Araştırma Makalesi **BAUN Fen Bil. Enst. Dergisi, 21(1), 81-93, (2019)**
DOI: 10.25092/baunfbed.522967 **J. BAUN Inst. Sci. Technol., 21(1), 81-93, (2019)**

An application of spatial statistics: Spatial analysis of simulated fault plane geodetic points

Özlem TÜRKŞEN*

Ankara University Faculty of Science Department of Statistics, Ankara

Geliş Tarihi (Recived Date): 25.09.2018
Kabul Tarihi (Accepted Date): 16.01.2019

Abstract

In this study, simulated fault plane geodetic points are analyzed by using spatial statistics. The synthetic geodetic points are generated to understand the basic spatial structure of the fault plane because of the difficulty about obtaining real data set. The spatial statistics are applied to geodetic point data with three main items: (i) spatial descriptive statistics, e.g. spatial mean (center mean-CM), standard distance (SD), standard deviational ellipse (SDE), (ii) spatial pattern analyses, e.g. quadrat count method, the nearest neighbor approach, (iii) spatial autocorrelation, e.g. Moran's I index. It is seen from the application results that spatial autocorrelation should be taken into consideration during the spatial analysis of geodetic point data to understand if the surface displacements on the locations are clustered or not.

Keywords: *Spatial point analysis, spatial descriptive statistics, spatial pattern analysis, fault plane, geodetic points.*

Mekansal istatistiklerin bir uygulaması: Simule edilmiş fay düzlemine ilişkin jeodezik noktaların mekansal analizi

Özet

Bu çalışmada, mekansal istatistikler kullanılarak simule edilmiş fay düzlemine ilişkin jeodezik noktalar analiz edilmiştir. Fay düzleminin temel mekansal yapısını anlamak için gerçek verilerin elde edilmesinin zor olması nedeniyle yapay jeodezik noktalar üretilmiştir. Mekansal istatistikler, jeodezik nokta verilerine üç temel başlıkta uygulanmıştır. (i) mekansal betimsel istatistikler (mekansal ortalama, standart uzaklık, standart sapmalı elips), (ii) mekansal örüntü analizi (kuadrat sayma yöntemi, en yakın komşuluk yaklaşımı), (iii) mekansal otokorelasyon (Moran' ın I indeksi). Analiz sonuçlarından, jeodezik nokta verilerine mekansal analiz uygulanırken, yüzey yer

* Özlem TÜRKŞEN, turksen@ankara.edu.tr, https://orcid.org/0000-0002-5592-1830

değişimlerinin jeodezik noktalardaki konumuna bağlı olarak kümelenip, kümelenmediği anlamak için mekansal otokorelasyonun dikkate alınması gerektiği görülmüştür.

Anahtar kelimeler: *Mekansal nokta analizi, mekansal betimsel istatistikler, mekansal örüntü analizi, fay düzlemi, jeodezik noktalar.*

1. Introduction

The field of spatial statistics is a relatively new area of development in statistical researches and comprises a set of methods for describing and modeling spatial data. It is very important to know about the structure of the spatial data related to phenomena that occur in many areas, e.g. in health, in environment, in geology, in astronomy, etc. The data structures in spatial analysis can be categorized in three main items: (i) point data, (ii) line data, and (iii) area data. These are often called vector-based structures. The point data structure is the most encountered one among the vector-based approaches.

The spatial point data is distinguished by observations that are obtained at spatial locations $l_1, l_2, ..., l_n$ where the l_i, $i = 1, 2, ..., n$ are coordinates in the plane or space. The main idea is taking into account the spatial localization of the phenomena for understanding the spatial distribution of the point data. In order to obtain a useful summary of spatial distribution, spatial descriptive statistics are used for point data. It is well-known that the measures of center and dispersion are the most commonly used descriptive measures. The spatial mean and standard distance with standard deviational ellipse are used as spatial measures of the central tendency and the dispersion, respectively. Quadrat analysis and nearest neighbor method are the most used basic approaches to analyze that if the spatial pattern is clustered, random, or uniform. Spatial dependency is a key concept to understand the spatial relationship of the point data. The spatial dependency is measured with the computation of spatial autocorrelation. The spatial autocorrelation is a special case of a crossed products statistics and defines how the spatial dependency varies by comparing the values of a sample and their neighbors. Even there are several spatial autocorrelation tool, one of the most popular metric is Global Moran's *I* which gives a single summary value that describes the degree of spatial concentration or dispersion for the measured variable. The detailed information about the spatial data structure and wide knowledge about spatial analysis can be obtained from the books of [1-4].

Zimeras (2007) analyzed the spatial point patterns through spatial statistics to understand if there is any pattern that might help to make predictions about future earthquakes [5]. Sarp et al. (2007) defined the relationships between earthquake epicenters and faults and also predicted probable fault segments in the Northwest of the Ankara province by using spatial pattern analysis [6]. Ahmadi et al. (2013) applied spatial pattern analysis methods to a seismic data catalog of earthquakes beneath the Red Sea and aimed to explore global-local spatial patterns in the occurrence of earthquakes [7]. Tağıl and Alevkayalı (2013) aimed to detect clusters and explore spatial patterns in the occurrence of earthquakes in the Egean Region in Turkey using Geographical Information System (GIS) for the 1900-2012 seismicity catalog with event magnitudes larger than four [8]. Affan et al. (2016) applied spatial pattern analysis to detect and to cluster spatial patterns of earthquakes in the western part of Samatra Island during the period 1921-2014 using GIS [9]. Menekşe and Tağıl (2016) detected clusters of earthquakes in Turkey for the 2005-2015 earthquakes data with the event of magnitude larger from four by using spatial statistics [10]. Al-Dogom et al. (2018) applied spatial pattern analysis to examine the spatiotemporal occurrence of earthquake throughout the Arabian plate and their effect on the United Arab Emirates [11].

In this study, it is aimed to analyze the geodetic points of an earthquake fault plane area by using spatial statistics since the earthquake studies have a crucial role in the real world. However, obtaining real data is difficult as in many earth science problems. In this case, simulation studies are preferred. Synthetic geodetic points of a simulated earthquake fault plane area is analyzed by using spatial statistics, which are titled spatial descriptive statistics, spatial point pattern analysis, and spatial autocorrelation metric. The calculation results emphasized that analyzing the data set with considering the spatial information gives realistic results according to the phenomena. The rest of the paper is organized as follows: Section 2 gives brief information about spatial statistics for point data analyses. Fault plane structure is defined in Section 3. Simulated study is given in Section 4 as an application. Section 5 is composed with conclusion.

2. Fault plane structure

Prediction of earthquake occurence time is one of the crucial real world problem among many earth science problems. Estimation of fault plane parameters play an important role for determination of an earthquake occurence time since an earthquake occurence time can be defined by using fault plane parameters [12-14]. In earthquake studies, modeling of the surface displacements on the crust is really hard work. In order to get the point data for surface displacements, the geodetic locations should be well-defined according to the fault plane geometry. The fault geometry in three dimensions can be seen in Figure 1.

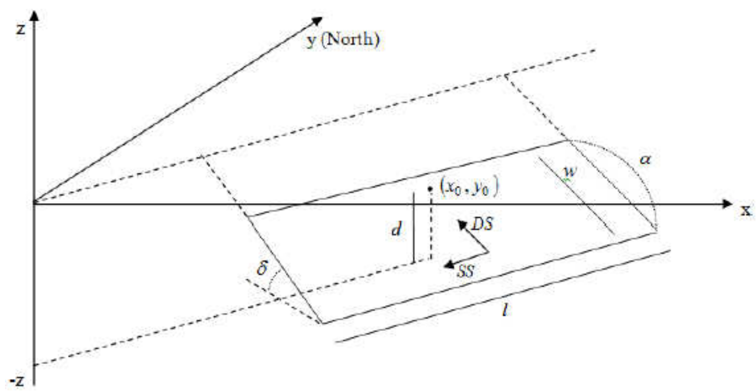

Figure 1. The fault plane geometry.

In Figure 1, a pair of coordinates, denoted as (x_0, y_0), can be considered as a geodetic point. Let's consider there are many geodetic points around the fault plane. The surface displacements, which are considered as the response values, are calculated by using the coordinates of geodetic points. Therefore, it is possible to say that the locations of geodetic points have importance for fault plane parameter estimation to predict earthquake occurence time. The spatial structure of geodetic points helps to analyze the fault plane area. However, sometimes, it is hard to obtain the real data. In this case, it is better to follow simulation studies as an alternative way of getting real data. From this perspective, in this study, an earthquake fault plane area is simulated and synthetic geodetic points are generated on the simulated fault plane area.

3. Spatial statistics for geodetic point data

A number of spatial statistics have been developed to understand the spatial structure of point data. Several earlier works are given in the book of [4]. It is well-known that the spatial patterns of a phenomena present fundamental clues about the nature of the spatial structure of geodetic points. The spatial distribution of the geodetic points is priorly quantified by incorporating *x*- and *y*- coordinates of the data structure. Then, it is better to apply spatial descriptive statistics to understand the basic spatial characteristics of simulated fault plane area. For this purpose, measuring centrality and dispersions of geodetic points over the fault plane area should be taken into consideration. The spatial point pattern analysis gives a visualization of geodetic points in fault plane if the points have uniform or random distribution or clustered. Besides, spatial autocorrelation of geodetic points should be defined. The spatial statistics, used for geodetic points, are summarized briefly in below.

3.1. Spatial descriptive statistics

3.1.1. Spatial mean (Mean center)

The spatial mean (mean center-*MC*) shows the central point of spatial distributions of events. It provides the average value of geodetic points for each of the *x*- and *y*-coordinates [4]. It is obtained by separately summing up the all *X*- and *Y*-values and dividing by the total number of geodetic points as follows:

$$\left(\bar{X}_{MC}, \bar{Y}_{MC}\right) = \left(\frac{\sum_{i=1}^{n} X_i}{n}, \frac{\sum_{i=1}^{n} Y_i}{n} \right) \tag{1}$$

where X_i and Y_i , $i = 1, 2, ..., n$ are the coordinates for geodetic point i and n is the total number of geodetic points.

3.1.2. Standard distance

Standard distance (*SD*) measures the spatial spread or variations of geodetic points. It also measures the extent to which geodetic points are dispersed around the *MC*. The mathematical formulation can be given as

$$SD = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}_{MC}\right)^2 + \sum_{i=1}^{n}\left(Y_i - \bar{Y}_{MC}\right)^2}{n}} . \tag{2}$$

3.1.3. Standard deviational ellipse

Standard deviational ellipse (*SDE*) is used to identify distributional trends of geodetic points. It is able to account for both distance and orientation. In order to obtain *SDE*, spatial mean, angle of rotation from the point of *MC*, and standard deviations along the *x*- and *y*- coordinates must be calculated [4]. The angle of rotation, θ, is obtained as follows:

$$\tan\theta = \frac{\left(\sum_{i=1}^{n} X_i'^2 - \sum_{i=1}^{n} Y_i'^2\right) + \sqrt{\left(\sum_{i=1}^{n} X_i'^2 - \sum_{i=1}^{n} Y_i'^2\right)^2 + 4\left(\sum_{i=1}^{n} X_i' Y_i'\right)^2}}{2\sum_{i=1}^{n} X_i' Y_i'} \quad (3)$$

where X_i' and Y_i', $i = 1, 2, ..., n$, are the deviations of *x*- and *y*- coordinates from the *MC*, calculated as $X_i' = X_i - \overline{X}_{MC}$ and $Y_i' = Y_i - \overline{Y}_{MC}$, respectively. The standard deviation along the *x*-axis is given by

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i' \cos\theta - Y_i' \sin\theta\right)^2}{n}} \quad (4)$$

and the standard deviation along the *y*-axis is given by

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i' \sin\theta - Y_i' \cos\theta\right)^2}{n}}\,. \quad (5)$$

3.2. Spatial point pattern analysis

3.2.1. Quadrat count method

The quadrat count method determines the geodetic point distribution by examining its density over the fault plane area. Analysis is based on subquadrats (or grid cells) that are constructed over a given fault plane area, denoted as *A*. After defining the fault plane area, the number of geodetic points per cell should be defined. Recall that count data is often modeled by a Poisson distribution where the rate parameter (λ) is also the mean and variance of the distribution. The ratio of mean count and the sample variance of the quadrat counts should be close to 1 in value if the counts are Poisson distributed. Deviations from 1 indicate deviations from spatial randomness [4]. The main steps for quadrat count method can be given below:

Step 1: Calculate the simulated fault plane area (*A*) and subquadrat area (κ) of *A*. The formula for κ is given as

$$\kappa = \frac{2A}{n} \quad (6)$$

where *n* is the number of geodetic points.

Step 2: By using κ, define the total number of quadrats, denoted with *m*.

Step 3: Determine the variance of geodetic points data by using the following formula

$$S^2 = \sum_{i=1}^{m} \frac{\left(X_i - \lambda\right)^2}{m-1} \quad (7)$$

where λ is equal to $\frac{n}{m}$ and X_i, $i = 1, 2, ..., n$, is the number of points in the *i*th quadrat.

Step 4: Compute the variance-mean ratio (R_{VM}) as follows:

$$R_{VM} = \frac{S^2}{\lambda}. \tag{8}$$

Step 5: Interpret R_{VM} statistics. If R_{VM} is too small (less than 1), the geodetic points appear more uniform than expected from a strictly random process. If R_{VM} is too big (greater than 1), the geodetic points are accumulated together which indicates the clustering.

Step 6: Statistical hypothesis test is applied to see that if the obtained result is statistically meaningful or not through using *t*-test, the test statistics is $t_C = \frac{R_{VM} - 1}{\sqrt{2/(n-1)}}$, with α nominal significance level.

3.2.2. The nearest neighbor approach

The nearest neighbor approach compares the distances between nearest geodetic points and distances that would be expected on the basis of chance or simply measures the distance between an individual geodetic point and its nearest neighbor. The approach computes the average distance between nearest neighbors in a point distribution (observed distance) and compares it to that of a theoretical pattern (expected distance) [4].

Step 1: Calculate the simulated fault plane area (*A*) and expected distance with the formula

$$r_{\exp} = \frac{1}{2\sqrt{\frac{n}{A}}} \tag{9}$$

where *n* is the number of geodetic points.

Step 2: Derive the distance between each geodetic point and its closest neighbor, denoted as d_i, $i = 1, 2, ..., n$.

Step 3: Determine the observed distance by using formula given below:

$$r_{obs} = \frac{\sum_{i=1}^{n} d_i}{n}. \tag{10}$$

Step 4: Compute the nearest neighbor ratio, denoted with R_{NN}, as follows:

$$R_{NN} = \frac{r_{obs}}{r_{\exp}}. \tag{11}$$

Step 5: Interpret R_{NN} statistics. If R_{NN} is equal to 1, the distribution of geodetic points is perfectly random. If R_{NN} is equal to 0, the distribution of geodetic points is completely clustered, and If R_{NN} is greater than 1, the distribution of geodetic points tends toward uniformity. It should be noted here that the R_{NN} statistics has range from 0 to 2.149.

Step 6: Statistical hypothesis test is applied to see that if the obtained result is statistically meaningful or not through using *Z*-test, the test statistics is $Z_C = \dfrac{r_{obs} - r_{\exp}}{0.26136 / \sqrt{n^2 / A}}$, with α nominal significance level.

3.3. Spatial autocorrelation

Spatial autocorrelation has crucial role for the use of statistical methods to analyze the spatial data. Strong spatial autocorrelation means that the surface displacements of the geodetic points are strongly related (whether positively or negatively). Therefore, it becomes possible to understand that how the spatial patterns change from the past to present or how the spatial patterns will change from present to the future [4]. The spatial autocorrelation can also be analyzed from either global or local perspective. The global autocorrelation is a whole-map property to understand if the spatial distribution of surface displacements presents clustering or not [15] The most widely used measure of global spatial autocorrelation in spatial statistics context is Moran's *I* index [16, 17]. For a set of *n* geodetic points for surface displacements, *I* is given as

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(U_i - \bar{U}\right)\left(U_j - \bar{U}\right)}{\sum_{i=1}^{n} \left(U_i - \bar{U}\right)^2} \tag{12}$$

where w_{ij}, $i, j = 1, 2, ..., n$ defines a priori which pairs of two locations i and j are likely to interact and called spatial weights, and U_i, U_j, $i, j = 1, 2, ..., n$, are surface displacements of the geodetic points. The collected weights are typically referred to as a spatial weights "matrix" W, of the same dimension as the number of observations $(n \times n)$ and with zero on the diagonal by convention. The term S_0 is then the sum of all the elements in the weights matrix, or $S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$ [15].

The computed Moran's *I* index value is compared with the expected value of *I*, denoted as $E(I)$, given as

$$E(I) = \frac{-1}{n-1}. \tag{13}$$

Note that, in large samples, the $E(I)$ will approach zero since the $n-1$ becomes larger with n. According to the obtained results, three basic comparisons are possible as follows:

i. If $I > E(I)$, the geodetic data points are clustered in fault plane area with respect to their surface displacement values.

ii. If $I < E(I)$, the geodetic data points are dispersed in fault plane area with respect to their surface displacement values.

iii. If $I \cong E(I)$, the geodetic data points have random pattern in fault plane area with respect to their surface displacement values.

In order to understand that if the obtained results have statistically significance, *Z*-test is applied. Here, the *Z*-test for Moran's *I* index can be given as

$$Z_I = \frac{I - E(I)}{\sqrt{Var(I)}} \tag{14}$$

where $Var(I) = E\left(I^2\right) - \left(E(I)\right)^2$ and

$$E(I) = \frac{\begin{pmatrix} n\left(\left(n^2 - 3n + 3\right)S_1 - nS_2 + 3S_0^2\right) - \\ \left(\dfrac{\sum_{i=1}^{n}\left(U_i - \bar{U}\right)^4}{\left(\sum_{i=1}^{n}\left(U_i - \bar{U}\right)^2\right)^2} \times \left(\left(n^2 - n\right)S_1 - nS_2 + 3S_0^2\right)\right) \end{pmatrix}}{(n-1)(n-2)(n-3)S_0^2} . \tag{15}$$

In Equation (15), the S_1 and S_2 are computed as $S_1 = \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left(w_{ij} + w_{ji}\right)^2\right) \Big/ 2$ and $S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2$.

4. Application

In this section, a numerical example is given to present the application of spatial statistics for geodetic point data. For this purpose, an operation region is simulated to present an earthquake fault plane area. The simulated earthquake fault plane area is formed in a study area, sized $(-30, 20) \times (-50, 20)$, with 50 geodetic points [13]. The simulated earthquake area and 50 locations of point data are signed on the graph in Figure 2.

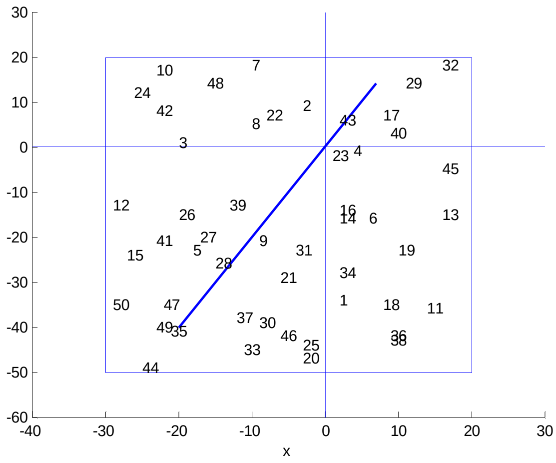

Figure 2. The locations of geodetic points of a simulated earthquake fault plane area with fault direction.

It can be easily seen in Figure 2 that the geodetic points are formed around the fault direction which is passed along the origin with a straight line. The surface displacements, calculated for each coordinate, are generated by using Matlab code taken from the geodynamics laboratory page of the Massachusetts Technology Institute for each coordinate [13]. The coordinates of geodetic points and surface displacement, denoted as *U* can be seen in Table 1. The detailed information about generating synthetic data set can be seen in the studies of [12-14].

Table 1. Coordinates and surface displacement values of simulated geodetic points [12-13].

Location number	**Coordinates**	**Surface displacements (*U*)**	**Location number**	**Coordinates**	**Surface displacements (*U*)**
1	(2, -34)	(-0.1275, 0.2515, -0.1019)	26	(-20, -15)	(0.0098, 0.1118, -0.0765)
2	(-3, 9)	(-0.0014, 0.0541, -0.0307)	27	(-17, -20)	(0.0128, 0.1496, -0.1237)
3	(-20, 1)	(0.0009, 0.0625, -0.0362)	28	(-15, -26)	(0.0101, 0.1950, -0.2330)
4	(4, -1)	(-0.0044, 0.0680, -0.0393)	29	(11, 14)	(-0.0020, 0.0417, -0.0222)
5	(-18, -23)	(0.0270, 0.1681, -0.1532)	30	(-9, -39)	(-0.3500, 0.7122, -0.2765)
6	(6, -16)	(-0.0197, 0.1046, -0.0600)	31	(-4, -23)	(-0.0339, 0.1756, -0.1282)
7	(-10, 18)	(-0.0008, 0.0432, -0.0231)	32	(16, 18)	(-0.0020, 0.0355, -0.0183)
8	(-10, 5)	(-0.0010, 0.0611, -0.0357)	33	(-11, -45)	(-0.6662, 1.0287, -0.1170)
9	(-9, -21)	(-0.0153, 0.1673, -0.1380)	34	(2, -28)	(-0.0680, 0.1887, -0.1022)
10	(-23, 17)	(-0.0001, 0.0385, -0.0204)	35	(-21, -41)	(0.0294, -0.7021, 1.0330)
11	(14, -36)	(-0.1098, 0.1485, -0.0293)	36	(9, -42)	(-0.1974, 0.2469, -0.0338)
12	(-29, -13)	(0.0144, 0.0736, -0.0367)	37	(-12, -38)	(-0.4933, 0.9630, -0.4768)
13	(16, -15)	(-0.0218, 0.0749, -0.0350)	38	(9, -43)	(-0.2121, 0.2581, -0.0308)
14	(2, -16)	(-0.0173, 0.1154, -0.0716)	39	(-13, -13)	(0.0001, 0.1155, -0.0820)
15	(-27, -24)	(0.0491, 0.1241, -0.0519)	40	(9, 3)	(-0.0041, 0.0562, -0.0309)
16	(2, -14)	(-0.0139, 0.1073, -0.0664)	41	(-23, -21)	(0.0309, 0.1325, -0.0877)
17	(8, 7)	(-0.0029, 0.0515, -0.0283)	42	(-23, 8)	(0.0005, 0.0484, -0.0264)
18	(8, -35)	(-0.1184, 0.1944, -0.0574)	43	(2, 6)	(-0.0023, 0.0569, -0.0323)
19	(10, -23)	(-0.0405, 0.1158, -0.0549)	44	(-25, -49)	(0.3093, -0.1076, 0.2482)
20	(-3, -47)	(-0.5171, 0.6714, -0.0800)	45	(16, -5)	(-0.0100, 0.0589, -0.0300)
21	(-6, -29)	(-0.0870, 0.2555, -0.1988)	46	(-6, -42)	(-0.3824, 0.6424, -0.1471)
22	(-8, 7)	(-0.0011, 0.0577, -0.0333)	47	(-22, -35)	(0.3027, 0.7685, -0.0648)
23	(1, -2)	(-0.0040, 0.0729, -0.0432)	48	(-16, 14)	(-0.0005, 0.0453, -0.0249)
24	(-26, 12)	(0.0005, 0.0412, -0.0217)	49	(-23, -40)	(-0.1164, 0.0907, 0.6267)
25	(-3, -44)	(-0.3982, 0.5752, -0.0957)	50	(-29, -35)	(0.1065, 0.1422, 0.1135)

The main aim of the study is to realize the structure of synthetic geodetic point data on the simulated fault plane area. The spatial mean center and the spatial standard distance are calculated as $\left(\bar{X}_{MC}, \bar{Y}_{MC}\right) = (-6.74, -17.08)$ and $SD = 24.4022$, respectively. The angle of rotation from the point of MC is calculated as 5.4899. The pair of standard deviations along the *x*- and *y*-coordinates are computed as $S_x = 17.959$ and $S_y = 17.8543$, respectively. The obtained values of spatial descriptive statistics are presented in Figure 3. It can be seen from the Figure 3 that the standard deviational ellipse is similar a circle which means that the distribution of geodetic points is uniform.

The quadrat count method is applied on simulated area to understand that if the geodetic points appear uniform, random or clustered. The total number of quadrats is calculated as 24 by following the algorithmic steps of quadrat count method given in Section 3.2.1. The fault plane area is divided the quadrats with 6 rows and 4 columns. It should be noted here that the number of rows and columns are defined by using area calculation given in the study of [4]. The quadrats with point pattern can be seen in Figure 4. The variance-mean ratio is computed as $R_{VM} = \frac{S^2}{\lambda} = \frac{0.6014}{2.0833} = 0.2887$. It is clear from the R_{VM} statistics value that the mean is bigger than variance. So, the geodetic points appear more uniform than expected from a strictly random process. In order to check that the obtained results is statistically meaningful or not, the calculated *t*-test value (t_C) is obtained as $t_C = -3.5208$. If it is compared with the *t*-table value at 0.05 significance level, the null hypothesis, which is defined as the distribution of point pattern is random, is rejected with 95% confidence level since $|t_C| > t_{0.025;49} = 2.010$.

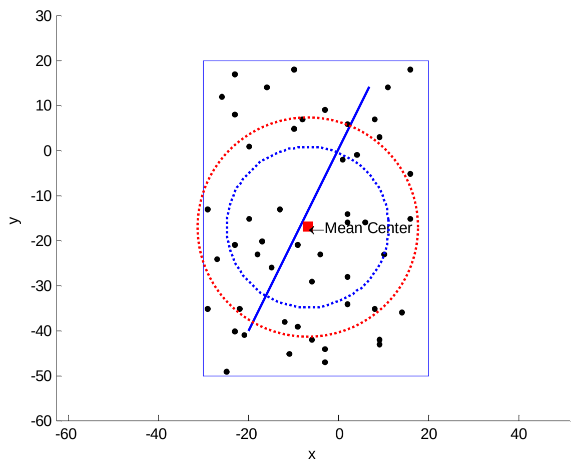

Figure 3. Spatial distribution of geodetic points in simulated earthquake area – meancenter (red square), standard distance (red dashed circle), standard deviational ellipse (blue dashed circle).

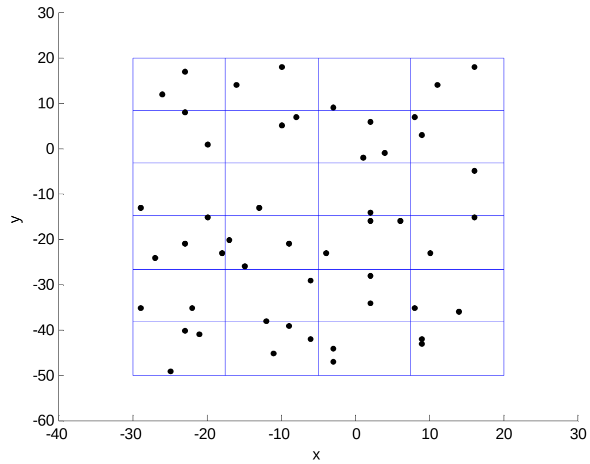

Figure 4. Quadrats with geodetic points on the simulated fault plane area.

If the nearest neighbor approach is applied to understand the spatial point pattern structure, the distance between each geodetic point and its closest neighbor are derived. The observed distance value, r_{obs}, and the expected distance value, $r_{\exp}$, are calculated as 5.0323 and 3.8827, respectively. Thus, the R_{NN} statistics is obtained as 1.2961. It can be said that the distribution of geodetic points tends toward uniformity since the R_{NN} statistics is greater than 1. For statistical hypothesis test, the calculated *Z*-test value (Z_C) is approximately obtained as equal to 4. The null hypothesis is rejected since $Z_C > Z_{0.025} = 1.96$ at 0.05 nominal significance level. It should be noted here that the null hypothesis is defined as the distribution of point pattern is random. The spatial descriptive statistics and spatial point pattern analysis present that the synthetic geodetic points have uniform distribution according to the locations of point data. However, it should be checked that if there is spatial autocorrelation between locations of geodetic points and the surface displacement values of these locations.

In order to compute the spatial autocorrelation, Moran's *I* index is preferred to use. The geodetic points which have similar features about surface displacements, along with the fault plane and dip angle with vertical direction, are chosen among the 50 geodetic points. These data points are numbered as 3, 5, 12, 15, 2, 26, 27, 28, 39, 41, 42 and 47 in Table 1. The locations of the grouped data points are also presented with red dots in Figure 5.

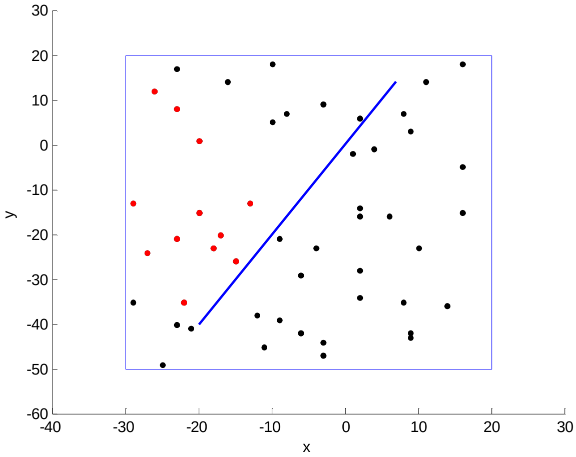

Figure 5. The geodetic data points with grouped geodetic points which are presented with red dots.

The spatial weight matrix, *W*, is calculated to obtain S_0 in Moran's *I* index formula given in Equation (12). By following this formula, the I index is obtained as 7.06×10^{-17}. The expected value of *I* is computed as -0.02 by using Equation (13). It is clear that the index value and the expected index value are approximately equal. So, it is possible to say that the geodetic data points have random pattern in fault plane area with respect to their surface displacement values. In order to understand if the obtained results have statistically significance, the *Z*-test value is calculated as 0.2938. It is clear from comparison with $Z_{0.025}=1.96$ value that the distribution of synthetic geodetic points is random in fault plane area with 0.05 nominal significance level. Here, the null hypothesis is also related to randomness.

5. Conclusion

This study presents the importance of spatial autocorrelation for spatial statistical analysis of geodetic points on fault plane area. The simulated study is applied since the difficulty of obtaining real earthquake data. The spatial descriptive statistics present that the distribution of geodetic points is uniform. The spatial pattern analysis results, quadrat count method and the nearest neighbor approach, indicate that the synthetic geodetic points have also uniform distribution. It is seen from the spatial statistical analysis results that the spatial descriptive statistics and spatial pattern analysis consider only location information of geodetic points. However, it is possible to obtain more realistic results with considering the spatial autocorrelation between geodetic points and the surface displacements. The global spatial autocorrelation index, Moran's *I*, considers the surface displacement values of the locations. It is understood from the calculations that the geodetic data points have random pattern in the fault plane according to the Moran's *I* index. This can be considered as an important result for the geodetic data set since the surface displacement values are analyzed related to the locations of geodetic points.

References

[1] Cressie, N.A.C., **Statistics for Spatial Data**, John Wiley and Sons, (1990).

[2] Moller, J., **Spatial Statistics and Computational Methods**, SpringerVerlag, NewYork, LectureNotes in Statistics, (2003).

[3] Fischer, M.M. and Wang, J., **Spatial Data Analysis, Models, Methods and Techniques**, Springer, (2011).

[4] Oyana, T.J. and Margai, F.M., **Spatial Analysis, Statistics, Visualization, and Computational Methods**, CRC Press, (2016).

[5] Zimeras, S., Modeling earthquake data using spatial statistics techniques, **Proceedings of the 1st IASME / WSEAS International Conference on Geology and Seismology (GES'07), Portoroz, Slovenia**, 33-36, (2007).

[6] Sarp, G., Düzgün, S. and Toprak, V., Spatial Analysis of Earthquake Epicenters in; North-West of Ankara, **International Conference on Environment: Survival and Sustainability, Near East University, Nicosia-Northern Cyprus**, 4619-4633, (2007).

[7] Al-Ahmadi, K., Al-Amri, A. and See, L., A spatial statistical analysis of the occurrence of earthquakes along the Red Sea floor spreading: clusters of seismicity, **Arabian Journal of Geosciences**, 7(7), 2893-2904, (2013).

[8] Tağıl, Ş. and Alevkayalı, Ç., Earthquake Spatial Distribution in the Egean Region, Turkey: The Geostatistical Approach, **The Journal of International Social Research**, 6(28), 369-379, (2013).

[9] Affan, M., Syukri, M., Wahyuna, L. and Sofyan, H., Spatial Statistic Analysis of Earthquakes in Aceh Province Year 1921-2014: Cluster Seismicity, **Aceh International Journal of Science and Technology**, 5(2), 54-62, (2016).

[10] Menteşe, S. and Tağıl, Ş., The Spatial Distribution of Earthquakes in Turkey: A Geostatistical and Spatial Statistical Approach, **The Journal of International Social Research**, 9(45), 408-414, (2016).

[11] Al-Dogom, D., Schuckma, K. and Al-Ruzouq, R., Geostatistical Seismic Analysis and Hazard Assessment; United Arab Emirates, **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,** 29-36, (2018).

[12] Türkşen, Ö., **The Adaptive Simulated Annealing Method in Estimation of Earthquake Source Parameters Through Surface Measurements**, Ms.Thesis, Ankara, (2005).

[13] Türkşen, Ö. and Apaydın, A., Estimating The Earthquake Source Parameters: Simulated Annealing Versus Nelder-Mead Simplex Algorithm, **Communications Faculty of Sciences University of Ankara**, Series A1, 62(2), 53-66, (2013).

[14] Türkşen, Ö. Estimation of Fault Plane Parameters by Using Stochastic Optimization Methods, **International Journal of Earthquake Engineering and Hazard Mitigation**, 2(2), 61-66, (2014).

[15] Anselin, L., Murray, A.T. and Rey, S.J., **Spatial Analysis, The Oxford Handbook of Quantitative Methods** (Edited by T.D. Little), Oxford University Press, (2013).

[16] Moran, P.A.P., The interpretation of statistical maps, **Biometrika**, 35, 255-260 (1948).

[17] Moran, P.A.P., Notes on continuous stochastic phenomena, **Biometrika**, 37, 17-23 (1950).