Monitoring Spatial Patterns around Fixed Points: Comparing Two Distance-Based Approaches

Yifei Sun

Department of Geography, California State University Northridge
18111 Nordhoff St. Northridge, CA 91330-8249, USA

Abstract

In this study, the weighted distance method, which combines the cumulative sum method and the inverse distance, is devised for monitoring spatial patterns of point events around fixed points. It is found the weighted-distance approach can better reveal the cluster scale and is less subject to observations far away from the monitor site than the distance-based method developed by Rogerson and Sun (1999). It is also found that combining the weighted-distance approach and the distance-based approach achieves better results for monitoring spatial patterns or changes in spatial patterns.

I. INTRODUCTION

With advances in computing technologies including geographical information systems, and the increased availability of large set of geo-referenced data, interests in exploratory spatial data analysis (ESDA) have been growing. Particularly, many methods have been developed for identifying spatial patterns, or detecting spatial clusters (Openshaw et al., 1986; Openshaw, 1994; Stone, 1988; Turnbull et al., 1990; Baseg and Newell, 1991; Getis and Ord, 1992; Fotheringham and Rogerson, 1994; Anselin, 1995; Kulldorf and Nagarwalla, 1995; Ord and Getis, 1995; Tango, 1995, 2000; Fotheringham and Zhan, 1996; Rogerson, 1999).

Four major dimensions can be observed from the existing literature, *global vs. local*, *absolute clusters vs. relative clusters*, *point data vs. area data, and retrospective vs. prospective.* Global tests are used to answer questions such as "whether or not there exists a cluster of points within the study area." Local statistics can be used for detection around focal points and for detection of clusters (Baseg and Newell, 1991). They are useful in answering the question "where the clusters are."

Methods for detecting *absolute clusters* assume that the study area is a homogeneous plain without any variation from place to place. The spatial patterns are tested against complete spatial random (CSR). Methods for detecting *relative clusters* acknowledge the spatial variations of population at risk. For example, in crime analysis, methods for detecting absolute clusters will note where the crime is clustering without considering the population distribution, while methods for detecting relative clusters will take the population distribution into account. It is naturally expected that crime is less likely to take place at locations where few people live. Methods for detecting absolute clusters include nearest neighbor statistics (Clark and Evans, 1954), K-functions (Ripley, 1981), and quadrat counts analysis. Examples of methods for detecting relative clusters include Tango's C (1995) and Rogerson's spatial Chi-square (1999).

Techniques for testing spatial clustering for both *area* and *point data* are available. Point pattern analysis is associated with area pattern analysis in the sense that point data can always be grouped into some area scheme. For example, crime data may be grouped by census tracts or they may be grouped by imposing a set of grids. Methods for testing patterns of point data have the advantage of using the exact location of the events, while methods for testing area data may provide convenience for further analyses if other attributes are available for the areal scheme.

Raubertas (1982) pointed out another dimension in testing spatial clustering: *retrospective* vs. *prospective*. Retrospective tests constitute a single attempt to decide on the presence or absence of clustering, based upon a set of past observations. In prospective approaches, observations are processed sequentially (Rogerson and Sun, 2001; 1999), and the effort of detecting spatial clustering becomes an on-going continuous process as new cases become available.

However, so far the major efforts have been focusing on developing methods from the retrospective side, though recent efforts are being devoted to the prospective side (Collica and Taam, 1996; Hansen et al., 1997; Rogerson, 1999; Rogerson and Sun, 2001, 1999). It is argued that such methods for continuously monitoring spatial patterns should be valuable in practice. For example, in crime analysis, residents and law-enforce agents might be interested in methods which can help identifying *new*, *emerging* hot spots, so that resources can be allocated more efficiently. Potential hot spots might be prevented from developing further into serious problems in certain areas.

The basic idea for these prospective approaches is to combine the cumulative sum method, which will be briefly reviewed as

1082-4006/00/0701-42$5.00

follows, with some statistics measuring spatial patterns such as nearest neighbor statistic (Clark and Evans, 1954), spatial joint count, and Tangs's *C* (Tango, 1995) among others. The primary purpose of this paper is to devise a weighted-distance approach for monitoring spatial patterns of point events around fixed points, follow the ideas from the prospective side. To begin with, the cumulative sum method will be briefly introduced. After that, the distance-based method developed by Rogerson and Sun (1999) will be presented and evaluated. Then the weighted-distance approach, a revised version of the distance-based approach, is devised and tested. Finally, both the distance-based and the weighted-distance methods are adopted to explore the 1996 data of residential burglaries in the City of Buffalo, New York. Finally, the findings of this paper will be summarized

II. THE CUMULATIVE SUM (CUSUM) METHOD

Cumulative sum (cusum) methods are widely used in industrial process control for monitoring variables over time (Wetherill and Brown, 1991; Hawkins and Olwell, 1997). These methods, summarized as bellow, can be adapted for use in monitoring spatial patterns of point data such as geocoded crime data, disease distribution, or new development of properties.

Suppose that successive independent values of a random variable (Z) come from a normal distribution with mean 0 and variance 1. One wishes to monitor Z and signal an alarm if the mean of Z becomes greater than zero. The one-sided cusum statistic at time t (for detecting upward trend), S_t^+ is defined as

$$S_t^+ = Max(0, S_{t-1}^+ + Z_t - k) \quad (1)$$

$$S_0^+ = 0$$

The value of k is a constant reference value chosen prior to the analysis, and is usually chosen to be equal to one half the magnitude (in standard deviation units) of the change one wishes to detect. Often k is chosen to be equal to 1/2.

As long as S^+ does not get too large, one can accept the idea that the mean of the Zs is zero; the process is said to be "in control." However, when the cusum is larger than a predetermined value, h, this indicates that large values of Z have accumulated and that the process is "out of control."

Similarly, to detect downward trend (decreasing means), cusum is defined as

$$S_t^- = Min(0, S_{t-1}^- + Z_t + k) \quad (2)$$

$$S_0^- = 0$$

with a signal when S_t^- is less than $-h$.

A measure of success of the monitoring system is the average run length (ARL) until an out of control signal is given. The value of h *is* chosen prior to monitoring so that the in control ARL will be equal to some desired level. In control ARLs define the times when a false alarm is sounded. Ideally one would like long ARLs when the process is in control, and short ARLs when the process goes out of control. The analyst must decide on a tradeoff: one can have quick rates of detection at the "price" of high false alarm rates, or one can have lower false alarm rates, along with the "penalty" of slower detection times when true change occurs. For example, when monitoring standard normal random variables, h = 4.0 and k=0.5 will yield an in control ARL of about 336, and when the mean changes from 0 to 1, the ARL will be about 17. With h=5.0 and k=0.5, the false alarm rate is lower (in control ARL is approximately 930), but the time to detection of a change when the mean shift from 0 to 1 is also higher (the out of control ARL in this instance is about 21). The other approach to determine the critical value of h is through simulations under the null hypothesis of no significant change of the variable under monitoring.

If the assumption of normal distribution is violated, the original data can be grouped with each group containing a couple of points (say, 5 points in each group). Then means of the groups can be assumed to come from a normal distribution.

III. THE DISTANCE-BASED METHOD

Rogerson and Sun (1999) devised a distance-based method for successively monitoring locations of points around a fixed focus. First, it calculates out the expected distance (d_{exp}) and variance (Var(d)) from a point to the focus under random patterns. Using these two parameters, the z value can be calculated for any successively observed point after the observed distance (d_t) from the point to the focus is obtained. That is

$$z_t = \frac{d_t - d_{\exp}}{\sqrt{Var(d)}} \quad (3)$$

where d is the observed distance between a point at time t and the focus; d_{exp} and Var (d) are the expected distance and variance between the point and the focus respectively under random patterns.

Then the variable z_t becomes the variable in the cusum equation (Equation 2).

As pointed out by Rogerson and Sun, the distance is not normally distributed and the observations can be grouped into batches of a certain number (b). It is found that the batch size of 3 approximately gives the normal distribution of the distance. Accordingly, the cusum formula for detecting downward trend of the distance becomes

$$s_t^- = Min(0, s_{t-1}^- + z_t + k/\sqrt{b}) \quad (4)$$

$$s_t^- = 0$$

and the critical value for s_t^- also changes from $-h$ to

$-h/\sqrt{b}$.

In the case where upward trend of the monitoring variable is needed, the corresponding cusum formula becomes

$$s^+ = Min(0, s_{t-1}^+ + z_t - k/\sqrt{b}) \qquad (5)$$

$$s_t^+ = 0$$

Rogerson and Sun also discussed how to monitor *changes* in point distribution around a focus, where it is necessary to decide the base period (training period) for which the mean ratio between the observed distance (d_t) and expected distance (d_{exp}) is obtained. Then d_{exp} and Var (d) in equation 3 are adjusted by the mean ratio calculated above for the base period.

Rogerson and Sun conducted series of tests on this method in the unit circle and found that the average run length (ARL) under the null hypothesis is about 755 and the median run length is about 514, with k and h set to be ½ and – 4.1 correspondingly. Under clustering, the average run length changes with different relative risk in the cluster. The higher the relative risk is, the quicker the cluster can be detected. It is also found that the method detects the clustering signal much quicker than the retrospective method of nearest-neighbor statistic.

The distance-based approach adopted the cusum formula for detecting downward trend in equation 4. If one puts aside the adjusting term "k" in equation 4 for the time being, what the distance-based method does is to cumulate the difference of the standardized distance between a point and the monitoring site. From the mathematical perspective, if z_t (adjusted by k) is negative, then s_t^- will decline; if z_t (adjusted by k) is positive, then s_t^- starts increasing. Therefore, if more negative z_ts are observed, the downward signals will be detected at certain point down the road. In the context of the distance-based method, when a cluster develops close to the monitoring site, more points than expected have shorter distance from the monitoring site; and hence, the clustering signal might be detected. On the other hand, when more points than expected are located farther away from the monitoring site, more z_ts will be positive, causing s^-_t to rise. Therefore, it would be harder to detect the clustering signal around the monitoring site. Furthermore, the closer a point is to the monitoring site, the bigger will be the absolute value for the negative z_t, leading to a quicker detection of the clustering signal. The reverse case is also true: the further a point is from the monitoring site, the bigger the value for the positive z_t, making it more difficult to detect the clustering signal at the monitoring site. When the distant cluster(s) is (are) strong enough, the distance-based method will fail to detect the clustering signal around the monitoring, even though there might exist a cluster developing around the monitoring site.

If one assumes the study area is a circle with radius R and the circle center is the monitoring site, then the expected distance between a random point in the circle and the circle center is $2R/3$ and the variance of the distance is $R^2/18$. Put it another way, the circle with the radius of R is divided into two areas, *the inner circle* and *the outer ring* (Figure 1), and the radius for the inner circle is $(2R/3 - R/2\sqrt{18})$. Therefore, the clustering signal around the circle center can be detected if more points than expected under random patterns fall within the inner ring no matter whether or not the cluster is centered at the unit circle center. However, when one or more clusters are under development, particularly when the cluster(s) in the outer ring is (are) stronger than the one within the inner circle, it might become harder to detect the clustering signal at the circle center. Therefore, the distance-based method might be interfered by those remote clusters.

To summarize, the distance-based method provides a means to monitor spatial patterns around a fixed site, and is most effective when there is one cluster under development. If the distance-based method detects a clustering signal around a site, there is no doubt that a cluster is occurring around the site up to the area of the critical ring, the radius of which is the expected distance minus one half of the standard deviation of the distance between a random point and the monitoring site. However, if there is no signal occurring, one needs to be careful in making further judgments, since a potential cluster might be still under development around the monitoring site, while the cluster is "covered" by a stronger cluster(s) developing in the outer ring.

Such analysis suggests a couple of approaches that can help develop more effective methods for monitoring spatial patterns. To use the distance-based method, the study areas maybe divided into a few smaller sub-regions, so that each sub-region may contain only one potential cluster, if any. Such an approach is particularly usefully when one is certain where the clusters might take place. For example, in studying impacts of pollution from nuclear plants on disease distribution, one may define a study area that contains only one nuclear plant so that there is only one possible cluster to occur. A second possible approach is to modify the distance-based method. Since the clustering

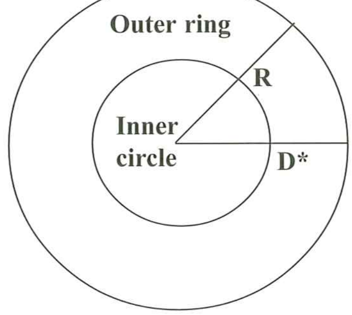

Figure 1. Defining the inner circle, and the outer ring

signal for a site might be affected by a distant cluster, one may want to lower the impacts of those observations far away from the monitoring site so that those closer to the focus get higher weights. The results following the latter approach will be reported in the next section.

IV. THE WEIGHTED-DISTANCE APPROACH

In the above distance-based method, the variable under monitoring is the distance between the observed point and the monitoring site. The essence of the method is to cumulate the difference between the observed distance (*ObsDis*) and the expected distance (*ExpDis*) from a point to the monitoring site. The weighted-distance method gives higher weight to the points close to the monitoring site, and then cumulate the difference between the observed weighted distance (*ObsWD*) and the expected weighted distance (*ExpWD*) under the null hypothesis of no clustering in the study area. Specifically, the monitoring variable is

$$WD_t = W_t /(D_t + 1) \qquad (6)$$

where WD_t is the weigted-distance for an observation at time t; W_t is the weight assigned to the observation at time t; and D_t is the distance from the monitoring site for a observed point at time t.

Essentially, the weighted-distance is inversely proportional to the distance between a point and the monitoring site. In the numerator the form "$D_t + 1$" is adopted instead of using "D_t", to avoid the problem with 0 values of D_t. Hence, the closer a point is to the monitoring site, the higher the value of WD_t will be. To detect clustering, I will use the cusum formula for detecting upward trend (equation 1) instead of the one for downward trend as in the distance-based method (equation 2).

W_t can be further obtained from the following formula,

$$W_t = W_c - (W_c - W_p) * D_t / D_{\max} \qquad (7)$$

where W_c is the weight assigned to the monitoring site; W_p is the weight assigned to the most distant location from the monitoring site in the study area; and D_{max} is the maximum distance from the monitoring site for a location in the study area.

Apparently, the weight for a point is in reverse proportion to its distance from the monitoring site. The closer a point to the monitoring site, the higher the weight value will be. If W_t in equation 7 is substituted for W_t into equation 6, the formula for WD_t in equation 6 becomes,

$$WD_t = (W_c - (W_c - W_p) * D_t / R)/(D_t + 1) \qquad (8)$$

This is a declining curve (Figure 1), given W_c, W_p, and R. As mentioned before, what the weighted-distance approach does is t_0 accumulate the difference between the observed weighted distance ($ObsWD_t$) and the expected value for WD_t (*ExpWD*) (under the null hypothesis of random patterns). $ObsWD_t$ can be calculated according to equation 8, given *Wc*, *Wp*, and *R*, and *ExpWD* can be obtained through the following formula,

$$ExpWD = \int_0^R \{ W_c - (W_c - W_p) * D_t / R \} * 2D_t / R^2 d_{D_t} =$$

$$\frac{W_c + W_p}{R} + \frac{2(W_c - W_p)}{R^2} - \frac{\ln(R+1)}{R^2} W_c - \frac{2\ln(R+1)}{R^3}(W_c - W_p)$$

$$\cong \frac{W_c + W_p}{R} \qquad (9)$$

For example, if one chooses W_c, W_p and R to be 10, 1 and 100 correspondingly and the expected value for *WD* ($ExpWD_{(10)}$) is approximately 0.11. Pulling together equations 8 and 9, once the observed value for weight-distance ($ObsWD_t$) is less than the expected one (*ExpWD*), the z value in the cusum will become negative, which probably will lead the cusum value to decline.

That is, the cusum will likely go up if the following holds

$$ObsWD_{(t)} < ExpWD \qquad (10)$$

or

$$(W_c - (W_c - W_p) * D_t / R)/(D_t + 1) < (W_c + W_p)/R \qquad (11)$$

Equation 11 is equivalent to the next equation,

$$D_t < D^* \qquad (12)$$

where

$$D^* = R/2 - 1/2 + \frac{W_p}{2W_c} \qquad (13)$$

In other words, when a point is less than D^* from the circle center, it will help develop a cluster signal at the monitoring site. Similarly, the circle with radius of R is divided into two areas (Figure 1): the inner circle and the outer ring, though the radius is reduced from approximately $2R/3$ to $R/2$. Therefore, once a point is located within the inner circle, $ObsWD_t$ will be less than *ExpWD*, likely leading to a clustering signal at the circle center. Whenever a point is in the outer ring, $ObsWD_t$ will be greater than *ExpWD*, making the clustering signal at the circle center harder to occur.

In comparison to the distance-based method, the radius of the inner circle in Figure 1 is pushed down from $2R/3$ to $R/2$, if R is really large, and the ratio between *Wp* and *Wc* is very small; hence, the cluster size is better defined by the weighted-distance method than the distance-based method.

In addition, *WD* declines rapidly at the beginning as a point moving away from the monitoring site. Even points away from the monitoring site in a short distance will get very small value of *WD*. Therefore, those distant points have little effects on the cusum value and the weighted-distance is less influenced by the remote clusters than the distance-based method. Such expectations are reconfirmed by results from simulations, which are not reported here to save space.

With those conclusions in mind, the next section examines one more issue about the weighted-distance method: what would happen if the weights (W_c and W_p) are changed?

From the above discussion, it is clear that the critical point for the weighted-distance method is how to define the inner circle and outer ring (Figures 1 and 2), and the issue is related to the value of D^* in equation 13. The effects of W_c and W_p can be explored by examining the relationship between D^* and W_c and W_p. From equation 13, it is clear that D^* will decline as W_c increases, while it will increase with increasing W_p. An implication of this is the size of inner circle may be manipulated by changing W_c or W_p. Unfortunately, such a capacity of weighted-distance approach is limited, and D^* can only be pushed down to ($R/2 - 1/2$).

*I*n conclusion, the weighted-distance approach allows us to make more accurate statements regarding the scope of the clustering area than the distance-based method. Furthermore, the weighted-distance approach is also less subject to distant clusters than the distance-based method, though it is still not very effective in dealing with multiple clusters.

As discussed by Rogerson and Sun (1999), the distance-based method can be used to monitor "*changes*" in spatial pattern around a fixed point as well by defining the base-period. Similarly, the weighted-distance approach can compare the observed weighted-distance to the expected weighted-distance, given the baseline pattern, instead of simply using the expected weighted-distance under null hypothesis of random patterns.

V. MONITORING SPATIAL PATTERNS OF RESIDENTIAL BURGLARIES IN THE CITY OF BUFFALO

As discussed above, the weighted-distance approach can achieve better delineation of the clusters, and both the methods can detect deviation from the *random* pattern and *changes* in spatial patterns. This section applies both the methods to the residential burglary data in City of Buffalo, NY. Due to the amount of computation, only selected 10 monitoring sites within City of Buffalo (Figure 3) were selected to illustrate how these methods can be applied in practice.

Figure 4 presents the results for monitoring the spatial patterns, using the distance-based method. Keep in mind that the critical value for the cusum is –2.38, with the observations grouped into 3 a batch. It is clear that residential burglaries became more clustered around the sites of 0, 1, 5, 7, and 8, and the patterns around site 2 and 4 experienced ups and downs, while the clustering around site 3, 6 and 9 was not sustained.

The weighted-distance approach gives different results (Figure

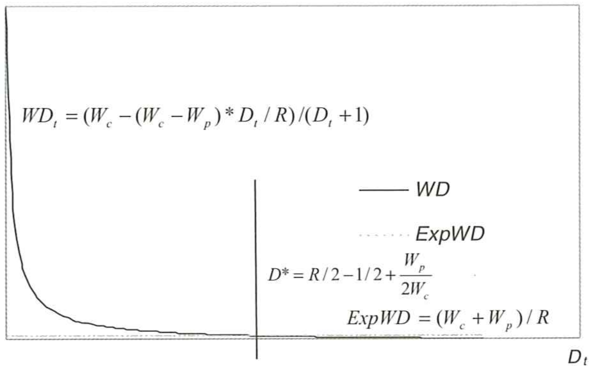

Figure 2. Relationship between weighted-Distance (WD_t) and the distance between a point and the monitoring site (D_t)

5), though similarities do exist among the two sets of results, particularly for the patterns around site 3, 6, 7, and 9. Such differences may indicate that the clustering took place in a broader area around these monitoring sites (up to 2/3 R in the unit circle), while no clustering was occurring in a more finely defined area around them (up to ½ R in the unit circle). The distance-based method shows that residential burglaries were clustering between observations 1400 and 2700 around site 4, while the weighted-distance method shows that the clustering lasted for the whole year. Such differences indicate that clustering might be happening in the more finely defined area around site 4, although such patterns were not so consistent for the broader area. Both methods show that clustering around 8 lasted for the whole year of 1996, while only occasional clustering signals occurred around site 3, 6, and 9. The results

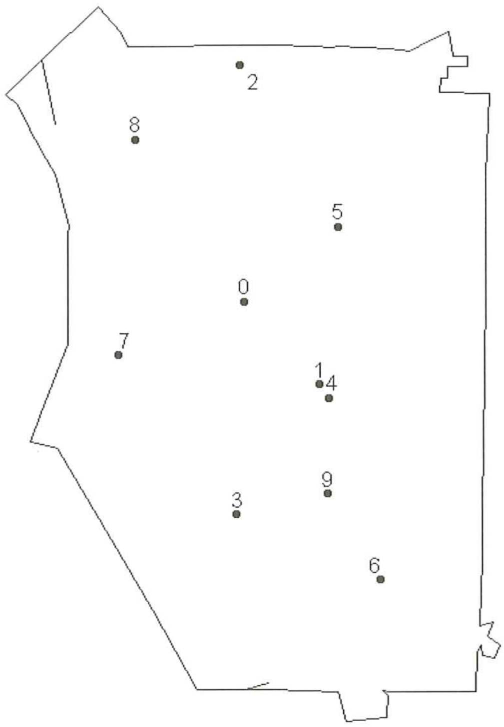

Figure 3. The monitor sites in the city of Buffalo

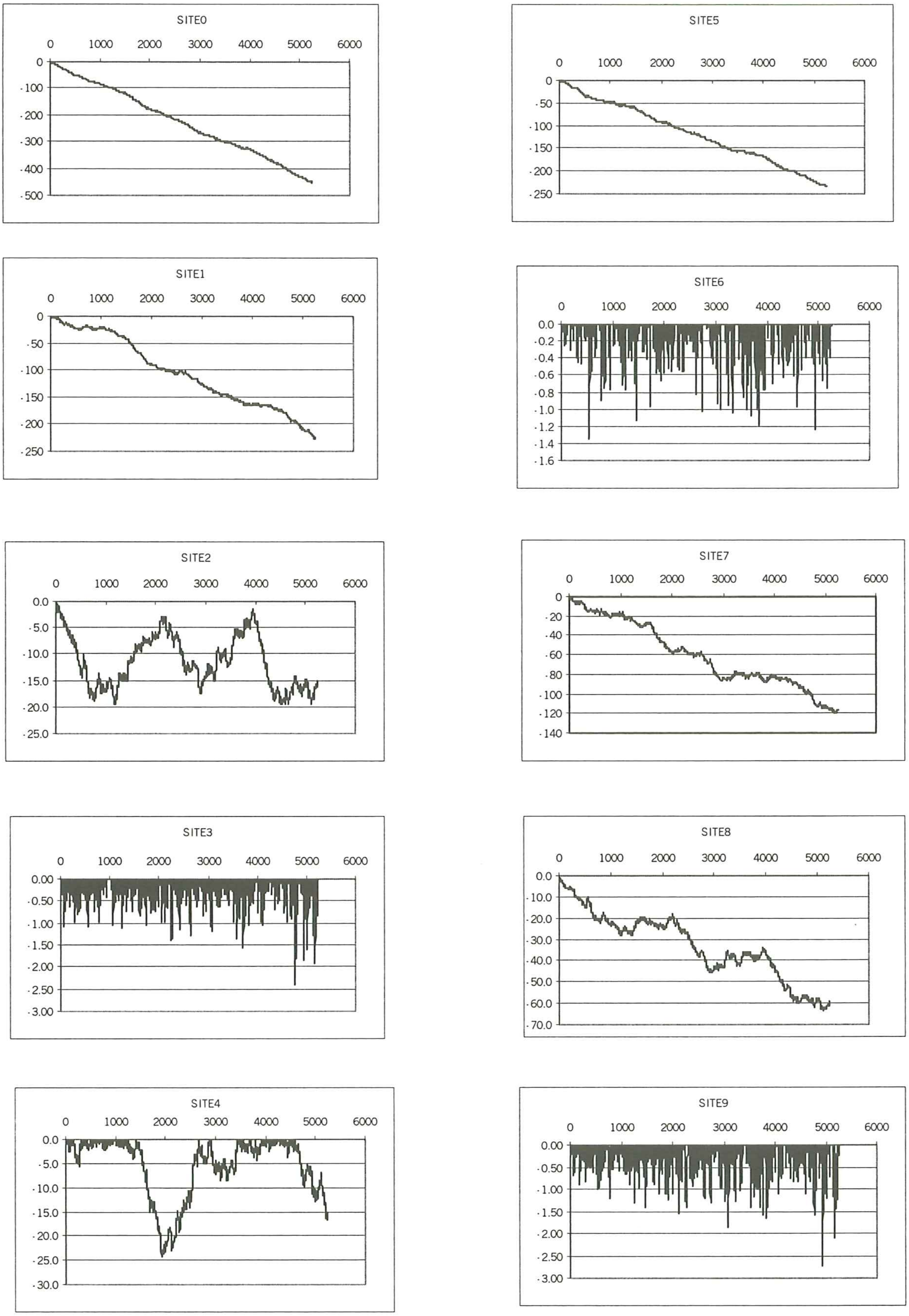

Figure 4. Monitoring spatial patterns of residential burglaries (distance-based method)

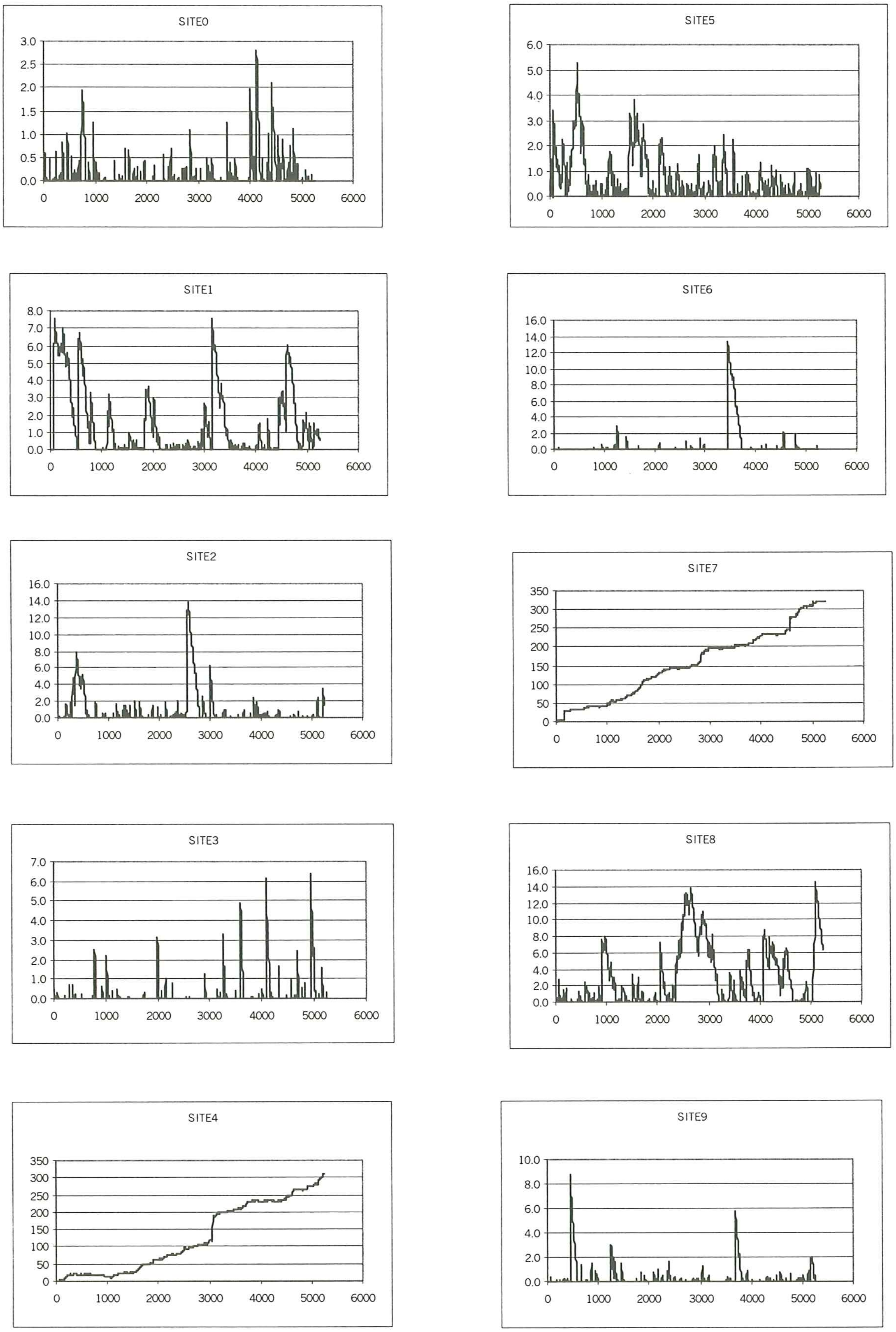

Figure 5. Monitoring spatial patterns of residential burglaries (weighted-distance method)

clearly demonstrate the advantage of combined usage of both the distance-based and the weighted-distance method: the distance-based method gives a broader picture, while the weighted-distance method depicts the picture more finely.

In the above tests, the two methods were used to monitor the deviation of the observations from random patterns. Here *changes* in spatial patterns will be monitored, where the first five months of 1996 were chosen as the base period (Figures 6 and 7). In other words, one can determine if the residential burglaries are becoming *more* clustered in the period from June 1996. The distance-based method does report occasional signals for some of the sites, though the signals do not sustain long enough for most sites (Figure 6). For instance, three individual signals were reported for sites 0, and all of the three signals did not last more than two continuous observations. Such results imply that no significant changes in residential burglaries occurred around the sites in the second half year of 1996. The only site that witnessed significant changes is site 7, for which the signal lasts from observation 2778 to 2880. However, the weighted-distance approach exhibits some different pictures (Figure 7). As in the distance-based method, the weighted-distance approach does not report signals lasting long or no signals at all for site 0, 1, 5, and 9. However, the weighted-distance approach does report lasting signals for the other site (2, 3, 4, 6, 7 and 8), the most of the signals are due to a few extreme observations. For example, for site 4, the cusum jumps up from 0 to 37.21 at the observation 3107, which causes the signal. The intensity of the signal is so high (37.21) that it lasts for 1920 observations (384 batches) until observation 4072, though the cusum value shows a declining trend during that period. Similar observations can be noted for site 6, 7, and 9. In comparison with the distance-based method, the weighted-distance approach seems to be more subjective to the observations extremely close to the monitoring focus. For example, if one chooses the values for R, Wc, and Wp to be 100, 10, and 1 correspondingly, the Z value for a point at the circle center would be 52. In other words, the cusum value would jump by about 52 point for the single observation. The clustering signal may last for a long time, though the points following the observation may not be clustering around the focus. Clearly this might be a weakness of the weighted-distance method.

VI. SUMMARY

A new method, the weighted-distance approach, is developed and evaluated for monitoring spatial patterns around fixed points. The effectiveness of this newly developed method is also compared with the distance-based method (Rogerson and Sun, 1999). It is found that both the distance-based method and the weighted-distance approach are effective when only one cluster is under development, while the signal will be affected by potential remote clusters. The weighted-distance approach better reveals the clustering area and is less influenced by remote clusters than the distance-based method, though it is strongly affected by close observations. Both these two methods can also be used for monitoring changes in spatial patterns, in addition to monitoring deviations of point distributions from pure random patterns, given the base-period. However, it seems that combining these two methods will achieve the best result. The illustration of monitoring the spatial patterns of residential burglaries in City of Buffalo did reveal some clustering signals and signals of changes in spatial patterns.

However, the weighted-distance method is still affected by the remote clusters, as in the case of distance-based method. Neither of these two methods addresses the issue of population at risk. It would be interesting to develop methods that can better reveal the cluster size, avoid the interference from the observations far away from the monitoring site, and take into account population at risk. It would also be valuable to explore questions such as why certain changes occur at specific time and locations.

ACKNOWLEDGMENTS

This paper is based upon work partially supported by Grant No. 98-IJ-CX-K008 from the National Institute of Justice, Office of Justice Programs, U. S. Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice. I want to thank Professor Peter A. Rogerson for his invaluable comments and suggestions. The reviewers' comments on the previous version also greatly improve the quality of the paper.

REFERENCES

[1] Anselin, L., 1995, Local indicators of spatial association-LISA. *Geographical Analysis*, 27: 93-115.

[2] Besag, J., and Newell, J., 1991, The detection of clusters in rare diseases. *Journal of Royal Statistical Society* A, 154: 143-155.

[3] Clark, P. J., and Evans, F. C., 1954, Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35: 445-453.

[4] Collica, R., R., and Taam, W., 1996, Process monitoring in integrated circuit fabrication using both yield and spatial statistics. *Quality and Reliability Engineering International*, 12: 195-202.

[5] Fotheringham, S. A., and Zhan, F. B., 1996, A comparison of three exploratory methods for cluster detection in spatial point patterns. *Geographical Analysis*, 28: 200-218.

[6] Fotheringham, S. A., and Rogerson, P. A. (eds), 1994, *Spatial Analysis and GIS* (London: UK: Taylor & Francis ltd).

[7] Getis, A., and Ord, J. K., 1992, The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24: 189-206.

[8] Hansen, M. H., Hair, V. N., and Friedman, D. J., 1997, Monitoring wafer map data from integrated circuit fabrication process for spatially clustered defects. *Journal of the American Statistics Association*, 39: 241-53.

[9] Hawkins, D. M., 1997, *Cumulative Sum Charts and Charting for Quality Improvement* (New York: Springer).

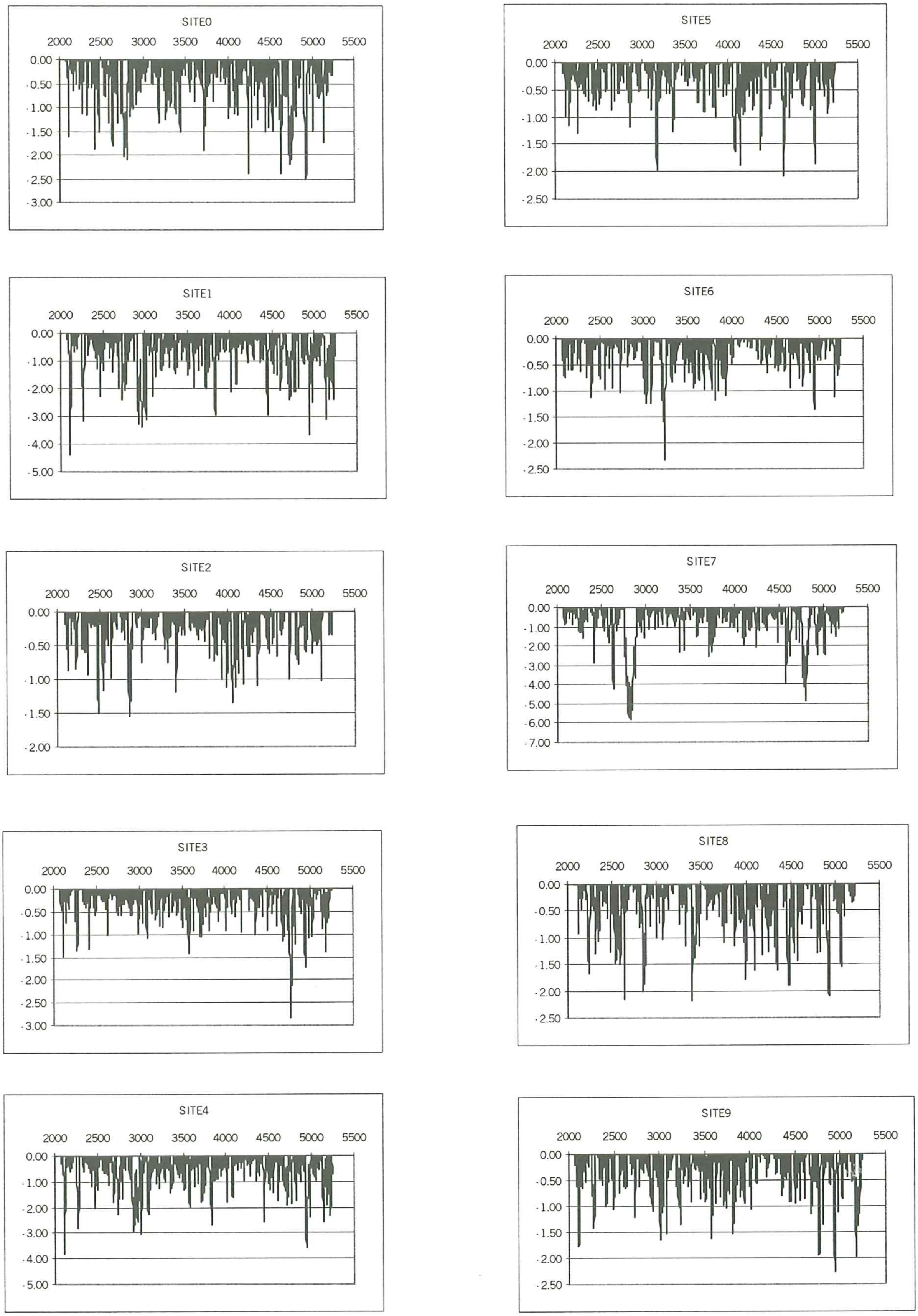

Figure 6. Monitoring changes in spatial patterns of residential burglaries (distance-based method)

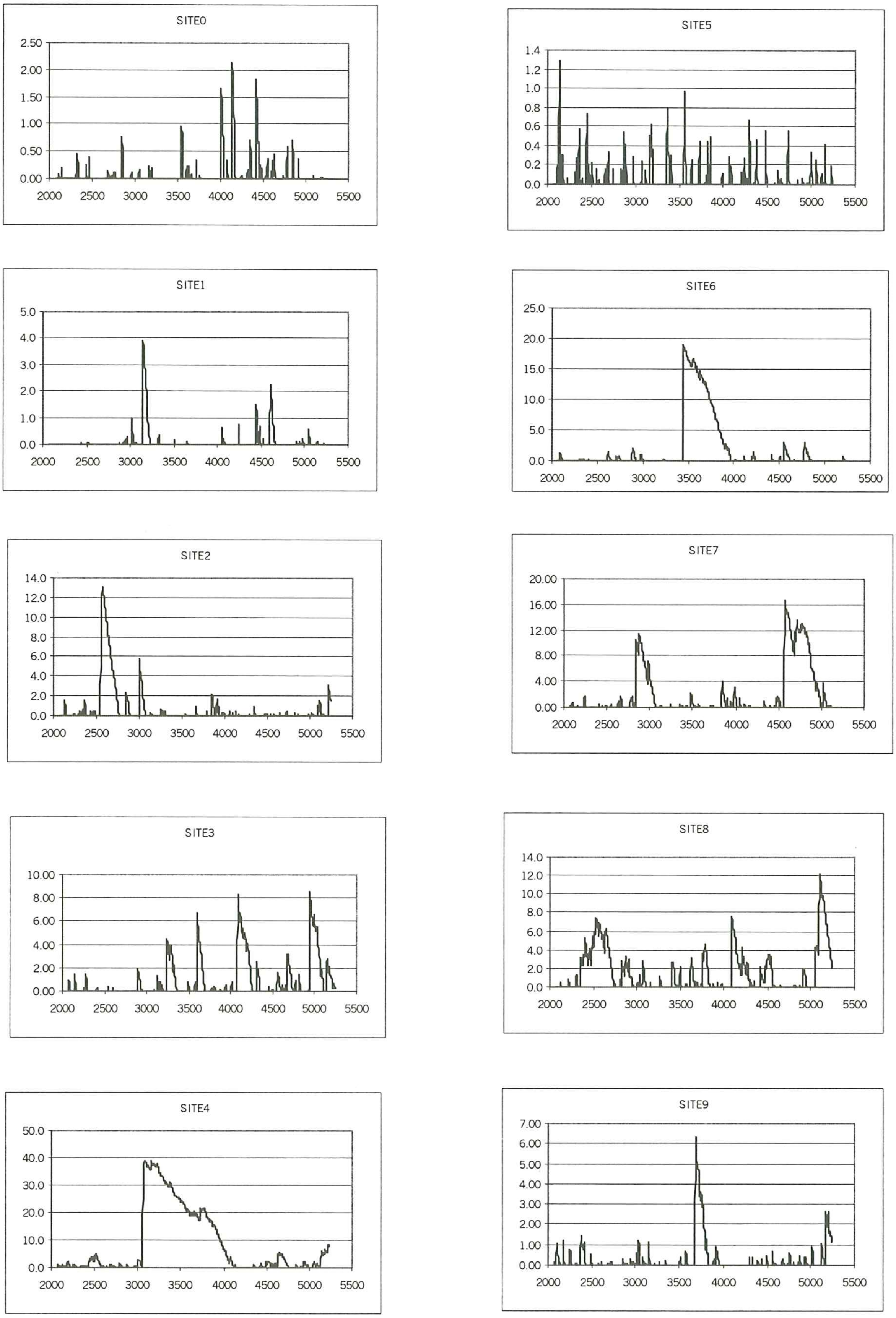

Figure 7. Monitoring changes in spatial patterns of residential burglaries (weighted-distance method)

[10]Kulldorf, M., and Nagarwalla, N., 1995, Spatial disease clusters: detection and inference. *Statistics in Medicine,* 14: 799-810.

[11]Openshaw, S. 1994, Two exploratory space-time-attribute pattern analyses relevant to GIS. In *Spatial Analysis and GIS*, edited by A.S. Fotheringham and P. A. Rogerson (London: Taylor & Francis ltd.), pp.83-104.

[12]Openshaw, S., Charton, M., Wymer, C., and Craft, A., 1987, A mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Sciences*, 1: 335-358.

[13]Ord, J. K., and Getis, A., 1995, Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27: 286-306.

[14]Raubertas, R., 1982, Analysis of disease surveillance data that uses the geographic location of the reporting units. *Statistics in Medicine*, 8: 267-71.

[15]Ripley, B. D., 1981, *Spatial Statistics* (New York: John Wiley & Sons).

[16]Rogerson, P. A., 1999, The detection of clusters using a spatial version of the chi-square goodness-of-fit statistic. *Geographical Analysis*, 31: 130-147.

[17]Rogerson, P. A., and Sun, Y., 2001, Spatial monitoring of geographic patterns: an application to crime analysis. *Computers, Environment, and Urban Systems, Forthcoming*.

[18]Rogerson, P. A. and Sun, Y., 1999, The monitoring of point incidents around a fixed location. Paper presented at GEOMED 99 in Paris.

[19]Stone, R. A., 1988, Investigations of excess environmental risks around putative source: statistical problems and a proposed test. *Statistics in Medicine*, 7: 649-660.

[20]Tango, T., 2000, A test for spatial disease clustering adjusted for multiple testing. *Statistics in Medicine*, 19: 191-204.

[21]Tango, T., 1995, A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Statistics in Medicine*, 14: 2323-2334.

[22]Turnbull, B. W, Iwano, E. J., Burnett, W. S., Howe, H. L., and Clark, L. C., 1990, Monitoring for clustering of disease: application to leukemia incidence in upstate New York. *American Journal of Epidemiology*, 132: s136-s143.

[23]Wetherill, B. G., and Brown B. W., 1991, *Statistical Process Control* (London: Chapman).