Spatial Statistics Analysis to Identify Hot Spots Using Accidental Event Calls Services

Samira Ghashghaie and Saeed Behzadi *

Shahid Rajaee Teacher Training University

Received: 3/27/2020   Approved: 1/24/2021

Abstract. Today, data is produced in large volumes, and from multiple sources, so this has caused problems in service. These problems can also affect the speed and accuracy of emergency services. Therefore, access to various resources and databases, information extraction to evaluate and analyze data and provide appropriate solutions for citizens is inevitable. In this paper evaluation of clusters is used for Getis-ord Gi* statistics and Anselin Local Moran's I statistics to identify the behavioral pattern of data. The data used in this article is a large free dataset of Spatio-temporal emergency call events from the United States. Accidental call events in five years are evaluated from this dataset. Moran statistics are used to identify and detect the events which have the pattern of spatial distribution. A high/low distribution pattern of accidental events was obtained through Hotspot maps, an annual comparison and evaluation are made by survey the distribution map of events. Clustering Hotspots Map with Gi* statistics represents the spatial correlation between positive and negative events. In these distribution patterns, clusters with high value are called Hot-spots, and low-value clusters are called Cold-spots. Similarly, clustering maps of accidental events get evaluated every five years; then the Gi indicator evaluates each cluster for every two years. A positive z-score and G-index indicate that the data have a positive spatial correlation; its results show that the distribution pattern is similar in each year with an average of 93 percent. Then, hot/cold spot

---

* Corresponding author

clustering maps of 5-year accidental events are also created with the general Moran indicator. Moreover, a confidence level is created after calculating the p-value and z-score. In all 5-year data calculations, the Moran coefficient of accidental events is greater than the expected coefficient. Evaluation of biennial clustering maps with Moran index showed that there is more than 96percent behavioral similarity of dispersion pattern in both years. Raster clustering maps are also created to evaluate the clustering of Moran and Gi indicates. The similarity of raster clusters is more than 95percent per year. The results show that the pattern of accident distribution is the same in 5 years. Although the number of accidents has decreased during this period, the hotspots of accidents have not changed significantly in the city. Furthermore, hotspots indicate a high density of accidental events with 95percent confidence in the study area, and most accidents occur on the South Claiborne and New Orleans highways and at intersections with major streets.

1 Introduction

Today, location-based call service, such as the police department is one of the many data sources in urban services. Since urban events are interrelated, creating the right platform to provide services to citizens is inevitable. Identifying important factors affecting urban services is one of the challenges in this field. Therefore, the needs of today's urban services are the evaluation and assessment of urban events, the identification of patterns, sequencing events, their classification, and so on for improvement of the quality of services, events prediction, and incidents elimination (Chohlas-Wood et al., 2015; Hamoudzadeh & Behzadi, 2020). Today, access to various databases for location-based information has been expanding. One of the most common methods for identifying and extracting patterns in big data is the use of clustering methods. Clustering algorithms are classified as the cluster model. One of the model-based clustering methods is the degree of clustering used in the Getis-ord-Gi* Statistics for High/Low Values. This method is inferential statistics, so the results of the analysis are based on the null hypothesis and density of the values (Sheikholeslami et al., 1998). One of the new studies

in the field of urban events is the identification of the events' behavioral patterns based on time (Chohlas-Wood et al., 2015). Knox Test is one of the methods for the evaluation of spatiotemporal analysis in criminal events. In this method, the classification of the normal breaks, the closest neighborhoods, and Ripley's k function is used to determine the critical distance from the mean distance. This test is a complementary method for hotspot cluster analysis, which is suitable for determining the quantity of both spatial and temporal data interaction of crime (Kalantari et al., 2016). One of the main issues in recent urban development studies is the occurrence of traffic accidents, which increases the rate of accidents and costs (Behzadi, 2020). It also reduces the provision of appropriate insurance, emergency, and police services (Ryder et al., 2018). One of the methods for analyzing accidental events is through the use of the Geographic Information System (GIS) and spatiotemporal patterns in accident-prone locations (Aghajani et al., 2017). The purpose of this paper is to identify the spatial pattern of accidental events using two algorithms in GIS: 1) identify the accidental points in the city, and 2) evaluate the distribution patterns of accidental events in a 5 year.

2 Definition

2.1 Clustering

Clustering is a process done by a set of objects which can be grouped into distinct groups. Each cluster segment has very similar properties, and the level of similarity among them is minimum (Jafarian & Behzadi, 2020). The clusters include groups such as the small distance between cluster members, a dense data zone, distances, and specific statistical distributions. Data clustering is used to identify or predict phenomena, which is an unsupervised method. The purpose of clustering is to group data with the maximum internal similarities, and minimum similarity among classes (Norouzi & Behzadi, 2019). There are more than 100 types of published algorithms for clustering, so it is not easy to categorize these algorithms. The best clustering algorithm is selected experimentally to solve a specific problem. Clustering algorithms can be categorized into 4 main groups: Partitioning Clustering algorithm, Hierarchical Clustering algorithm, Density-Bases Clustering algorithm, and Model-Based Clustering algorithm (Sheikholeslami et al., 1998).
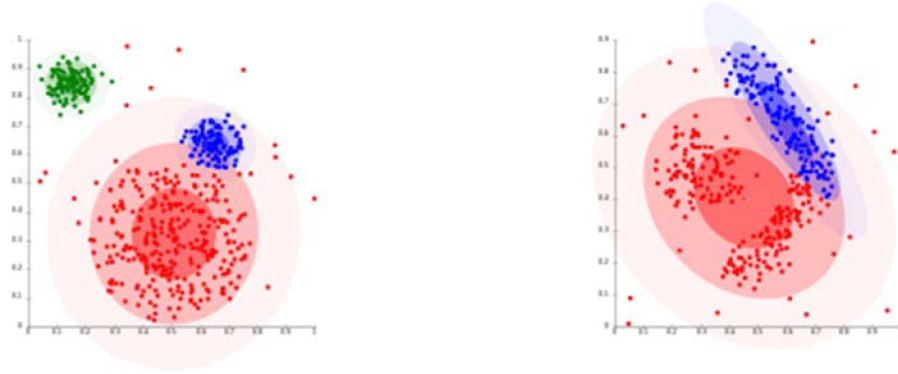
Figure 1. Density-based clustering does not suit Gaussian distribution well while EM has worked well for Gaussian data

2.2   Model-Based Clustering Algorithm

In this method, the data is supposed to follow a statistical distribution, and a hidden variable is introduced as a cluster label in the model. A likelihood function is written and checked to find the cluster to facilitate the probability of dependence of the i-th point to the j-th cluster. One of the known methods is the Gaussian Mixture Model. This method uses the Expectation-Maximization algorithm (Forero et al., 2012). The dataset is randomly selected with the number of fixed Gaussian distributions. The parameters which are repeatedly optimized converge to local optimization, so they can generate different results over several runs. Distributed clustering creates complex models for clusters, which can show the correlation and dependency of a feature (Bradley et al., 1998). Figure 1 shows the density-based clustering using the Gaussian distribution;

2.3   Spatial Autocorrelation

Evaluating clustering models is done in different ways. One of these methods is to investigate the directional of the clusters. In this case, the data of a cluster can be non-random if the data is random; as a result, the cluster is not valid. Hopkins statistics are used to evaluate the non-randomness of the cluster (Pfitzner et al., 2009). This statistic has two assumptions null hypothesis, and one hypothesis. In the null hypothesis, the data is distributed monotonic, and there are no significant clusters in them, while in

one hypothesis, the data is not distributed monotonic and there are significant clusters in them. If the test of the statistic is greater than 0.5, we can reject hypothesis one and assume that the data is monotonically distributed. Another method for assessing clusters is evaluating through inner evaluation methods (Lee & Wong, 2001). If data clustering uses algorithms for extraction of the elements of potential statistical clusters, the existence of autocorrelation is one of the issues that may lead to bias. This point requires an earlier scheme of analytical models to determine if all of the spatial and temporal data are dependent or not (Melecky, 2015). The spatial correlation mechanism indicates whether related data exhibit similar behavior or not. In spatial autocorrelation, apart from examining the relationship between the two elements, the neighborhood of the data is also checked.

General G-Statistic

General G statistic (Getis and Ord, 1992) is a local index spatial autocorrelation. The G statistics are calculated for each zonal region, and it implies that how the unit value of the region studied is related to the neighborhood value, which is defined by the distance (d) of the threshold (Peeters et al., 2015). This single-variable is obtained from equation (1):

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j}{\sum_{j=1}^n x_j} \tag{1}$$

Where Gi* describes the spatial dependency of the incident i over all n events, xj is the magnitude of variable X at incident location j over all n (j may equal i), wij is the weight value between event i and j, which represents their spatial interrelationship.

Wij is usually calculated based on the conceptualized spatial relationship about d (Songchitruksa & Zeng, 2010). Equation (2) is for expected coefficient (E) and required variance (V) statistics. The expected coefficients are obtained as:

$$Z_G = \frac{G - E[G]}{\sqrt{V[G]}} \tag{2}$$

$$E[G] = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij}}{n(n-1)}.ji \tag{3}$$

$$V[G] = E[G^2] - E[G]^2 \tag{4}$$

A z-score is obtained when spatial clustering is formed by high similar coefficients. If the spatial classification is low, z-score is negative. A null z score implies that there is no spatial correlation pattern statistically (Ding et al., 2015). If the z-score is positive, it indicates that the clustering is more intense in higher values. If the z-score is negative, it indicates that the intensity of the lower is higher in the dataset.

2.4   Anselin Moran I Statistics

Spatial Autocorrelation Local Moran's Statistics is used to detect spatial patterns of local differences. This indicator measures the space deference of the values between each cell and contiguous cells; it also evaluates its significance (Zhang & Lin, 2016). Local Moran statistics is decomposed of General Moran statistics. Local Moran's I statistics calculates z-score, pseudo-p-value, and it displays a variety of clusters with a code that has its statistical attributes. Results of the local Moran statistics are tested with z-score; as a result, a confidence level is determined. If cell i has a positive sign, the value of the cell i is similar to its neighbor cells value. If the value of i is a large positive number, it indicates a strong clustering range. If the value of i is negative and significant, the amount of surface cell property i has a high difference with its neighbor cells, which indicates a negative spatial correlation (Chen, 2013). The Local Moran's I statistics associations are obtained from equations (5) and (6):

$$I_i = \frac{X_i - \overline{X}}{S_i^2} \sum_{j=1?ji}^{n} w_{ij}(X_i - \overline{X}) \tag{5}$$

$$S_i^2 = \frac{\sum_{j=1?ji}^{n} (X_i - \overline{X})^2}{n-1} - \overline{X}^2 \tag{6}$$

The z-score is calculated as:

$$Z_{I_i} = \frac{I_i - E[I_i]}{\sqrt{V[I_i]}} \tag{7}$$

$$E[I_i] = -\frac{\sum_{j=1?ji}^{n} w_{ij}}{n-1}? \tag{8}$$

Table 1. Number of Accidental call events Datasets for each year

Year	call event	Calls Related with Accident	Outliers	Final calls Accident
2012	493375	19017	49	18968
2013	463624	19899	23	19876
2014	435125	20239	2	20237
2015	419738	21436	34	21402
2016	389942	21634	58	21576

$$V[I_i] = E[I_i^2] - E[I_i]^2 \qquad (9)$$

3 Materials and Methods

3.1 Case Study

In this article, the database is a criminal call event for the various services of the New Orleans Urban Services in Louisiana, USA. The data includes events that are reported to the New Orleans Police Department over a year through various call services. The collection is being compiled by the Orleans Parish Communications District (OPCD) for New Orleans and the 911 Office New Orleans Police Department (NOPD). This dataset has been collected from various communication call services such as 911, 21, 18, 24, 103, and so on. This database contains the geographic location, the type of call, the block address, and the type of service of the contact in each event. These events are based on unconfirmed reports to the police, and the classification of the events may change after further investigation. Therefore, there is no responsibility for the error or omission of information.

3.2 Data

There are about 187 call events in the dataset and the calls associated with accidental events, such as auto accident, auto accident with injury for each year. Table (1) represents the dataset accidental call events in the years 2012-2016.

Urban maps such as streets, urban blocks, and land use are also used in this research, which is shown in Figure (2).
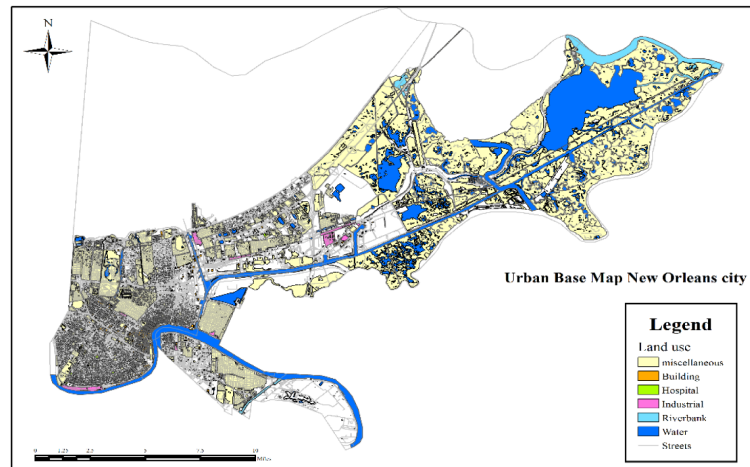
Figure 2. Urban Base Map from New Orleans City

3.3   Methodology

At first, the call event dataset is converted to a point layer. This point layer is investigated in GIS and all random events such as auto accident, traffic accident, an accident with injury, and so on are separated from the dataset. Next, low/high clustering is performed using G general statistics. P-value, z-score values, the G index, the Hotspot Cluster Map are created to detect the distribution pattern of accidental events in each year. Then, the clustering maps are converted to raster data using Gi values as a pixel value. Raster maps are then compared for every two years. The similarity or difference between the behavior patterns of accidental events is then used to General Moran statistics by calculating spatial autocorrelation and analyzing the z-score and Moran index. Next, a Hotspot cluster map is created using Moran statistics to detect the pattern of event distribution and identify random clusters for each year, the maps of each two years are compared and evaluated based on the Moran index. Finally, Hotspot clustering maps are compared, evaluated, and analyzed for each year with Moran and G indexes.

4 Results

4.1 Clustering Map of Accidental Events and Analysis Based on General G Index

G statistics clustering calculates the p-value, z-score, and G index for all data in 5 years. According to G-statistic calculations, the p-value obtained for each year indicates the probability of random data being null. A p-value of zero indicates that the data is not random, and there should be a statistically significant relationship between the data. Positive values of z-score and G index also indicates that the data has a positive spatial autocorrelation. The p-value obtained for each year is grate than 0 and the value of z is also positive (for example, z-score is 15.416 in 2014, and z-score is 15.029 in 2016). The results show that the data is not normal because data volumes are large and single-variable data are also numeric and have no scale. According to previous studies, the distance threshold was considered the same as 300 feet. The expected coefficient General G and the observed coefficient General G are then obtained. The result shows that the observed coefficient General G is greater than the expected coefficient G. For example in 2012, the expected coefficient G and the observed coefficient G are respectively equal to 0.000004 and 0.000013. Calculations of G statistics are shown in figures (3), (4) and (5) for each year.

4.2 Evaluation, and Analysis of Raster's Cluster Map of Events Based on the Gi Index

For evaluation, G statistics is used to reveal the type of semantic relationship between accidental events and the distribution pattern of events. At first, the p-value and z-score are calculated, then Gi's index calculates the data clustering map based on the G-statistics. In the cluster maps of each five years, the Gi index is grown from 0 to +3. Gi index with high values represents a statistically significant positive event that has reached the value of +3 which is the hot spot of the region. A value of 0 Gi index indicates that there is no statistically significant relationship between events. Moreover, the negative value of the Gi index is not observed in the maps, so there are no cold spots in clusters in these areas. A 5-year accidental events cluster map is presented in figure (6) using the Gi index.

These maps are then converted to raster to make a more accurate comparison among accidental events over five years. The comparison of raster
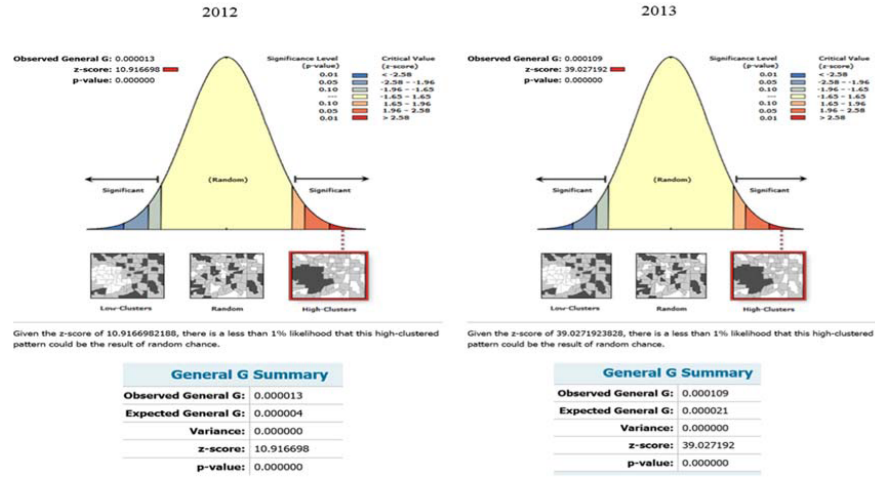
Figure 3. High / Low Clustering Chart with General G statistics (2012 and 2013)
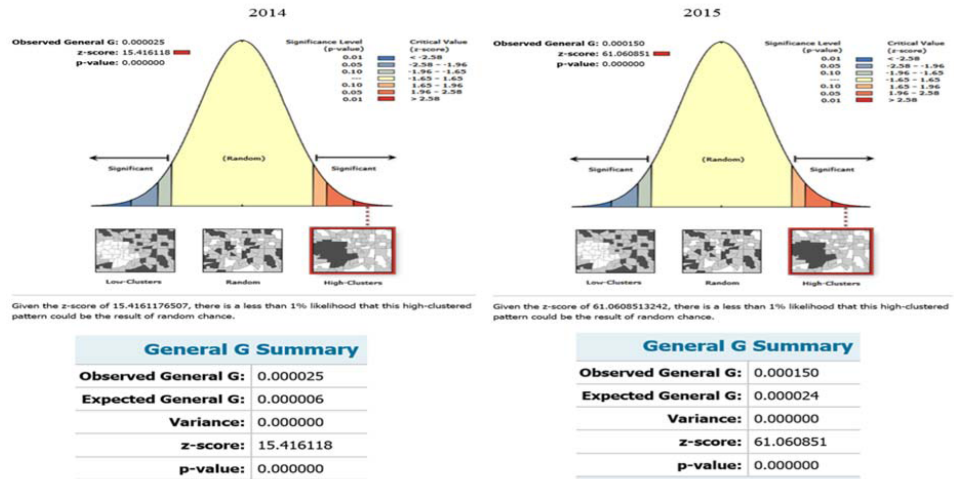

Figure 4. High / Low Clustering Chart with General G statistics (2014 and 2015)
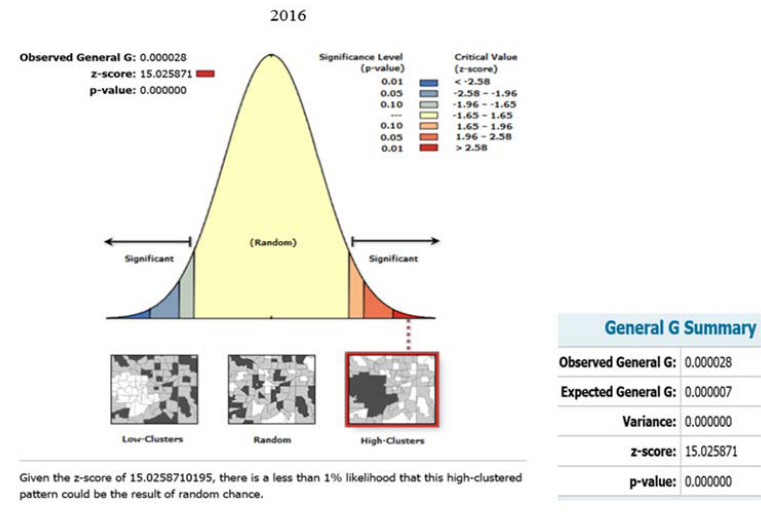
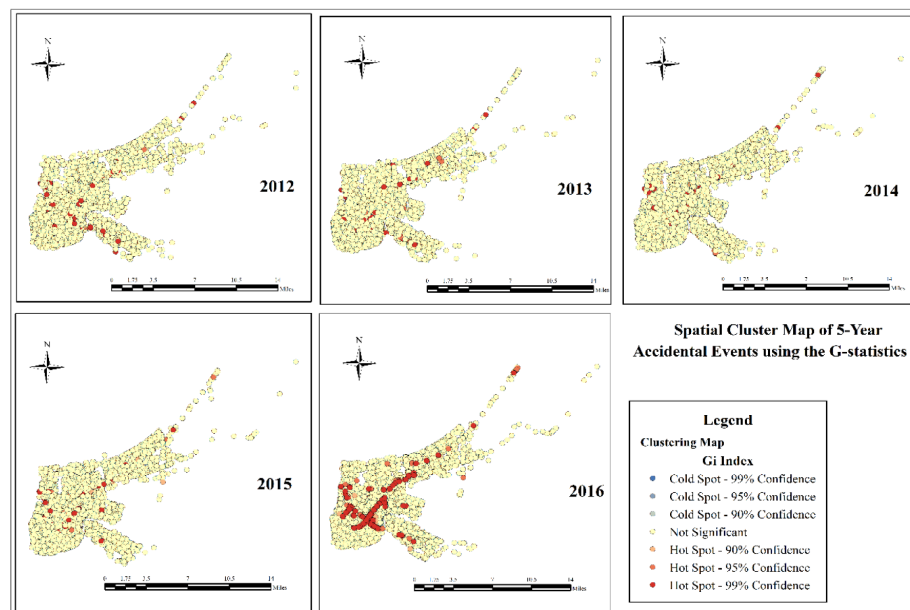**Figure 5.** High / Low Clustering Chart with General G statistics(2016)

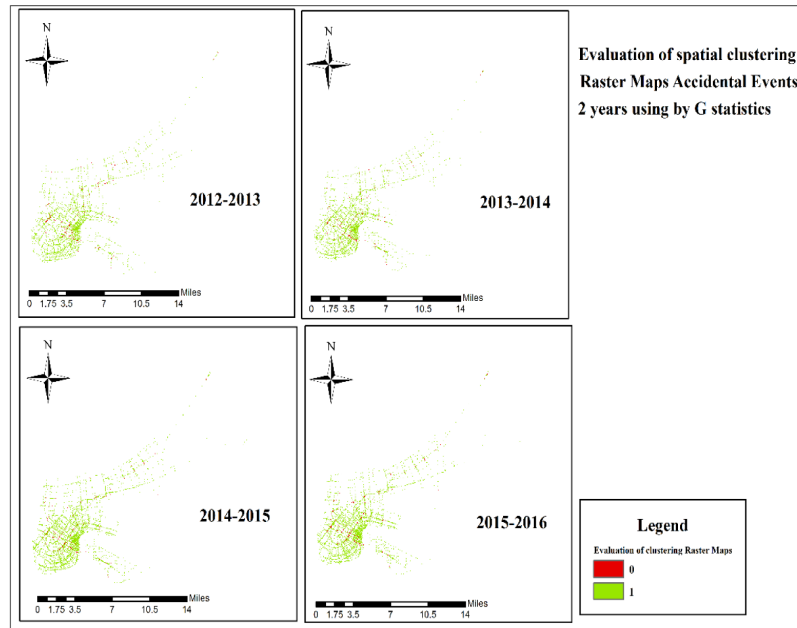**Figure 6.** Spatial Cluster Map of 5 Years Accidental Events using the G-statistics

Figure 7. Raster maps comparing 2-year Accidental events

maps is based on their defined properties and distance tolerance. The index Gi is given to each pixel as a numeric value. The G index is the main criterion for comparing the raster. The number of pixels with identical and non-identical values is then obtained based on the pixel value for every two years of the Gi index. Finally, the output map comes with values 0 and 1. Pixels with a value of 1 represent the same Gi index. Pixels with a value of 0 represent different Gi index. Raster maps comparing 2-year accidental events are presented in figure (7).

Table (2) represents the percentage of similarity in the accidental events cluster map for every two years.

4.3   Clustering Map of Accidental Events and Analysis Based on Moran Index

General Moran statistics are used to identify the geographic distribution pattern. The behavior of the cluster is based on the predetermined property and it does not indicate the type of spatial behavior. So local Moran statistics are to detect the behavior of the phenomenon in terms of the spatial distri-

Table 2. Percentage Similarity of Cluster Map of Accidental Events Based on Gi Index

No.	Compare two annual raster maps	Similarity percentage
1	2012-2013	94.43
2	2013-2014	95.29
3	2014-2015	95.88
4	2015-2016	94.80

bution pattern on the map (Wang & Fang, 2016). In this statistic, p-value and z-score are calculated first. Then a confidence level is defined for the z-score. Each cell of the map has a significant number. If the value of i is positive, it indicates a strong clustering range. If the value of i is negative, it has many differences with its neighborhood cells, which indicates a negative spatial correlation (Zhang & Lin, 2016). The p-value of zero represents the probability of non-random data, and the z-score indicates a statistically significant positive correlation between the data. The local spatial correlation coefficient must be greater than the expected coefficient so that the pattern of spatial distribution becomes acceptable (Nadian et al., 2018). Figures (8), (9) and (10) show the spatial autocorrelation Moran index. According to these Figures, the I index is larger than the expected coefficient. For example, the expected coefficient m Moran and the Moran index are equal to -0.000152 and 0.4186 respectively in 2014. The z-score is also positive for example the z-score is equal to 8.118 in 2012.

4.4 Evaluation and Analysis of Raster's Cluster Map of Events Based on the Moran Index

After the implementation of the General Moran Index, a significant map of local Moran is created. Beside the map of significant local Moran, the Moran scattering chart helps spatial autocorrelation type. This chart divides the data into 4 groups: High-value points surrounded by the other high-value points, which are known as High-High (HH) clusters. Low value points surrounded by high-value points, which are known as Low-High (LH) clusters. Low-value points surrounded by the other low-values points, are known as Low-Low (LL) clusters. High-value points surrounded by low-value points, are known as High-Low clusters (HL). Using these concepts, cluster maps of accidental events are created for each year based on the Moran index, which is shown in figure (11).
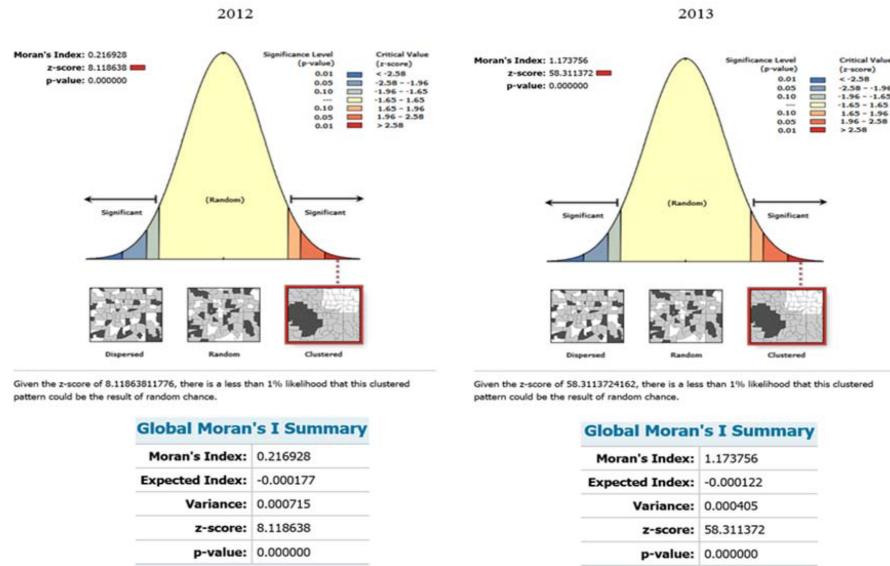
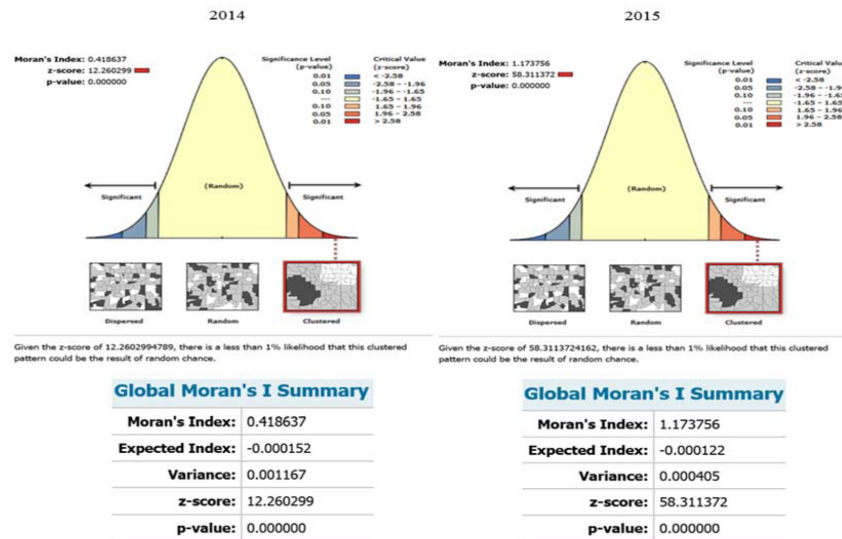**Figure 8**. Chart spatial Autocorrelation accidental events is based General Moran I Statistics (2012 and 2013)

Figure 9. Chart spatial Autocorrelation accidental events is based General Moran I Statistics (2014 and 2018)
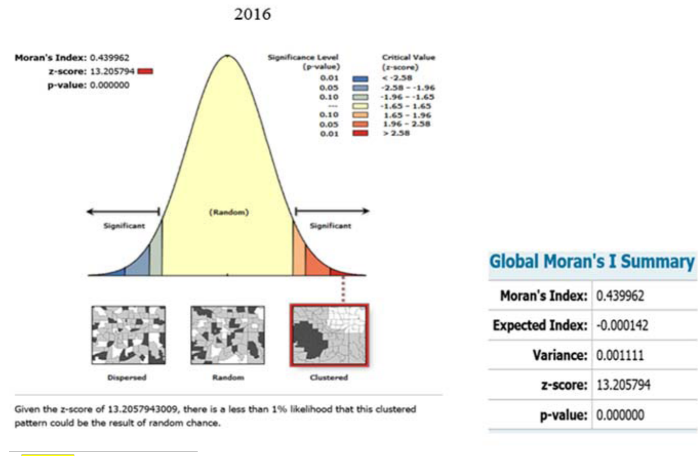
Figure 10. Chart spatial Autocorrelation accidental events is based General Moran I Statistics (2016)
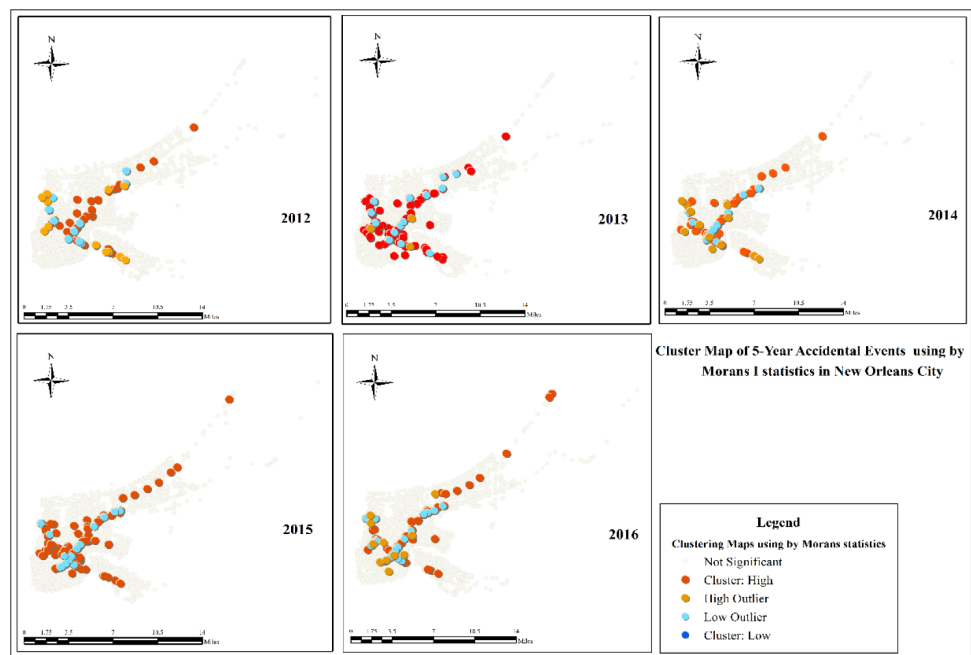

Figure 11. Spatial Clustering Map of 5 Years Accidental Events Using Moran Index
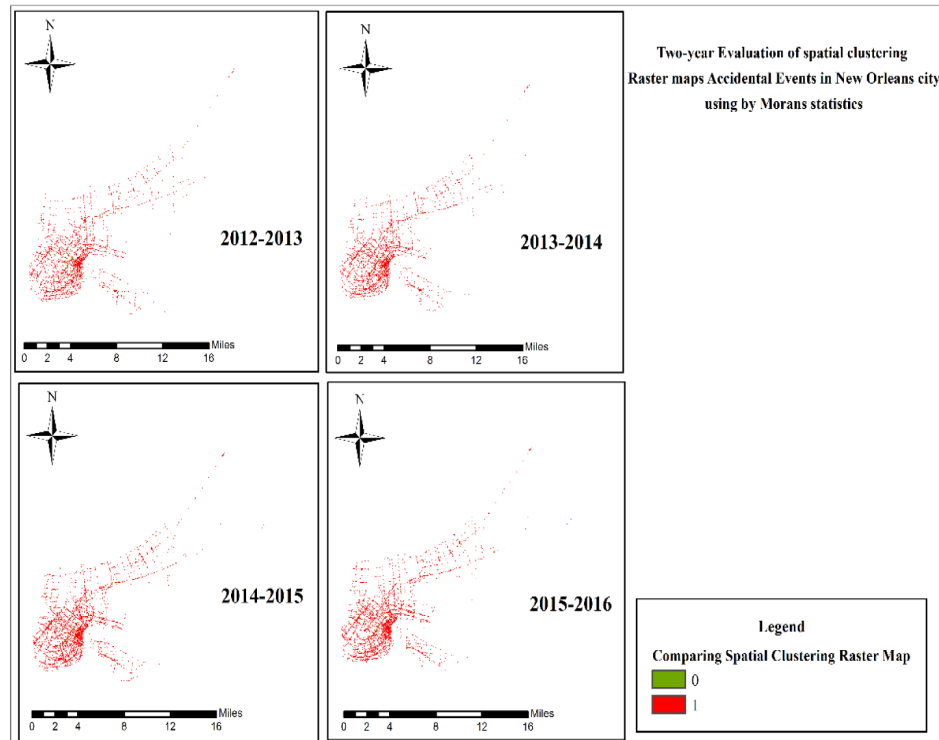
Figure 12. Raster maps comparing 2-year Accidental events using by Moran Index

In this statistic, except clusters with high values such as hot spots, and clusters with low values (cold spots), high/low outlier clusters are also detected. These clusters have high/low values; however, they have no high/low cluster in the neighborhood. Given that the result of the hot/cold analysis is descriptive with Moran statistics, the numerical values are attributed as follows: The cluster with no statistical correlation is considered as 0, the cluster with High-High value is considered as 2, the cluster with High values but outlier (H-L) is assigned with number 1, the clusters with Low-Low value is assigned with number -2, and the cluster with Low values but outlier (L-H) is assigned with number -1. Then, the maps of each 2-years are compared and evaluated by converting the clustering maps to raster, given a hypothetical numerical value. Raster maps obtained by comparing 2-year accidental events are presented in figure (12).

The comparison of the raster maps for two years showed that clustering

Table 3. Percentage Similarity of Cluster Map of Accidental Events Based on Moran Index

No.	Compare two annual raster maps	Similarity percentage
1	2012-2013	96.17
2	2013-2014	96.79
3	2014-2015	96.77
4	2015-2016	96.81

of accidental events has more than 96 percent similarity in every two years. The similarity for every two years are given in table (3).

4.5 Evaluation and Analysis of Raster's Cluster Map of Events Based on the Moran Index and G Index

The differences between the Moran and G statistics can be interpreted in p-value and z-value. When p has no statistical significance, there is no difference between the Moron statistics and the G statistics. In both cases, the null hypothesis cannot be rejected, and the statistical distribution indicates that the values set can be the result of a random process. In other words, the spatial pattern obtained from the samples can be one of the numerous possible random versions of the data. If the significant value of p and z are positive the Gi statistic displays the rejection of the null hypothesis, and spatial distribution has a high value in the dataset; as a result, the random spatial trend is expected. The result also indicates that the hypothesis is rejected by null; the spatial distribution has High or Low values in the dataset and the spatial process is completely random. If the p-value is significant and the z-value is negative, the null hypothesis is rejected. In the G statistics, this represents the spatial distribution of low values, which the random spatial trend is expected in this case. But in the Moran statistics, this indicates that the spatial distribution of High or low values in the dataset is scattered and the random spatial trend is expected. The distributed spatial pattern is often a reflection of many competitive processes; for example, high values reject other high values, and similarly, a low value rejects other low-value effects. Cluster accidental event maps of each year are compared based on the Gi index and Moran index. Raster maps are compared, and the numerical values of each pixel are then compared. This similarity is due to the similarity among z-values of the data, and the identifying of spatial correlation. The comparison of raster maps is shown in figure (13).
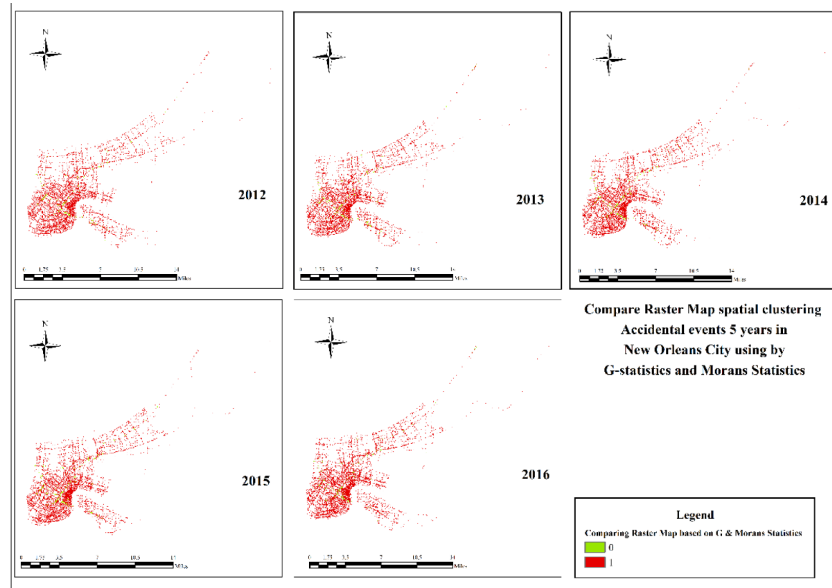
Figure 13. Comparison of cluster Raster maps of accidental events based on Moran statistics and G statistics

By comparing the numerical values of the G and Moran indexes, the similarity of the spatial distribution patterns are obtained for each year (Table (4)).

5   Conclusion

It is found in the evaluations that there are statistically partial differences between the clustering of Moran and G statistics. When the numerical univari-

Table 4. Comparison of clustering based on Moran and Gi index

No.	Compare two annual raster maps	Similarity percentage
1	2012	95.86
2	2013	95.53
3	2014	95.42
4	2015	95.1
5	2016	95.59

ate data are the same, the two statistics provide a similar scattering pattern. The lack of negative spatial correlation in the data is another reason for this similarity. This shows that the similarity of the behavioral pattern reduces the number of accidental events during the five years; however, accidental incident ranges remain constant in the city. This behavior pattern repeated in the last 5 years. It can be said that hot spots represent a high density and cold spots show a low density of accidental events in the spatial range with a confidence of 95 percent. The largest number of accidental events occurred on several highways and main roads and it includes their intersections of the city. The result shows that South Highway Claiborne and the intersection with main streets such as the 1st, 2nd and 30th streets, the main highway of the city of New Orleans with the intersection of the main streets, and so on has high accidental events. Considering that accidental events were investigated and evaluated solely in terms of location and number of calls to urban services, it is necessary that events are analyzed temporally. Also, we can use the other datasets associated with accidental events such as Traffic Apps, traffic control urban cameras, etc. for more accurately identifying the incident hotspots. Furthermore, in future research, the role of various parameters such as environmental factors, technical and geometric factors, etc. can be examined. Another suggestion that can be considered for future researches is the study and evaluation of the urban geometric structures' impact on the occurrence of accidental events.

References

Aghajani, M. A., Dezfoulian, R. S., Arjroody, A. R., and Rezaei, M. (2017). Applying GIS to Identify the Spatial and Temporal Patterns of Road Accidents Using Spatial Statistics (Case Study: Ilam Province, Iran). *Transportation Research Procedia*, **25**, 2126-2138.

Behzadi, S. (2020). An Intelligent Location and State Reorganization of Traffic Signal. *Geodesy and Cartography*, **46**, 145-150.

Bradley, P. S., Fayyad, U., and Reina, C. (1998). *Scaling EM (Expectation-maximization) Clustering to Large Databases*.

Chen, Y. (2013). New approaches for Calculating Moran's Index of Spatial Autocorrelation. *PloS one*, **8**, e68336.

Chohlas-Wood, A., Merali, A., Reed, W., and Damoulas, T. (2015). *Mining 911 Calls in New York City: Temporal Patterns, Detection, and Forecasting*.

Ding, L., Chen, K.-L., Liu, T., Cheng, S.-G., and Wang, X. (2015). Spatial-temporal Hotspot Pattern Analysis of Provincial Environmental Pollution Incidents and Related Regional Sustainable Management in China in the Period 1995–2012. *Sustainability*, **7**, 14385-14407.

Forero, P. A., Kekatos, V., and Giannakis, G. B. (2012). Robust Clustering Using Outlier-sparsity Regularization. *IEEE Transactions on Signal Processing*, **60**, 4163-4177.

Hamoudzadeh, A., and Behzadi, S. (2020). Predicting User's Next Location Using Machine Learning Algorithms. *Spatial Information Research*, 1-9.

Jafarian, H., and Behzadi, S. (2020). Evaluation of PM2. 5 Emissions in Tehran by Means of Remote Sensing and Regression Models. *Pollution*, **6**, 521-529.

Kalantari, M., Yaghmaei, B., and Ghezelbash, S. (2016). Spatio-temporal Analysis of Crime by Developing a Method to Detect Critical Distances for the Knox Test. *International Journal of Geographical Information Science*, **30**, 2302-2320.

Lee, J., and Wong, D. W. (2001). Statistical Analysis with ArcView GIS. John Wiley & Sons.

Melecky, L. (2015). Spatial Autocorrelation Method for Local Analysis of the EU. *Procedia Economics and Finance*, **23**, 1102-1109.

Nadian, M., Mirzaei, R., and SoltaniMohammadi, S. (2018). Application of MoransI Autocorrelation in Spatial-Temporal Analysis of PM2. 5 Pollutant (A Case Study: Tehran City). *Journal of Environmental Health Enginering*, **5**, 197-213.

Norouzi, E., and Behzadi, S. (2019). Evaluating Machine Learning Methods and Satellite Images to Estimate Combined Climatic Indices. *International Journal of Numerical Methods in Civil Engineering*, **4**, 30-38.

Peeters, A., Zude, M., Kathner, J., Unlu, M., Kanber, R., Hetzroni, A., Gebbers, R., and Ben-Gal, A. (2015). Getis–Ord's Hot-and Cold-spot Statistics as a Basis for Multivariate Spatial Clustering of Orchard Tree Data. *Computers and Electronics in Agriculture*, **111**, 140-150.

Pfitzner, D., Leibbrandt, R., and Powers, D. (2009). Characterization and Evaluation of Similarity Measures for Pairs of Clusterings. *Knowledge and Information Systems*, **19**, 361.

Ryder, B., Dahlinger, A., Gahr, B., Zundritsch, P., Wortmann, F., and Fleisch, E. (2018). Spatial Prediction of Traffic Accidents with Critical Driving Events–Insights from a Nationwide Field Study. Transportation Research Part A: Policy and Practice.

Sheikholeslami, G., Chatterjee, S., and Zhang, A. (1998). *Wavecluster: A Multi-resolution Clustering Approach for Very Large Spatial Databases*, VLDB.

Songchitruksa, P., and Zeng, X. (2010). Getis-Ord Spatial Statistics to Identify Hot Spots by Using Incident Management Data. Transportation Research Record: Journal of the Transportation Research Board (2165), 42-51.

Wang, Z.-b., and Fang, C.-l. (2016). Spatial-temporal Characteristics and Determinants of PM2. 5 in the Bohai Rim Urban Agglomeration. *Chemosphere*, **148**, 148-162.

Zhang, T., and Lin, G. (2016). On Moran's I Coefficient under Heterogeneity. *Computational Statistics & Data Analysis*, **95**, 83-94.

Samira Ghashghaie

Shahid Rajaee Teacher Training University,

Lavizan, Tehran, Iran.

email: *Samira.qashqaie.84@gmail.com*

Saeed Behzadi

Shahid Rajaee Teacher Training University,

Lavizan, Tehran, Iran.

email: *behzadi.saeed@gmail.com*