and de-icing. This research adds to this knowledge by providing statistical measurements that suggest that ice-related crashes cluster around bridges at several locations. Thus, for those counties where ice-related crashes cluster around bridge locations, winter maintenance activities should be focused at bridges to improve the safety. The results indicate bridges where ice-related crashes cluster as hotspots enable stakeholders to focus their maintenance and safety improvements at locations containing features that are causing the clustering of crashes. County winter maintenance personnel can use the 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 the 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. The methodology can be easily applied to other types of crash data either by individual types or severity against different geometric features such as intersection locations, and segment mid-points. Tools used in this research are readily available online and would only require basic GIS knowledge and the use of ArcGIS software.

ACKNOWLEDGEMENTS

The network cross K-function analysis was conducted using the SANET tool developed by Mr. Atsuyuki Okabe, Kei-ichi Okunuki, and Shino Shiode at the Center for Spatial Information

Science, University of Tokyo. The authors appreciate and thank them for providing the tool and all the help and necessary guidance to implement its use.

REFERENCES

1. Jones, A. P., Langford, I. H. and Bentham, G. The Application of K-Function Analysis to the Geographical Distribution of Road Traffic Accident Outcomes in Norfolk, England. *Social Science & Medicine,* Vol. 42, Issue 6, 1996, 879-885.

2. Kim, Karl E., and Yamashita, Eric Y. Using K-Means Clustering Algorithm to Examine Patterns of Pedestrian Involved Crashes in Honolulu, Hawaii. CD-ROM *Transportation Research Board*, National Research Council, Washington, D.C., 2004.

3. Levine, N., Karl E. Kim and Lawrence H. Nitz, Spatial Analysis of Honolulu Motor Vehicle Crashes: I. Spatial Patterns. *Accident Analysis & Prevention,* Vol. 27, Issue 5, October 1995, 663-674.

4. Thomas, I. Spatial Data Aggregation: Exploratory Analysis of Road Accidents. *Accident Analysis & Prevention,* Vol. 28, Issue 2, 1996, 251-264.

5. Abdel-Aty, M. and Wang, X. Crash Estimation at Signalized Intersections Along Corridors. In *Transportation Research Record; Journal of the Transportation Research Board, No. 1953. Transportation Research Board*, National Research Council, Washington, D.C. 2006, pp 98-111.

6. Okabe, A and Yamada, I. 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 Thill, J. 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 Thill, J. 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., Parker, S., T., Qin, X., Qiu, Z. and Noyce, D., A. System for Digitizing Information on Wisconsin’s Crash Locations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019*. Transportation Research Board, National Research Council, Washington, D.C. 2007, pp. 256-264.

11. Anselin, L. Local Indicators of Spatial Association - LISA. *Geographical Analysis,* Vol. 27, 1995, 93-115.

12. Ripley, B., D. Modelling Spatial Patterns. *Journal of Royal Statistics*. Vol. 39, 1977, pp. 172-212.

13. Okabe, A. and Yamada, I. The K-function Method on a Network and its Computational Implementation: CSIS Discussion Paper # 26. Center for Spatial Information Science, University of Tokyo, Japan. April 20, 2000.

14. Okabe, A., Okunuki, K. and Shiode, S. SANET: A Toolbox for Spatial Analysis on a Network. *Geographical Analysis*, Vol. 38(1), 2006, 57-66.