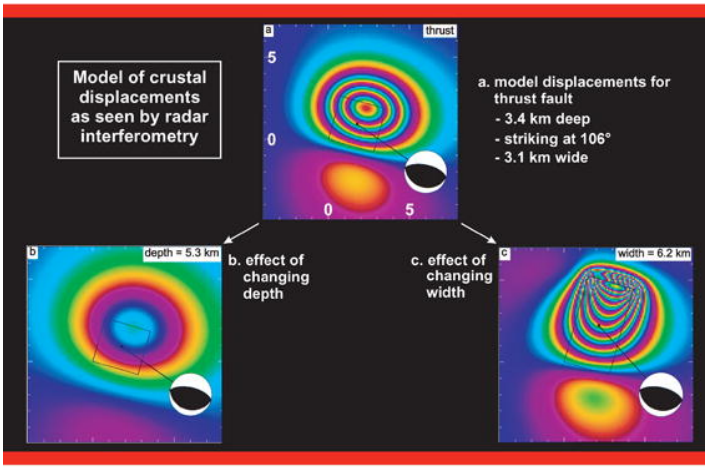
Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use.

Volume 33 Issue 7 July 2007 ISSN 0098-3004

COMPUTERS & GEOSCIENCES

An International Journal

Editor-in-Chief
Eric C. Grunsky
Geological Survey of Canada



See Editor's Home Page
http://www.iamg.org/CGEditor/index.htm
to download codes for programs discussed

This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

ELSEVIER

Computers & Geosciences 33 (2007) 952–965

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Delineation of support domain of feature in the presence of noise

Tao Pei[a,*], A-Xing Zhu[a,b], Chenghu Zhou[a], Baolin Li[a], Chengzhi Qin[a]

[a]*State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, CAS, 11A, Datun Road Anwai, Beijing 100101, China*

[b]*Department of geography, University of Wisconsin Madison, 550N, Park Street, Madison, WI 53706-1491, USA*

Received 15 May 2006; received in revised form 15 November 2006; accepted 17 November 2006

Abstract

Clustered events are usually deemed as feature when several spatial point processes are overlaid in a region. They can be perceived either as a precursor that may induce a major event to come or as offspring triggered by a major event. Hence, the detection of clustered events from point processes may help to predict a forthcoming major event or to study the process caused by a major event. Nevertheless, the locations of existing clustered events alone are not sufficient to identify the area susceptible to a potential major future event or to predict the potential locations of similar future events, so it is desirable to know the shape and the size of the region (the "territory" of feature events) that the feature process occupies. In this paper, the support domain of feature (SDF), the region over which any feature event has the equivalent likelihood to occur, is employed to approximate the "territory" of feature events. A method is developed to delineate the SDF from a region containing spatial point processes. The method consists of three major steps. The first is to construct a discrimination function for separating feature points from noise points. The second is to divide the entire area into a regular mesh of points and then compute a fuzzy membership value for each grid point belonging to the SDF. The final step is to trace the boundary of the SDF. The algorithm was applied to two seismic cases for evaluation, one is the Lingwu earthquake and the other is the Longling earthquakes. Results show that the main earthquakes in both areas as well as most aftershocks triggered by them fell into the estimated SDFs. The case study of Longling shows that the algorithm can deal with a region containing more than two processes.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Nearest-neighbor; Cluster; Spatial point process; Poisson process; Spatial data mining; Earthquake prediction; EM algorithm

1. Introduction

Many natural phenomena manifest themselves as spatial point processes which produce numerous events in space, such as earthquakes, landslides and craters. For clarification, we define an occurrence of a phenomenon located at a single point as an event in contrast to a simple geometric point. An event set is often not completely randomly distributed. Some events assemble in a restricted region while other events are dispersed over the remaining area. The former, distributed with higher intensity, is viewed as a cluster or "hotspot" (Brimicombe, 2003) and probably reveals some meaningful pattern (Jemal et al., 2002; Steenberghen et al., 2004), while the latter, distributed with lower intensity, is considered

*Corresponding author. Tel.: +86 10 64888960; fax: +86 10 64889630.

E-mail addresses: peit@lreis.ac.cn (T. Pei), axing@lreis.ac.cn, axing@geography.wisc.edu (A.-X. Zhu), zhouch@lreis.ac.cn (C. Zhou), libl@lreis.ac.cn (B. Li), qincz@lreis.ac.cn (C. Qin).

0098-3004/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2006.11.010

Author's personal copy

as noise or background (Fraley and Raftery, 2002). In many natural processes, the clustered events can be deemed either as a precursor that may induce a huge event to come or as offspring triggered by a huge event. Hence, the detection of clustered events from point processes may help us to predict a forthcoming major event or allow us to study the process caused by a major event (Wu et al., 1990; Ogata, 2001; Pei et al., 2006). The issue of how to detect a feature/cluster from a data set in the presence of noise has been extensively discussed (Ester et al., 1996; Byers and Raftery, 1998; Pei et al., 2006; Wang et al., 2006). Nevertheless, the locations of clustered events may not reveal the region in which a major event or successive events may take place, and the region may be very important in the case of constructing a natural disaster prevention plan, such as a quakeproof plan. It is thereby desirable to know the region (the "territory" of feature events) where future similar events might occur not only for identifying the area susceptible to a potential major future event but also for predicting the potential locations of future events. However, the locations of existing feature events alone are not sufficient to reveal the shape and size of the "territory" of clustered events. Therefore, it is necessary to construct a method that can precisely delineate the support domain of feature (SDF), which is referred to as the "territory" occupied by feature (clustered) events, in the presence of noise (background).

The problem of how to estimate the boundary of a point process, which is referred to as the convex hull of a point process, has been widely discussed (Ripley and Rasson, 1977; Davis et al., 1988; Hall et al., 2002; Chiu and Molchanov, 2003). However, their methods may overestimate the area of the support domain of a point process when it is distributed over a region with a concave shape, and also are restricted to only one process. If two Poisson processes with different intensities are superimposed, the support domain of one is included in that of the other. The methods for estimating the convex hull, because they are designed for approximating the boundary of one point process, will not be applicable due to the interference caused by the noise.

There are a number of studies addressing the issue on contouring clustered events in the presence of background noise. Banfield and Raftery (1993) proposed a model-based clustering method in which the clustered events are presumed to be mixtures of Gaussian density. The clustered events can be contoured at a specific value after the parameters of each Gaussian model have been derived. Fraley and Raftery (1998) constructed a multivariate normal mixture model for the sake of accommodating the different components originating from noise and feature. In their method, the Bayesian Information Criterion is utilized to determine the number of components. Jin et al. (2005) established a scalable model-based clustering framework. This algorithm is superior in analyzing large sets of complicated data and also significantly reduces the runtime as opposed to traditional algorithms. Although these methods may contour the clustered events in light of the difference in densities between the feature and noise, all of these ideas are built on the predefined model and therefore suffer from two limitations: a point process is rarely a Gaussian mixture and results from these methods are sensitive to departure from the model.

Allard and Fraley (1997) constituted a maximum likelihood estimator for a mixture of uniform point processes using the Voronoï tessellation defined by the data themselves. Although the method based on the Voronoï tessellation can determine the support domain of feature events by connecting the Voronoï polygons containing the feature events, it is limited in that the support domain of feature must be restricted to a single connected component without holes, and the boundary of features is of a specific geometry. Huo and Lu (2004) presented a digraph-based algorithm to estimate the boundary of the higher concentration regions (HCRs) from point processes. Although the method can adapt to HCRs with arbitrary shapes, users have to predefine the geometric constraint to the boundary of the HCR and specify the center(s) of the underlying region in advance.

In this paper, the support domain of feature is regarded as the "territory" occupied by feature events. In fact, the SDF is defined as the region in which a given feature event is equally likely to occur anywhere. An algorithm is developed in this paper to delineate the SDF from a region containing spatial point processes. The algorithm assumes that the entire region contains two point processes with different intensities. Any point in the region can be classified as feature or noise based on the distance between itself and its kth nearest event. Then, the entire region is divided into a mesh grid and each node is endowed with a fuzzy membership value of belonging to the SDF. Finally, the boundary of the

Author's personal copy

SDF is traced out from the entire region based on the mesh grid.

This paper is arranged in six sections. In Section 2, the definition of the SDF and the probability density functions (pdfs) of the kth nearest event (point)-to-event distances of Poisson process are reviewed. The algorithm for delineating the SDF is described in Section 3. Section 4 discusses the algorithm in terms of prior conditions, complexity, its extension to more than two processes and factors affecting the precision of the estimation of the SDF. In Section 5, two seismic cases, one in Northern China and the other in Southwestern China, are studied using the algorithm for delineating the potential zone in which strong earthquakes may occur. Conclusions and future work are discussed in Section 6.

2. Key concepts

2.1. The definition of the support domain of feature

A spatial point process can be defined as a stochastic model that governs the locations of events $\{s_i\}$ in A, where A is a subset of R^d. A Poisson process has the property that: conditional on $N(A)$ the number of events in the bounded region $A \subset R^d$, the events of the process are independently and uniformly distributed over A. That is, given $N(A) = n$, the ordered n tuple of events $(s_1, s_2, \ldots, s_n)$ in A^n satisfies

$$\Pr(s_1 \in B_1, \ldots, s_n \in B_n) = \prod_{i=1}^{n} (|B_i|/|A|), \quad B_1, \ldots, B_n \subset A,$$

where $|B| \equiv \int_B \mathrm{d}s$. Intuitively, this says that events are equally likely to occur anywhere within A and do not interact with each other, either repulsively (regularity of events) or attractively (clustering of events) (Cressie, 1991). It can be shown that $N(B)$ has a Poisson distribution with mean $\mu(B)$, for all $B \in A$:

$$P\{N(B) = n\} = \frac{(\mu(B))^n \mathrm{e}^{-\mu(B)}}{n!}, \quad n = 0, 1, \ldots,$$

where $\mu(B)$ is the intensity of the process. If the process is a homogeneous one, $\mu(B)$ is a constant number. A homogeneous Poisson process is a special case of a (inhomogeneous) Poisson process.

Region A is the support domain of the point process. If a region is occupied by two Poisson processes with different intensities, the area containing the process with higher intensity can be treated as the SDF, and the remaining area is viewed as the support domain of noise (SDN). The feature events, distributed with higher intensity, are equally likely to occur anywhere within the SDF, while noise events, distributed with lower intensity, have the equivalent likelihood to appear anywhere within the SDN. Our target is to trace the boundary of the SDF from the region in which a feature process and noise are superimposed.

2.2. The distance of event-to-event and that of point-to-event

In this context, we define the kth nearest event-to-event distance as the one from event s_i to its kth nearest-neighbor event $s_j (i \neq j)$ and the kth nearest point-to-event distance as the one from point $p \in R^d$ to its kth nearest event s_h. For clarification, the kth nearest event-to-event distance is denoted as W_k and the kth nearest point-to-event distance as X_k. The distance order of point (p_i) or event (s_i) refers to the ordinance of the nearest neighbor from p_i or s_i.

2.3. The probability density function of W_k and that of X_k

If a spatial point process is a homogenous Poisson process and is distributed in the region S, then we can find the distribution of the kth nearest event-to-event distance (W_k) from a randomly chosen event in the process to its kth nearest neighbor. For $x \in [0, \infty)$,

$$P(W_k \geqslant x) = \sum_{m=0}^{k-1} \frac{\mathrm{e}^{-\lambda\pi x^2} (\lambda\pi x^2)^m}{m!} = 1 - G_{W_k}(x), \quad (1)$$

where k is the parameter referring to the distance order, $G_{W_k}(x)$ is the cumulative distribution function (cdf) of W_k, λ is the intensity of the Poisson process (Byers and Raftery, 1998). The equation is obtained by conceiving a circle of radius x centered at the event under consideration. If W_k is greater than x, there must be one of 1,2,…,$k-1$ events in this circle. Its pdf $g_{W_k}(k, \lambda)$ is the derivative of $G_{W_k}(x)$:

$$g_{W_k}(k, \lambda) = \frac{\mathrm{d}G_{W_k}(x)}{\mathrm{d}x} = \frac{\mathrm{e}^{-\lambda\pi x^2} 2(\lambda\pi)^k x^{2k-1}}{(k-1)!}, \quad (2)$$

where λ and k are the same as those in Eq. (1).

Author's personal copy

Because the cdf of X_k shares the same form as (1) (Thompson, 1956), we can derive the pdf of X_k as

$$f_{X_k}(k,\lambda)=\frac{e^{-\lambda\pi x^2}2(\lambda\pi)^k x^{2k-1}}{(k-1)!}, \quad (3)$$

where k and λ have the same meanings as those in Eq. (1). In the region S containing a homogenous point process, the pdf of X_k and pdf of W_k share the same parameter, namely, the intensity λ.

3. The algorithm for delineating the support domain of feature

3.1. The Bayesian function for classifying the feature points and noise points

Suppose the feature process and noise in the region S are distributed within their own support domains with intensities of λ_1 and λ_2, respectively. X_ks of points within the SDF are distributed under the function of $f_{X_k}(k,\lambda_1)$. Similarly, X_ks of points in the SDN are distributed under the function of $f_{X_k}(k,\lambda_2)$. Because any point in S corresponds to its kth nearest point-to-event distance (X_k), points in S can be classified into feature and noise based on the difference in their X_k. In this paper, we employ the Bayesian function to accomplish the classification (see Eq. (4)).

$$P(\omega_i|X_k)=\frac{p(X_k|\omega_i)P(\omega_i)}{\sum_{j=1}^{2}p(X_k|\omega_j)P(\omega_j)} \quad (i,j=1,2), \quad (4)$$

where X_k represents the kth nearest point-to-event distance, ω_1 represents the feature and ω_2 represents the noise, $P(\omega_1)$ is the prior probability of belonging to SDF, $P(\omega_2)$ is the prior probability of belonging to SDN. $p(X_k|\omega_1)$ refers to the posterior pdf of belonging to SDF, and $p(X_k|\omega_2)$ represents the posterior pdf of belonging to SDN. Actually, $p(X_k|\omega_1)$ is the pdf of X_k of feature points and $p(x|\omega_2)$ is the pdf of X_k of noise points.

Three parameters remain to make the Bayesian function available, they are: the intensity (λ_1) of $p(X_k|\omega_1)$, the intensity (λ_2) of $p(X_k|\omega_2)$ and the prior probability $P(\omega_1)$ (prior probability $P(\omega_2)$ equals to $1-P(\omega_1)$). Our next job is to evaluate these three parameters.

3.2. Estimation of the parameters of the Bayesian function

3.2.1. Estimation of λ_1 and λ_2

Byers and Raftery (1998) proposed an Expectation–Maximization (EM) algorithm to estimate the intensity (λ_1) of feature events and that (λ_2) of noise events when two point processes are superimposed in a restricted region. The main idea is described as follows.

For a region S containing two point processes with different intensities, the pdf of W_k can be treated as a mixture pdf with different intensities, that is,

$$W_k \sim p g_{W_k}(k,\lambda_1)+(1-p)g_{W_k}(k,\lambda_2), \quad (5)$$

where p is the proportion coefficient, λ_1 and λ_2 are intensities of each distribution, k is the order of distance. The EM algorithm can be employed to evaluate the parameters of λ_1, λ_2 and p. For more details about the EM algorithm and its applications, please refer to Moon (1996) and Chawla et al. (2001).

The E-step in this context is

$$E(\hat{\delta}_i^{(t+1)})=\frac{\hat{p}^{(t)}g_{W_k}(k;\hat{\lambda}_1^{(t)})}{\hat{p}^{(t)}g_{W_k}(k;\hat{\lambda}_1^{(t)})+(1-\hat{p}^{(t)})g_{W_k}(k;\hat{\lambda}_2^{(t)})},$$

while the M-step is

$$\hat{\lambda}_1^{(t+1)}=\frac{k\sum_{i=1}^{n}\hat{\delta}_i^{(t+1)}}{\pi\sum_{i=1}^{n}w_{k,i}^2\hat{\delta}_i^{(t+1)}} \quad \text{and}$$

$$\hat{\lambda}_2^{(t+1)}=\frac{k\sum_{i=1}^{n}(1-\hat{\delta}_i^{(t+1)})}{\pi\sum_{i=1}^{n}w_{k.i}^2(1-\hat{\delta}_i^{(t+1)})}$$

$$\text{with } p^{(t+1)}=\sum_{i=1}^{n}\hat{\delta}_i^{(t+1)}/n, \quad (6)$$

where n is the number of points, $w_{k,i}$ is the kth nearest event-to-event distance of event q_i ($i=1,2,\ldots,n$) and t is the number of iterations. If we define the component with λ_1 representing the feature, then events with $\hat{\delta}_i^{(t+1)}\geqslant 0.5$ belong to feature and events with $\hat{\delta}_i^{(t+1)}<0.5$ can be viewed as noise. For more details, please refer to Byers and Raftery (1998).

3.2.2. Estimation of $P(\omega_1)$ and $P(\omega_2)$

For simplicity, we denote the area of the entire region, that of the SDF and that of the SDN as $|S|$, $|B_f|$ and $|B_n|$, respectively. Theoretically, the prior

Author's personal copy

probability of belonging to the SDF is equal to the ratio between $|B_f|$ and $|S|$, and the prior probability of belonging to the SDN is equal to the ratio between $|B_n|$ and $|S|$. However, the areas of those two support domains cannot be calculated directly until the membership of each point in the entire region is known.

Fortunately, the areas of the support domains of a process can be derived from the equation

$$|B_f| = \frac{N_f}{\lambda_f} \quad \text{and} \quad |B_n| = \frac{N_n}{\lambda_n},$$

where λ_f and λ_n are the intensities of feature and noise, respectively, N_f is the number of feature events, N_n is the number of noise events. All of them can be acquired from the EM algorithm which was discussed in Section 3.2.

If $|B_f|$ and $|B_n|$ are known, the prior probability of belonging to the SDF may be approximated by the equation below:

$$P(\omega_f) = \frac{|B_f|}{|B_f| + |B_n|} = \frac{N_f/\lambda_f}{N_f/\lambda_f + N_n/\lambda_n}, \tag{7}$$

and the prior probability of belonging to the SDN may be approximated by

$$P(\omega_n) = \frac{|B_n|}{|B_f| + |B_n|} = \frac{N_n/\lambda_n}{N_f/\lambda_f + N_n/\lambda_n}. \tag{8}$$

With the intensities and the prior probabilities being estimated, the fuzzy membership value of belonging to the SDF can be calculated from Eq. (4). The curve of Eq. (4) is a z-shape function. With this function, each point in the research region can be assigned a membership value of belonging to the SDF. As each node is endowed with a fuzzy membership value, a mesh grid covering the research region can be constructed. Then, from the mesh grid we can trace out the boundary of the SDF. For any point within the SDF, its fuzzy membership value of belonging to the SDF should be greater than 0.5.

3.3. Selection of the value of k in Eq. (4)

So far we assume that the k in Eq. (4) has been chosen by users. In fact, the estimated SDF is greatly influenced by the value of k in terms of the size, the location and shape.

Next an example is given to demonstrate how the estimated SDF evolves along with the increase of k. Fig. 1 shows the simulated data which is composed of a rectangle feature and a set of noise events. The

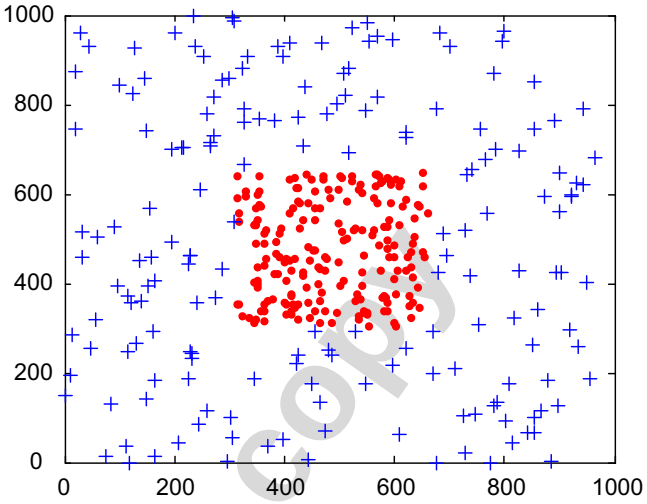


Fig. 1. Simulated data containing feature and noise. (Solid dots represent feature events and crosses represent noise events.)

intensity of feature is 0.00162 and that of noise is 2.0057×10^{-4}. Due to the edge effect, events or points near the edge (or the corner) of the region will have fewer neighbor events than those in the center of region. This may cause the W_ks (X_ks) of those events (points) to be overestimated and thereby increase the error of the classification. To correct the edge effect, the events in the region have been transformed into toroidal edge-corrected data. Interested readers are referred to Byers and Raftery (1998) for details.

In order to quantify the accuracy of the estimated SDF, we adopt three indices, namely, the error rate of the entire region, the false rate of the SDF and the missing rate of the SDF. The error rate of the entire region is referred to as the proportion of the misclassified area in the area of the entire region. The false rate of the SDF is defined as the ratio of the area of the false SDF to the area of the actual SDF. The missing rate of the SDF I s used to measure to what extent the actual SDF is excluded outside the estimated SDF.

Fig. 2 presents curves of those three indices versus k, which is calculated from the data in Fig. 1. It is found that the error rate of the entire region decreases below $k = 8$ and increases above it. The false rate of the SDN decreases from $k = 2$ to 5 and keeps stable at $k = 5$–8, then increases dramatically as k increases. The missing rate of the SDF decreases at all times until zero along with the increase of k.

The explanation for the trend of the error rate of the entire region mainly lies in that the increase of k exerts a two-side effect on the X_ks of points in the

Author's personal copy

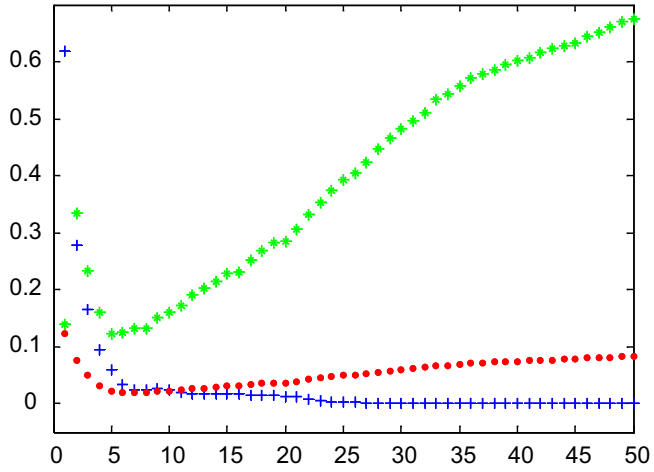


Fig. 2. Error rate of the entire region, false rate of SDF and missing rate of SDF versus k. (The dots represent the error rates of the entire region; the asterisks represent the false rate of the SDF; the crosses represent the missing rate of the SDF.)

region. We will illustrate the effect based on a comparison between the pdfs of X_k computed and the theoretical ones at $k = 3$, 7 and 35, which are presented in Fig. 3.

On one hand, the increase of k tends to reduce the error rate of the entire region. This can be illustrated by the overlapping area between the pdf of X_k of feature and that of noise, which changes with the increase of k and is shown in Fig. 3. When k is small, say, $k = 3$, the estimated pdf of X_k of noise points and that of feature points are not well separated and thereby produce a large overlapping area (the shade area in Fig. 3a) between them. This area represents the loss of the misclassification using Eq. (4) (interested readers are referred to Gelman et al., 1995 for more details). However, as k increases, those pdfs appear more distinct in bimodal and the overlapping area significantly shrinks at $k = 7$ and 35. The reason lies in that the increase of k enlarges the gap between the expectation of X_ks of feature and that of noise and breaks the mixed pdfs away by degrees. Hence the overlapping area is reduced, so is the classification error. We define this effect as the "separated effect".

On the other hand, the increase of k will distort the estimated pdfs as opposed to the actual pdfs. Compared with those at $k = 7$, the pdfs estimated from X_ks (the dashed line) at $k = 35$ significantly deviate from the actual pdfs (the solid line). This can be explained as follows: the noise points, especially those near the boundary of the SDF, own more feature events as their kth nearest neighbor with the increase of k. This will make the X_k s of these noise

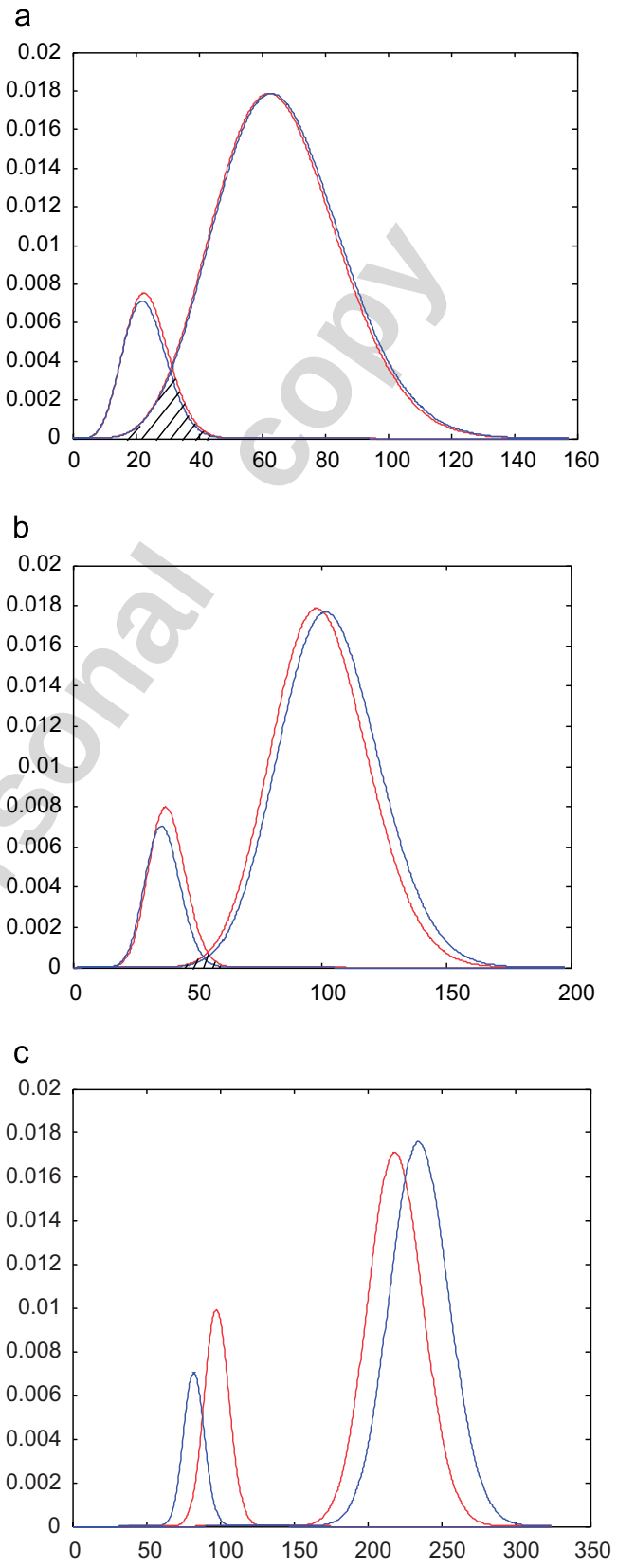


Fig. 3. Pdfs of X_k of feature and noise at different k values: (a) $k = 3$; (b) $k = 7$; (c) $k = 35$. (The solid line represents the theoretic pdf of X_k; the dashed line represents the pdf of X_k computed; the shade area represents loss of the misclassification using Eq. (4).)

Author's personal copy

points less on average compared with X_ks of those far away from the SDF. On the contrary, as k increases, the X_ks of the feature points near the boundary of SDF become larger on average compared with X_ks of those located in the center of the SDF. Hence, the X_ks of noise points and those of feature points, which are both near the boundary of the SDF, will be respectively underestimated and overestimated. Thus, the pdfs of X_ks for both noise points and feature points will deviate from the theoretical pdfs and the deviation of the pdfs affirmatively increases the error rate of the entire region. We define this phenomenon as the "inner edge effect". The increase of k will magnify the "inner edge effect" and will thereby increase the error rate of the entire region.

The "separated effect" is prone to decreasing the error rate of the entire region whereas the "inner edge effect" is prone to increase it. Only if these two effects reach a compromised point at some value of k, for example, $k = 7$ in this case, can the entire region achieve the minimum error rate. Furthermore, the "separated effect" plays a leading role in reducing the error rate of the entire region when k is at a small value, while the "inner edge effect" dominates the process and increases the error rate when k is of a moderate value or larger.

In addition, "the separated effect" and "the inner edge effect" also have a strong impact both on the missing rate of the SDF and the false rate of the SDF. When k is of a small value, both the false rate of the SDF and the missing rate are dominated by the "separated effect". The false rate of the SDF decreases dramatically at $k = 2$–5 while the missing rate declines rapidly at $k = 1$–6. These trends are also consistent with the change of the SDF in terms of the shape, area and the number of polygons. This can be found in Fig. 4 which demonstrates the growing process of the SDF along with k. From Fig. 4a we notice that at $k = 3$ the estimated SDF is composed of many polygons, among which the smaller ones undoubtedly are the false SDF. When $k = 7$ those small polygons almost disappear and the number of polygons is significantly reduced. Although a few small polygons, i.e. the false SDF, still remain, we can easily remove them by checking if each of their areas is smaller than a specific threshold. Then, the big polygon at the center of the region highly approximates the actual SDF in terms of the shape and the size.

When k is of a moderate value or larger, for example, $k > 8$, "the inner edge effect" begins to

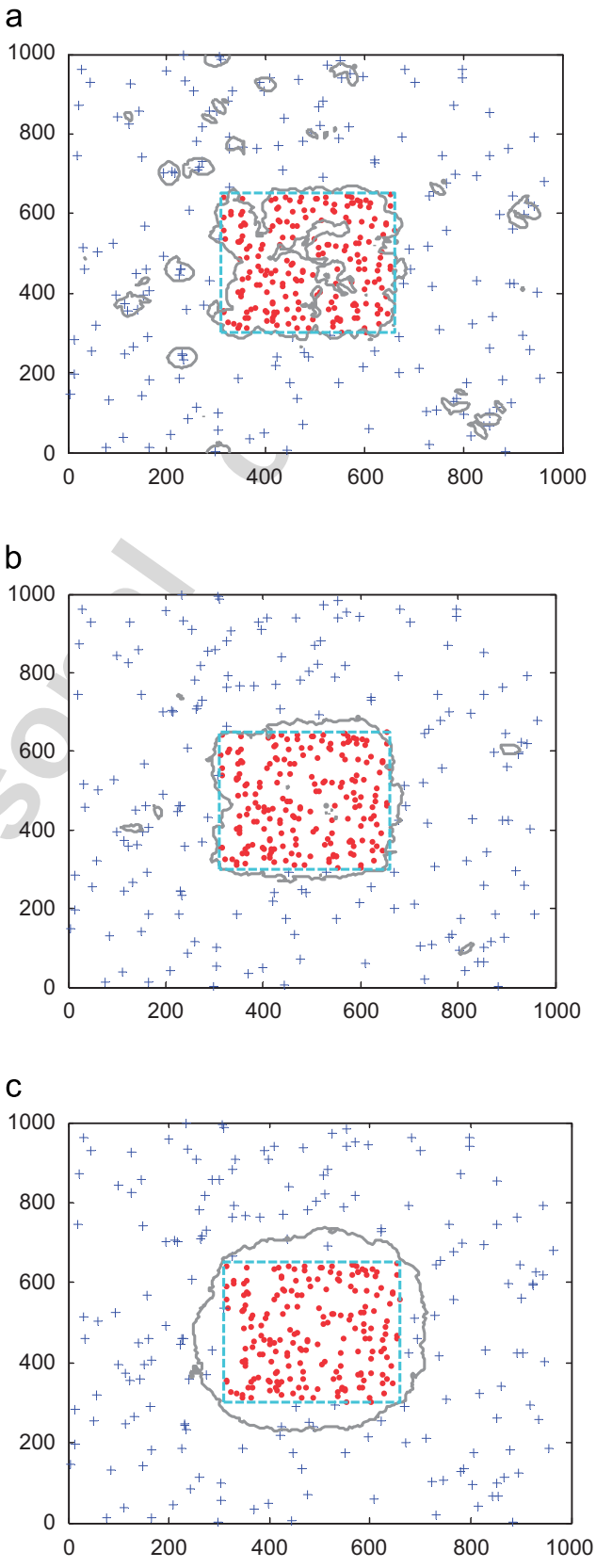


Fig. 4. Boundary of support domain of feature at different k values: (a) $k = 3$; (b) $k = 7$; (c) $k = 35$. (The rectangle bounded with a dashed line is the actual support domain of feature; the dots represent the feature events; the crosses represent the noise events.)

Author's personal

dominate the process. The effect tends to transform noise points outside the SDF into feature points as k increases rather than transform the feature points within SDF into noise points. The false rate of the SDF increases sharply when $k>8$, whereas the missing rate of the SDF keeps decreasing until zero at $k=28$. These curves also vary consistently with the behavior of the SDF in two dimensions. Compared with that at $k=7$ (Fig. 4b), the estimated SDF at $k=35$ (Fig. 4c) significantly expands and completely covers the actual SDF at the expense of producing a larger false positive area (the false rate arrives at 0.52).

In conclusion, the estimated SDF is an outcome of the interaction between "the separated effect" and "the inner edge effect" over varied values of k, and k is the only parameter to estimate the SDF when using this algorithm. In this paper, two aspects, i.e. the error rate of the entire region and the missing rate of the SDF, are taken into account when finding the optimum value of k. If one cares more about the error rate of the entire region or the accuracy of the SDF (namely, the size, the location and the shape of the SDF), then a moderate value, for example, $k=4$–8, is appropriate. Nevertheless, if one cares more about the missing rate of the SDF, a larger value ($k=9$–30) is appropriate, especially when predicting the susceptible area of the disaster that occurs in a point pattern. In that case, we would rather decrease the missing rate of the SDF at the expense of producing a larger false area.

3.4. Several other examples

In order to evaluate the algorithm, we apply it to four different simulated data sets, containing

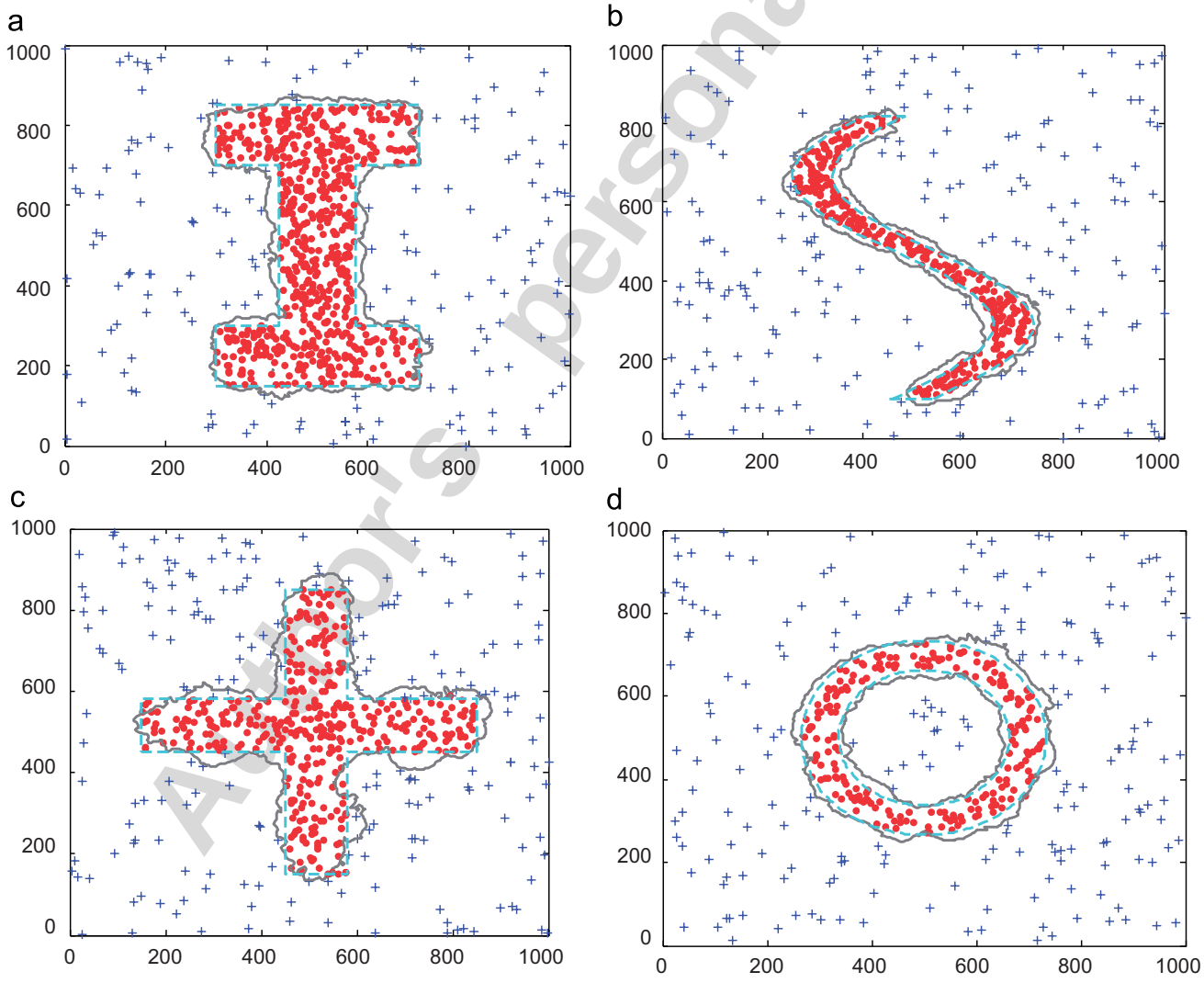


Fig. 5. Delineating SDFs from several simulated data sets at $k=6$. (The regions bounded with a solid line are the estimated SDFs and the regions bounded with a dashed line are the actual SDFs.)

Author's personal copy

features in the shape of 'I', 'S', cross and annulus. Fig. 5 shows the classification results of different simulated data at $k = 6$ using the algorithm. It is found that all the estimated SDFs are close to the actual ones. The results indicate that the algorithm can adapt to an arbitrarily shaped SDF.

4. Analysis of the algorithm

Four more issues need to be illustrated when applying the algorithm to identify the SDF. They are: (i) how to determine if a data set contains clusters; (ii) how to deal with a data set containing more than two point processes; (iii) the major steps of the algorithm and its complexity and (iv) the impact of the resolution of the mesh grid on the SDF.

4.1. How to determine if the data set contains clusters

There is a tremendous amount of literature focusing on how to determine if locations of a set of events are in complete spatial randomness (CSR, synonymous with a homogenous Poisson process), cluster or regular pattern (Eberhardt, 1967; Johnson and Zimmer, 1985; Cressie, 1991; Prayag and Deshmukh, 2000; Lucio and Brito, 2004). The methods for the determination are categorized into several types, i.e. the quadrat-based, the distance-based and the distance-based coupled with angle. Compared with the methods based on quadrat, the distance-based and those coupled with angle make use of precise information on the locations of events and have the advantages of not depending on arbitrary choices of quadrat size and shape, hence they have been adopted by more and more researchers (Liu, 2001). A number of test statistics based on the distance method have been proposed to determine if an event pattern is in CSR, aggregation (cluster) or regularity (Cressie, 1991). In this paper, the statistic of Clark and Evans (1954) is employed for the verification of aggregation because of their simplicity. In detail, the test statistic of $R_1 = 2(\lambda)^{1/2}\sum_{i=1}^{n} w_{1,i}/n$ has the asymptotic distribution of $N(1,(4-\pi)/n\pi)$ under CSR when n is large, where n is the number of events in the event set, $w_{1,i}$ is the first nearest event-to-event distance of event q_i $(i = 1,2,\ldots,n)$, λ is given by $n/(\pi\sum_{i=1}^{n} w_{1,i}^2)$. If R_1 is significantly greater than 1, then aggregation is indicated; if R_1 is significantly less than 1, then regularity is indicated. Interested readers may consult Cressie (1991) for details.

For the event set in Fig. 1, we obtain $R_1 = 0.8819$ with $n = 376$ and $\lambda = 3.7309 \times 10^{-4}$. When setting significant level (α) to 0.05, we can derive $c_{\min}$ and $c_{\max}$ from $F_N(c_{\min}) = \alpha$ and $F_N(c_{\max}) = 1-\alpha$, where the $F_N(\cdot)$ is the cdf of $N(1,(4-\pi)/n\pi)$. Here only $c_{\min}$ is needed for the test of the aggregation. Because $c_{\min} = 0.9904$ and $R_1 < c_{\min}$, we can conclude that the event set in Fig. 1 contains cluster(s).

4.2. The multi-scale algorithm for detecting SDFs

When a data set contains more than two processes, say, three processes, the feature/noise separated from the algorithm is still a mixed process in which two or more components remain. The remnant can be classified by reapplying the entire procedure until each remnant at the final stage is confirmed as a homogenous Poisson process, and we call this process the multi-scale algorithm. Thus, the data set can be segmented into several homogenous Poisson processes grouped as a binary tree.

Fig. 6 shows a simulated data set consisting of three Poisson processes and the result at $k = 7$ using the multi-scale algorithm. The segmentation scheme clearly reveals the SDFs at different scales.

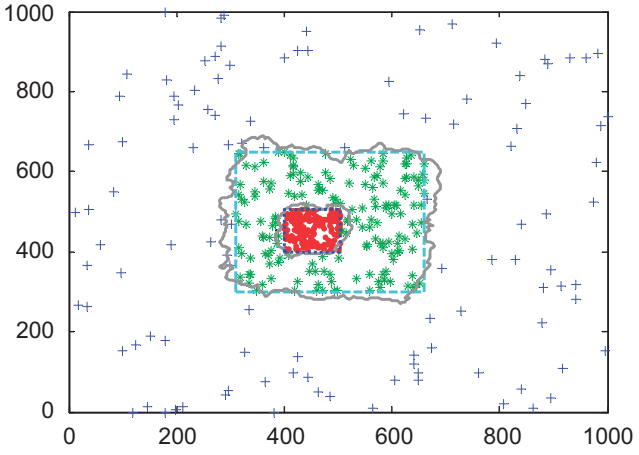


Fig. 6. Support domains of feature at $k = 7$. (The solid line indicates the boundaries of SDFs determined by the algorithm at different scales; the rectangle bounded with a dashed line represents the actual SDF occupied by the asterisks; the one bounded with a dotted line represents the actual SDF occupied by the dots.)

Author's personal copy

4.3. The major steps of the algorithm and its complexity

For the convenience of programming, we summarize the steps for delineating the SDF as follows:

(1) Compute the kth nearest event-to-event distances of each event (W_ks) within the study region.
(2) Determine if the data set contains clusters. If the data set contains clusters, then go to the next step, otherwise, end the algorithm.
(3) Evaluate the parameters of each component at k with the EM algorithm and derive the prior probability of belonging to the SDF and that of belonging to the SDN, then build the discrimination function (Eq. (4)) at k for the classification.
(4) Construct the mesh grid at a resolution of xn*yn, and compute X_k for each node. Then, assign each node with a fuzzy membership value acquired from Eq. (4).
(5) Trace out the boundary of the SDF based on the mesh grid.
(6) Determine if the subgroups, namely the feature events and noise, contain clusters. If yes, then go back to Step 2; otherwise, end the algorithm by outputting the boundary of the SDF.

The runtime of the algorithm is dominated by Steps 1, 3, 4 and 5. The substantial part of Step 1 is to query for the kth nearest neighbor of each event, where n is the number of events. The function of the query can be implemented by spatial access methods, such as R-tree (Beckmann et al., 1990), whose runtime is $O(\log n)$. Therefore, the complexity of Step 1 is k*n*$O(\log n)$. The complexity of Step 3 approximates to $O(m$*$n)$, where m is the number of iterations for the E-step and the M-step. Step 4 needs to query the kth nearest neighbor for each node of the mesh grid, and its complexity is $O(xn$*yn*k*$\log(n))$, where xn is the resolution of the mesh grid in x-direction and yn is the resolution of the mesh grid in y-direction. Step 5 involves tracing out the boundary of the SDF, and its complexity is $O(xn$*$yn)$. Consequently, the total complexity of the algorithm is k*n*$O(\log n)+O(m$*$n)+O(xn$*yn*$(k$*$\log(n)+1))$ if the data set under consideration contains only two point processes. When the data set contains more than two point processes, the increment of complexity will be less than the complexity of the first classification (for simplicity, we omit the proof). When xn*$yn \gg n$, the last item of the complexity formula, i.e. $O(xn$*yn*$(k$*$\log(n)+1))$, will dominate the runtime of the algorithm. The resolution is defined by users and will significantly influence the complexity of the algorithm.

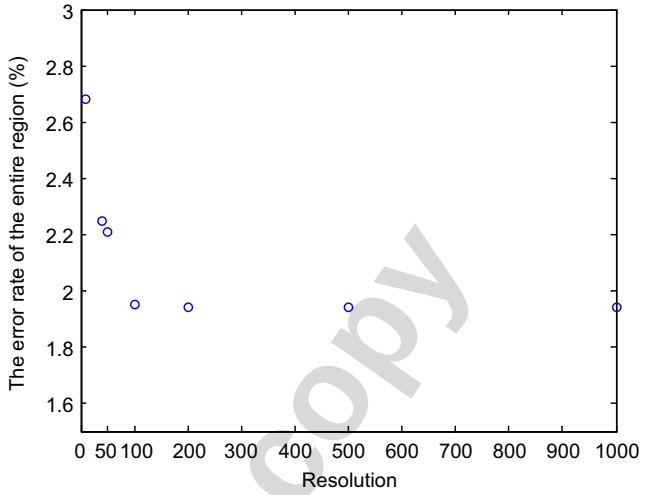


Fig. 7. Error rate of the entire region versus resolution of mesh grid.

4.4. The impact of the resolution of the mesh grid on the SDF

The resolution of the mesh grid affects not only the complexity of the algorithm but also the error rate of the entire region. Taking the data set in Fig. 1 as an example, we display the scatter plot of the error rate of the entire region versus the resolution of the mesh grid in Fig. 7. The plot shows that the error rate decreases sharply when the resolution is less than 100 and keeps stable when the resolution is greater than 100. The relationship between the error rate and the resolution implies that the finer resolution, or rather, more runtime, may not bring the equivalent accuracy after it exceeds some value, say, 100 in this case. Hence, we may adopt a moderate resolution to delineate the boundary of the SDF after balancing its precision against the complexity of the algorithm.

5. The case studies: identifying the susceptible area of strong earthquakes and aftershocks

5.1. The support domain of foreshocks and the susceptible area of strong earthquakes

Background earthquakes are referred to as small earthquakes that occur with a stable intensity within

Author's personal copy

Fig. 8. Locations of two research regions.

a certain region (Wyss and Toya, 2000; Pei et al., 2003). Clustered earthquakes are referred to as a swarm of earthquakes that occur within a limited region (Matsu'ura and Karakama, 2005). In this regard, background earthquakes and clustered earthquakes can be seen as two overlaid spatial Poisson processes with different intensities. As opposed to background earthquakes, clustered earthquakes occur with a higher intensity and are probably associated with a strong earthquake in the gestation mechanism (Reasenberg, 1999; Ogata, 2001; Kagan and Houston, 2005; Zhuang et al., 2005). Clustered earthquakes are viewed as foreshocks if they occur before a strong earthquake or as aftershocks if they occur after a strong earthquake. Therefore, clustered earthquakes could serve as a very important clue to predict strong earthquakes if the possibility of them being aftershocks can be excluded (Chen et al., 1999; Ripepe et al., 2000; Umino et al., 2002).

Although clustered earthquakes may be the precursors of a strong earthquake, their locations may not provide the exact shape, precise position and accurate area of the susceptible region of a strong earthquake. The determination of the potential location of strong earthquakes is still a subjective process when the foreshocks alone are employed for the prediction. If the foreshocks are presumed to be a homogeneous point process, the support domain of the foreshocks is the SDF and can be thought of as the region for generating a strong earthquake and is thereby deemed as the area susceptible to a strong earthquake. The algorithm discussed in this paper can supply an automatic and practical way to delineate the area susceptible to a strong earthquake. Its efficiency will be validated by two seismic cases in the next section.

5.2. The data set

China is one of the most seismically active countries in the world. There were about 21 strong earthquakes ($M \geqslant 6.5$) on the Chinese mainland from 1990 to 2005 and most of them were concentrated in several regions, namely, Northwestern China, Southwestern China, Northern China and Southeastern China.[1] Two regions were selected to test our algorithm, one is located in Northern China (region A) and the other is located in Southwestern China (region B) (see Fig. 8).

The seismic data presented in this paper were collected from the published seismic catalogs and imported into a database. Data about region A were acquired from the Institute of Geophysics China Earthquake Administration (1989a, b, 1991) and data about region B were from Feng and Huang (1980, 1989). All the records are after the year 1974. Due to the occurrence of the Xingtai earthquake in 1966, a devastating earthquake ($M = 7.2$) in Northern China with a death toll of 8064, the Chinese government paid more attention to monitoring and

[1]China Seismograph Network (CSN) Catalog. Available online at: http://www.csndmc.ac.cn (accessed in 2006).

Author's personal copy

predicting seismic activity and began to set up a nationwide seismographic network (Zhang, 1988). Over the following few years, more than 400 seismograph stations were built. The monitoring ability of seismic activity, especially in densely populated areas including Northern China and Southwestern China, has been greatly improved (Jiao et al., 1990). Hence, the quality of the data set involved in the case studies meets the requirement of this research.

5.2.1. *Detection of the susceptible area of the Lingwu earthquake*

Region A is located between 105° and 108°E, 36.5° and 39.5°N. The selected seismic data are from January 1st, 1985 to August 9th, 1987 and no less than 1.5 in Richter intensity (M). Thus, 131 epicenters have been obtained altogether.

When we identify the region in which strong earthquakes are likely to happen, what we are concerned with is the missing rate of the susceptible area of the main earthquake, namely the missing rate of the SDF in this context. In other words, we would rather make a larger false susceptible area of the main earthquake than miss a fraction of it. Therefore, we set $k = 20$ when employing the algorithm to trace out the boundary of the SDF for both of these two seismic cases. The support domain of the clustered earthquakes in region A is displayed in Fig. 9. Because no strong earthquakes occurred ahead of those clustered earthquakes in region A, we think they may be the foreshocks and their support domain may imply the susceptible area of a strong earthquake. This was proved by a strong earthquake that occurred later. The Lingwu earthquake ($M = 5.5$), occurred at 106.41°E and 38.11°N on August 10th, 1987, was later confirmed as the main earthquake of the foreshocks (Meng, 1996). Fig. 9 shows that the epicenter of the Lingwu earthquake was included in the support domain of the foreshocks. Interestingly, the epicenter of the Lingwu earthquake is not located in the central part of the clustered earthquakes, but to the east edge of the epicenters of the foreshocks. If the epicenters of the foreshocks alone are employed to predict the epicenter of strong earthquake, then the epicenter of the main earthquake will possibly be missed. The SDF may not only indicate the epicenter of the main earthquake but also delineate the possible area in which aftershocks will appear. The subsequent seismic records show that 43 aftershocks occurred during the next 90 days and all of them fell within the estimated SDF.

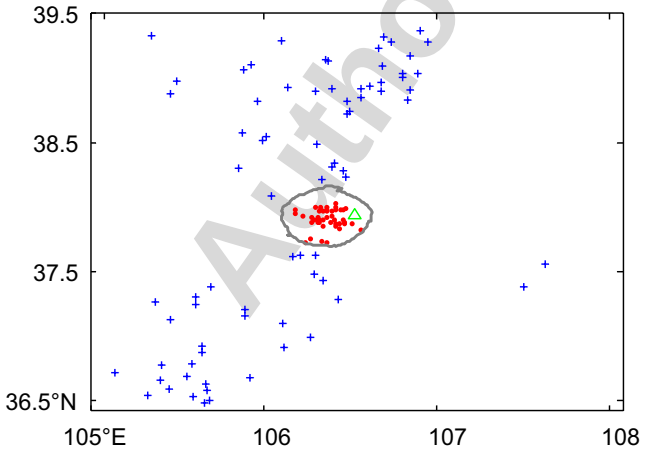


Fig. 9. Support domain of clustered earthquakes at $k = 20$. (The triangle represents main earthquake with $M = 5.5$.)

5.2.2. *The detection of the susceptible area of the Longling earthquakes and that of the aftershocks of the Simao earthquakes*

Region B is located between 98° and 102°E, 22° and 26°N, including a small part of Burma. The selected seismic data are from December 20th , 1975 to May 28th, 1976. A total of 543 epicenters were obtained.

We identified the clustered earthquakes with the multi-scale algorithm and delineated two SDFs with

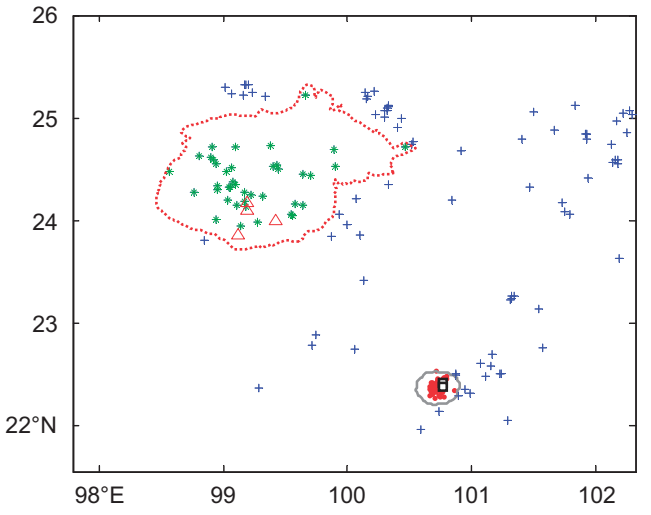


Fig. 10. Support domains of clustered earthquakes at $k = 20$ in different scales. (The squares represent the Simao earthquakes; the solid dots represent the aftershocks of the Simao earthquakes; the triangles represent the Longling earthquakes; the asterisks represent the foreshocks of the Longling earthquakes; the crosses represent background earthquakes; the region bounded with a dotted line is the support domain of the foreshocks; the region bounded with a solid line is the support domain of the aftershocks.)

Author's personal copy

different intensities (Fig. 10). Obviously, the northern one is larger in area but lower in intensity and the southern one is smaller but with a higher intensity. We may note them as domain N and domain S. The clustered earthquakes in these two domains are in different stages of seismic sequence. The earthquakes within domain S can be treated as aftershocks because two strong earthquakes ($M = 5.4$ and 5.1), which are called the Simao earthquakes, occurred on February 16th, 1976 and February 19th, 1976 (Zhang, 1990) and both of them appeared before the clustered earthquakes in domain S. On the contrary, since no strong earthquakes occurred before the clustered earthquakes in domain N, they can be deemed as foreshocks. Strong earthquakes that occurred later also proved that this was the case. In domain N, four main earthquakes ($M > 6$) occurred after 28th May, 1976, and all of the four epicenters fell into the domain N. These four strong events were located near the county of Longling and were thereby named the Longling earthquakes (Zhang, 1990).

According to seismic catalogs of the aftershocks of the Longling earthquakes, 2355 earthquakes were recorded during the 30 days after the occurrence of the main earthquakes and 2293 earthquakes fell into domain N. Since the foreshocks, the main shock(s) and the aftershocks are in the same local stress field, they are the outcomes originating from the same tectonic movement at different stages. Therefore, domain N, from which the foreshocks are assumed to come, may not only imply the epicenters of the strong earthquakes but also the epicenters of aftershocks.

6. Conclusions and future work

In the community of spatial data mining, the determination of feature events from a data set has attracted much attention and many different methods have been constructed. However, people often neglect to determine where these feature events originated and what the spatial relationship is between these events and forthcoming events if they belong to spatio-temporal processes.

The algorithm proposed in this paper is designed to delineate the boundary of the SDF from the research region. It is built on the theory of nearest-neighbor distance and requires only one parameter (k). Moreover, the algorithm cannot only accommodate an SDF of arbitrary shape, but can also adapt to event sets containing more than two point processes. Since the SDF delineated from a region containing point processes can be viewed as the "territory" of feature events, it can thus be employed to predict the location of successive events. The successful application of the algorithm to seismic data reveals that it can also adapt to other point process, such as landslides and distribution of epidemics. The algorithm proposed in this paper is limited to 2D point processes. Future work will be focused on extending the algorithm into higher dimensions.

Acknowledgments

This study was funded through support from the National Key Basic Research Special Foundation Project of China (Project Number: 2006CB701305), the Grants (Project Number: 40601078 and 40225004) from National Natural Science Foundation of China and the 'Hundred Talents' Program of Chinese Academy of Sciences. Support from the University of Wisconsin–Madison is also appreciated.

References

Allard, D., Fraley, C., 1997. Nonparametric maximun likelihood estimation of features in spatial point process using voronoi tessellation. Journal of the American Statistical Association 92, 1485–1493.

Banfield, J.D., Raftery, A.E., 1993. Model based Gaussian and non-Gaussian clustering. Biometrics 49, 803–821.

Beckmann, N., Kriegel, H.P., Schneider, R., Seeger, B., 1990. The R*-tree: an efficient and robust access method for points and rectangles. In: Proceedings of ACM SIGMOD International Conference on Management of Data, Atlantic City, NJ, pp. 322–331.

Brimicombe, A.J., 2003. A variable resolution approach to cluster discovery in spatial data mining. Lecture Notes in Computer Science 2669, 1–11.

Byers, S.D., Raftery, A.E., 1998. Nearest-neighbor clutter removal for estimating features in spatial point processes. Journal of the American Statistical Association 93, 557–584.

Chawla, S., Shekhar, S., Wu, W.L., Ozesmi, U., 2001. Modelling spatial dependencies for mining geospatial data: an introduction. In: Miller, H.J., Han, J.W. (Eds.), Geographic Data Mining and Knowledge Discovery. Taylor & Francis, London, New York, pp. 131–159.

Chen, Y., Liu, J., Ge, H.K., 1999. Pattern characteristics of foreshock sequences. Pure and Applied Geophysics 155, 395–408.

Chiu, S.N., Molchanov, I.S., 2003. A new graph related to the directions of nearest neighbours in a point process. Advances in Applied Probability 35, 47–55.

Clark, P.J., Evans, F.C., 1954. Distance to nearest neighbour as a measure of spatial patterns in biological populations. Ecology 35, 445–453.

Cressie, N.A.C., 1991. Statistics for Spatial Data, first ed. Wiley, New York, 900pp.

Author's personal copy

Davis, R.A., Mulrow, E., Resnick, S.I., 1988. Almost sure limit sets of random sample in R^d. Advances in Applied Probability 20, 573–599.

Eberhardt, L.L., 1967. Some developments in 'distance sampling'. Biometrics 23, 207–216.

Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, Oregon, pp. 324–331.

Feng, H., Huang, D.Y., 1980. A catalogue of earthquake in western China (1970–1975, $M \geqslant 1$). Seismological Press, Beijing, 679pp (in Chinese).

Feng, H., Huang, D.Y., 1989. Earthquake catalogue in West China (1976–1979, $M \geqslant 1$). Seismological Press, Beijing, 671pp (in Chinese).

Fraley, C., Raftery, A.E., 1998. How many clusters? Which clustering method? Answers via model-based cluster analysis. Computer Journal 41, 578–588.

Fraley, C., Raftery, A.E., 2002. Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association 97, 611–631.

Gelman, A., Carlin, J.B., Stern, H.S., Bubin, D.B., 1995. Bayesian Data Analysis, second ed. Chapman & Hall, London, 668pp.

Hall, P., Park, B.U., Turlach, B.A., 2002. Rolling-ball method for estimating the boundary of the support of a point-process intensity. Annales De L Institut Henri Poincare-Probabilites et Statistiques 38, 959–971.

Huo, X.M., Lu, J.C., 2004. A network flow approach in finding maximum likelihood estimate of high concentration regions. Computational Statistics & Data Analysis 46, 33–56.

Institute of Geophysics China Earthquake Administration, 1989a. The annal of the catalogue of earthquakes in China (1985). Seismological Press, Beijing, 461pp (in Chinese).

Institute of Geophysics China Earthquake Administration, 1989b. The annal of the catalogue of earthquakes in China (1986). Seismological Press, Beijing, 781pp (in Chinese).

Institute of Geophysics China Earthquake Administration, 1991. The annal of the catalogue of earthquakes in China (1987). Seismological Press, Beijing, 543pp (in Chinese).

Jemal, A., Kulldorff, M., Devesa, S.S., Hayes, R.B., Fraumeni, J.F., 2002. A geographic analysis of prostate cancer mortality in the United States, 1970–1989. International Journal of Cancer 101, 168–174.

Jiao, Y.B., Wu, K.T., Yang, M.D., 1990. Assessment about capability and quality of our country earthquake observation network. Earthquake Research in China 6, 1–7 (in Chinese with English abstract).

Jin, H.D., Leung, K.S., Wong, M.L., Xu, Z.B., 2005. Scalable model-based cluster analysis using clustering features. Pattern Recognition 38, 637–649.

Johnson, R.B., Zimmer, W.J., 1985. A more powerful test for dispersion using distance measurements. Ecology 6, 1669–1675.

Kagan, Y.Y., Houston, H., 2005. Relation between mainshock rupture process and Omori's law for aftershock moment release rate. Geophysical Journal International 163, 1039–1048.

Liu, C.R., 2001. A comparison of five distance-based methods for spatial pattern analysis. Journal of Vegetation Science 12, 411–416.

Lucio, P.S., Brito, N.L.C., 2004. Detecting randomness in spatial point patterns: a "stat-geometrical" alternative. Mathematical Geology 36, 79–99.

Matsu'ura, R.S., Karakama, I., 2005. A point-process analysis of the Matsushiro earthquake swarm sequence: the effect of water on earthquake occurrence. Pure and Applied Geophysics 162, 1319–1345.

Meng, G.K., 1996. Discussion on intensity assessments of recent earthquakes in Ningxia: Inland Earthquake 10, 330–336 (in Chinese with English abstract).

Moon, T.K., 1996. The expectation–maximization algorithm. IEEE Signal Processing Magazine 13, 47–60.

Ogata, Y., 2001. Increased probability of large earthquakes near aftershock regions with relative quiescence. Journal of Geophysical Research 106, 8729–8744.

Pei, T., Yang, M., Zhang, J.S., Zhou, C.H., Luo, J.C., Li, Q.L., 2003. Multi-scale expression of spatial activity anomalies of earthquakes and its indicative significance on the space and time attributes of strong earthquakes. Acta Seismologica Sinica 3, 292–303.

Pei, T., Zhu, A.X., Zhou, C.H., Li, B.L., Qin, C.Z., 2006. A new approach to the nearest-neighbour method to discover cluster features in overlaid spatial point processes. International Journal of Geographical Information Science 19, 153–168.

Prayag, V.R., Deshmukh, S.R., 2000. Testing randomness of spatial pattern using Eberhardt's index. Environmetrics 11, 571–582.

Reasenberg, P.A., 1999. Foreshock occurrence rates before large earthquakes worldwide. Pure and Applied Geophysics 155, 355–379.

Ripepe, M., Piccinini, D., Chiaraluce, L., 2000. Foreshock sequence of September 26th, 1997 Umbria–Marche earthquakes. Journal of Seismology 4, 387–399.

Ripley, B.D., Rasson, J.P., 1977. Finding the edge of a Poisson forest. Journal of Applied Probability 14, 483–491.

Steenberghen, T., Dufays, T., Thomas, I., Flahaut, B., 2004. Intra-urban location and clustering of road accidents using GIS: a Belgian example. International Journal of Geographical Information Science 18, 169–181.

Thompson, H.R., 1956. Distribution of distance to nth nearest neighbour in a population of randomly distributed individuals. Ecology 27, 391–394.

Umino, N., Okada, T., Hasegawa, A., 2002. Foreshock and aftershock sequence of the 1998 $M \geqslant 5.0$ Sendai, northeastern Japan, earthquake and its implications for earthquake nucleation. Bulletin of the Seismological Society of America 92, 2465–2477.

Wang, M., Leung, Y., Zhou, C.H., Pei, T., Luo, J.C., 2006. A mathematical morphology based scale space method for the mining of linear features in geographic data. Data Mining and Knowledge Discovery 12, 97–118.

Wu, K.T., Jiao, Y.B., Lu, P.L., Wang, Z.D., 1990. Panorama of Seismic Sequence, first ed. University Press, Beijing, 250pp (in Chinese).

Wyss, M., Toya, Y., 2000. Is background seismicity produced at a stationary Poissonian rate. Bulletin of the Seismological Society of America 90, 1174–1187.

Zhang, Z.C., 1988. Earthquake cases in China (1966–1975). Seismological Press, Beijing, 222pp (in Chinese).

Zhang, Z.C., 1990. Earthquake cases in China (1976–1980). Seismological Press, Beijing, 421pp (in Chinese).

Zhuang, J.C., Chang, C.P., Ogata, Y., Chen, Y.I., 2005. A study on the background and clustering seismicity in the Taiwan region by using point process models. Journal of Geophysical Research-Solid Earth 110 (B05S18).

Author's personal copy