

Spatio-temporal point pattern analysis on Wenchuan strong earthquake*

Peijian Shi¹ Jie Liu^{2,†} and Zhen Yang³

¹ Agricultural College, Jiangxi Agricultural University, Nanchang 330045, China

² China Earthquake Networks Center, Beijing 100036, China

³ Department of Precision Instrument and Mechanology, Tsinghua University, Beijing 100084, China

Abstract For exploring the aftershock occurrence process of the 2008 Wenchuan strong earthquake, the spatio-temporal point pattern analysis method is employed to study the sequences of aftershocks with magnitude $M \geq 4.0$, $M \geq 4.5$, and $M \geq 5.0$. It is found that these data exhibit the spatio-temporal clustering on a certain distance scale and on a certain time scale. In particular, the space-time interaction obviously strengthens when the distance is less than 60 km and the time is less than 260 h for the first two aftershock sequences; however, it becomes strong when the distance scale is less than 80 km and the time scale is less than 150 h for the last aftershock sequence. The completely spatial randomness analysis on the data regardless of time component shows that the spatial clustering of the aftershocks gradually strengthens on the condition that the distance is less than 60 km. The results are valuable for exploring the occurrence rules of the Wenchuan strong earthquake and for predicting the aftershocks.

Key words: Wenchuan earthquake; completely spatial randomness; spatio-temporal point pattern; K-function

CLC number: P315.08 **Document code:** A

1 Introduction

A strong earthquake with magnitude 8.0 occurred in Wenchuan region of Sichuan province of China on 12 May 2008. According to the data of Wenchuan aftershocks issued by China Earthquake Networks Center, the total number of aftershocks with $M \geq 4.0$ went up to 294 on 4 March, 2009. Among these aftershocks, the number of aftershocks with $M \geq 4.5$ was 127, and the number of aftershocks with $M \geq 5.0$ was 43. Exploring the characteristics of aftershock occurrence is practically valuable for reducing the loss caused by aftershocks. It can also provide a useful reference to study earthquakes in other regions.

Statistical seismology was originated when many statistical methods and stochastic models were applied to the seismic investigation. From the statistical side, an important impulse has been the development of the theory of stochastic point processes in statistical seismology (Vere-Jones et al, 2005). Some methods deriving from

space-time point processes can be employed to analyze the spatio-temporal characteristics of aftershock occurrence. Among these studies, the modified Omori formula is very famous for finding the descending law of aftershock frequency versus time. Since the 1980s, the epidemic-type aftershock sequences models (ETAS models) (Ogata, 1999 and the references therein) have been widely applied to the investigation of aftershock sequence because aftershocks can excite aftershocks and because there are some aftershock sequence deviating from the Omori formula (Department of Monitoring and Prediction, China Earthquake Administration, 2007).

The aftershocks could interact on a certain spatio-temporal scale. For example, at the early stage of Wenchuan aftershocks, if the quietude time of the aftershocks with $M \geq 4.7$ is above 16 h, the occurrence risk of aftershocks with $M \geq 5.5$ would increase in a few hours around the areas of the quietude-broken aftershock when the quiet state is broken. The stress triggering model in seismology is used to study strong earthquake interaction (Harris, 1998; Stein, 1999). The results from many earthquake cases show that the increase of Coulomb failure stress is beneficial to trigger the occurrence of

* Received 24 November 2008; accepted in revised form 9 April 2009; published 10 June 2009.

† Corresponding author. e-mail: liujie@seis.ac.cn

following middle-strong earthquakes and to cause influence of spatial distribution of aftershocks (Robinson and Zhou, 2005). In practice, there probably exists a particular interaction between different magnitude-level earthquakes, however, the extent of interaction is closely related to the stress environment and the media state in the region. Following a large earthquake, the stress field comes into a state of adjustment. Then a small aftershock can trigger another strong aftershock in a short time.

Thus, the key point of this paper is to find a suitable statistical method for describing the characteristics of spatio-temporal distribution of the aftershock activities. Diggle et al (1995) extended the K -function (Ripley, 1977) to the space-time K -function, and proposed the spatio-temporal point pattern analysis method. Considering that practical methods for space-time point processes have been accepted (Vere-Jones, 2009), we use the method of spatio-temporal point pattern analysis to examine whether the data of Wenchuan aftershock sequence exhibit space-time clustering (i.e., space-time interaction).

2 Basic principle of spatio-temporal point pattern analysis

2.1 Ripley's K -function and the method of complete spatial randomness (CSR) analysis

Theorem 1: For a stationary, isotropic, orderly process,

$$K(s) = \lambda^{-1} E(n_s),$$

where λ denotes the intensity; s denotes the distance scale; $E(\cdot)$ denotes the expectation; n_s denotes the number of further events occurring within distance s of an arbitrary event.

Theorem 2: For a homogeneous, planar Poisson process,

$$K(s) = \pi s^2.$$

Let n represent the total number of events in the studied region A , we have

$$\hat{\lambda} = \frac{n}{|A|}$$

and

$$\hat{E}(s) = n^{-1} \sum_{i=1}^n \sum_{j \neq i} w_{ij} I(d_{ij} \leq s),$$

where the hat represents the estimated value, $|A|$ denotes the area of the studied region A , and $I(\cdot)$ denotes the in-

dicator function. If the distance d_{ij} between event i and event j is more than the given distance scale s , $I(\cdot)=0$; if $d_{ij} \leq s$, $I(\cdot)=1$. w_{ij} denotes the edge-corrected weight, and it is reciprocal of proportion of circumference of circle with center x_i and radius d_{ij} , which is contained in A . Then the estimated $K(s)$ can be obtained

$$\hat{K}(s) = \frac{|A|}{n^2} \sum_{i=1}^n \sum_{j \neq i} w_{ij} I(d_{ij} \leq s). \quad (1)$$

In practice, the calculation of $\hat{K}(s)$ is not easy since, for arbitrary-shaped regions, the edge-corrected weights w_{ij} are hard to derive. Explicit formulae for w_{ij} can be handled for simple shapes such as rectangular or circular R (Diggle, 1983; Gatrell et al, 1996; Tang et al, 2003). Rowlingson and Diggle (1993) and Shi et al (2009) proposed different computer calculation algorithms about the weights w_{ij} for arbitrary-shaped regions. The method of CSR analysis is based on the assumption of spatial homogeneity. On a certain distance scale s , the theoretical values of $K(s) - \pi s^2$ should be zero for a homogeneous process without spatial dependence. Under regularity, $K(s) - \pi s^2$ should be less than zero. Under clustering, $K(s) - \pi s^2$ should be greater than zero. In spite of the homogeneous process without spatial dependence, $\hat{K}(s) - \pi s^2$ against s is actually in the state of oscillation around zero due to the existence of estimated errors. For analyzing that, the Monte-Carlo test is generally done. In the studied region, n random events are simulated with r times. Every time $\hat{K}(s) - \pi s^2$ is computed. Then r values of $\hat{K}(s) - \pi s^2$ are obtained in total. The maximum of $\hat{K}(s) - \pi s^2$ is regarded as the upper bound of the envelope; the minimum of $\hat{K}(s) - \pi s^2$ is regarded as the lower bound of the envelope. If $\hat{K}(s) - \pi s^2$ lies in the envelope on a certain distance scale s , the observed data are considered to exhibit the random distribution; if $\hat{K}(s) - \pi s^2$ is beyond the upper bound of the envelope, the observed data are considered to exhibit the clustering distribution; if $\hat{K}(s) - \pi s^2$ is behind the lower bound of the envelope, the observed data are considered to display a spatial pattern of regularity. $\hat{L}(s)$ is also employed to examine the clustering (Zhang, 1998).

$$\hat{L}(s) = \sqrt{\frac{\hat{K}(s)}{\pi}} - s \quad (2)$$

2.2 Space-time K-function and the method of spatio-temporal point pattern analysis

Obviously the hypothesis of CSR neglects the time factor. If the lapse time between two events is enough long, the time factor can affect the clustering of events. To analyze the effect, Diggle et al (1995) defined the space-time *K*-function:

$$K(s, t) = \lambda^{-1} E(n_{st}).$$

Here the intensity λ is defined as the expected number of events per unit space per unit time; n_{st} denotes number of further events occurring within distance s and time t of an arbitrary event. If the spatial component and the temporal component are independent in the spatio-temporal process on a large but finite region with the form $A \times (0, T)$, it then follows that:

$$K(s, t) = K_1(s)K_2(t), \tag{3}$$

where $K_1(\cdot)$ and $K_2(\cdot)$ denote the *K*-function of the spatial and the temporal component processes, respectively.

$$K_1(s) = \lambda_1^{-1} E(n_s)$$

$$K_2(t) = \lambda_2^{-1} E(n_t)$$

Here λ_1 and λ_2 denote the spatial and temporal intensities, respectively, where $\lambda_1 = \lambda T$ and $\lambda_2 = \lambda |A|$; n_s denotes number of further events occurring within distance s of an arbitrary event; n_t denotes number of further events occurring within time t of an arbitrary event.

The estimated formulae of three *K*-functions are

$$\begin{cases} \hat{K}(s, t) = \frac{|A|T}{n(n-1)} \sum_i \sum_{j \neq i} w_{ij} I(d_{ij} \leq s) v_{ij} I(u_{ij} \leq t) \\ \hat{K}(s) = \frac{|A|}{n(n-1)} \sum_i \sum_{j \neq i} w_{ij} I(d_{ij} \leq s) \\ \hat{K}(t) = \frac{T}{n(n-1)} \sum_i \sum_{j \neq i} v_{ij} I(u_{ij} \leq t) \end{cases}, \tag{4}$$

where w_{ij} is the spatial edge-corrected weight, which has been mentioned above; v_{ij} denotes the temporal edge-corrected weight. Let $v_{ij} = 1$ if both ends of the interval of length $2u_{ij}$ and centre t_{ij} lie within $(0, T)$, $v_{ij} = 2$ otherwise. According to Diggle et al (1995), the temporal edge effects are much less important in practice.

For diagnosing space-time clustering, Diggle et al (1995) defined the following two important functions:

$$\hat{D}(s, t) = \hat{K}(s, t) - \hat{K}_1(s)\hat{K}_2(t) \tag{5a}$$

and

$$\hat{D}_0(s, t) = \frac{\hat{D}(s, t)}{\hat{K}_1(s)\hat{K}_2(t)}. \tag{5b}$$

The perspective plot of the surface of $\hat{D}(s, t)$ can show the dependence between the spatial and temporal components. $\hat{D}_0(s, t)$ has a nice physical interpretation as an estimated risk factor owing to space-time interaction. See Diggle et al (1995) for further explanation.

The plot of $R(s, t) = \hat{D}(s, t) / \text{Var}\{\hat{D}(s, t)\}$ versus $\hat{K}_0(s, t) = \hat{K}_1(s)\hat{K}_2(t)$ can also be used to diagnose space-time clustering, which is analogous to a plot of standardized residuals against fitted values in regression modeling. Using $R(s, t)$, the positive or negative space-time interaction can be obtained:

$$U = \sum_s \sum_t R(s, t). \tag{6}$$

The significantly positive or negative values of U indicate the positive or negative space-time interaction. On the basis of equation (6), the Monte-Carlo test is done, where the observed value U_0 is ranked amongst values U_1, U_2, \dots, U_r generated by recomputing the $R(s, t)$ value after each of r independent random permutations of the reference times. If U_0 ranks r_0 th largest, the one-sided attained significance level is $r_0/(r+1)$. Then we can judge whether the null hypothesis of no space-time interaction holds.

3 Spatio-temporal point pattern analysis on Wenchuan aftershocks

3.1 Pretreatment for the data of Wenchuan mainshock and aftershocks

According to the latitudinal and longitudinal data of Wenchuan mainshock and aftershocks, the package Map and the package Mapproj in the statistical package R are used to convert these data into the Euclidean planar coordinates. The distance unit in the Euclidean plane is km, where the coordinate of the mainshock is (243.754 8, 423.678 2) (Figure 1). In the next context, the package Splanx in the statistical package R and MatLab 6.5 are employed to execute CSR analysis and spatio-temporal point pattern analysis.

3.2 Result of CSR analysis

Regardless of the occurrence time of aftershocks, $\hat{L}(s)$ of the observed data and the envelope can be obtained by the method of CSR analysis (Figure 2). Obviously, the aftershocks exhibit the clustering distribution

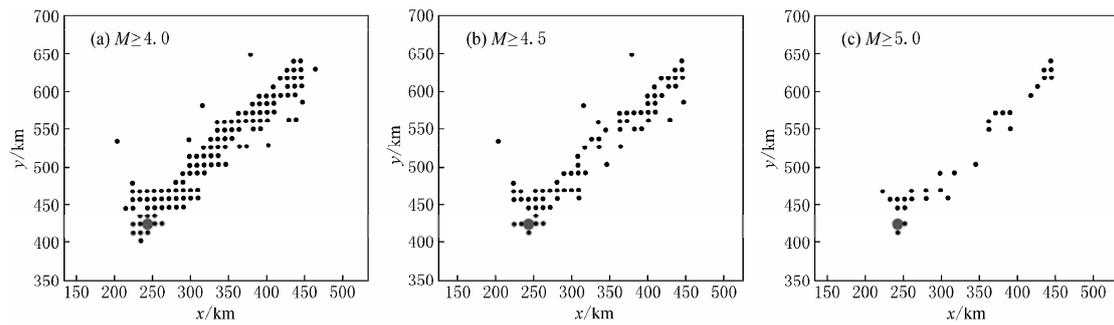


Figure 1 Distribution of Wenchuan mainshock (represented by gray dot) and aftershocks (presented by small dots) with $M \geq 4.0$ (a), $M \geq 4.5$ (b) and $M \geq 5.0$ (c) in the Euclidean plane. Horizontal coordinates are converted from latitudinal coordinates, and vertical coordinates are converted from longitudinal coordinates.

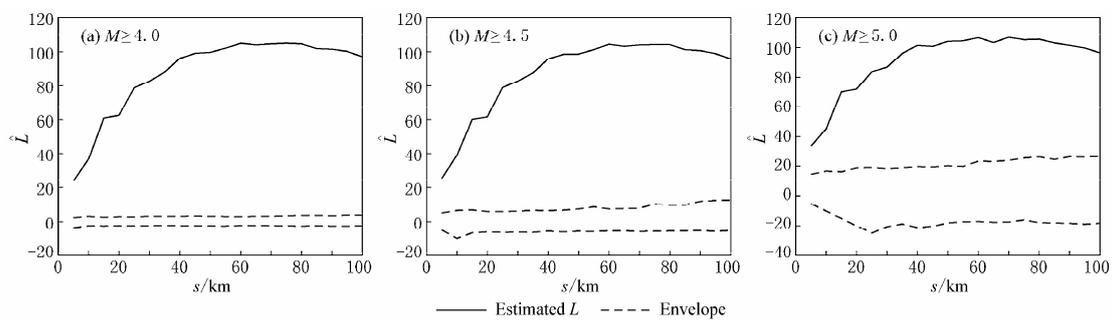


Figure 2 Estimated L of Wenchuan aftershocks of $M \geq 4.0$ (a), $M \geq 4.5$ (b) and $M \geq 5.0$ (c) with the envelope. A set of 9 999 Monte-Carlo simulations is performed. The distance scales s are chosen from 5 km to 100 km with step of 5 km.

because the $\hat{L}(s)$ curve of the observed data is beyond the upper bound of the envelope. The characteristic of spatial clustering gradually strengthens under 60 km. The widths between envelopes in Figures 2a–2c are different because the numbers of aftershocks with different magnitude are different.

3.3 Result of spatio-temporal point pattern analysis

If the occurrence time of aftershocks is considered, Wenchuan aftershocks form space-time events. Figures 3–5 show the results of spatio-temporal point pattern analysis. Figures 3a, 4a and 5a show the result according to equation (5a). $\hat{D}(s, t)$ grows with increasing s and t at first, then descends, which implies that the space-time clustering exists on a certain distance scale and on a certain time scale. Figures 3b, 4b and 5b show the results according to equation (5b). The values of $\hat{D}_0(s, t)$ are apparently large on a certain distance scale and on a certain time scale. From Figure 3b, we can see the Wenchuan mainshock and aftershocks exhibit space-time clustering on a certain distance scale and on a certain time scale, especially under 60 km and under 260 h. The result is in accord with the result calculated according to

equation (5a). Figures 3c, 4c and 5c show the results calculated by the formula $R(s, t) = \hat{D}(s, t) / \text{Var}\{\hat{D}(s, t)\}$. According to the suggestion of Diggle et al (1995), informally, if the spatial and temporal component processes are independent, approximately 95% of the $R(s, t)$ values are expected to lie within the limits ± 2 . Most $R(s, t)$ values of the observed data are higher than the positive 2. Figure 3d exhibits the result of Monte-Carlo simulations. U_0 calculated by the data of aftershocks with $M \geq 4.0$ ranks 17th among the descending sequence $\{U_1, U_2, \dots, U_{9999}\}$. In the package Splancs, the sequence calculated by the Monte-Carlo simulations is increasing. Thus, U_0 ranks the 9983rd among the sequence. The one-sided test attained significance level is $17/(9999+1) = 0.0017 < 0.05$, therefore the null hypothesis of no space-time interaction is refused. In Figure 4d, the one-sided test attained significance level is $0.0003 < 0.05$, illustrating that the aftershocks with $M \geq 4.5$ also exhibit the space-time interaction. In Figure 5d, the one-sided test attained significance level is $0.0002 < 0.05$, illustrating that the aftershocks with $M \geq 5.0$ exhibit the space-time interaction. From the first aftershock sequence to the third aftershock sequence, the numbers of

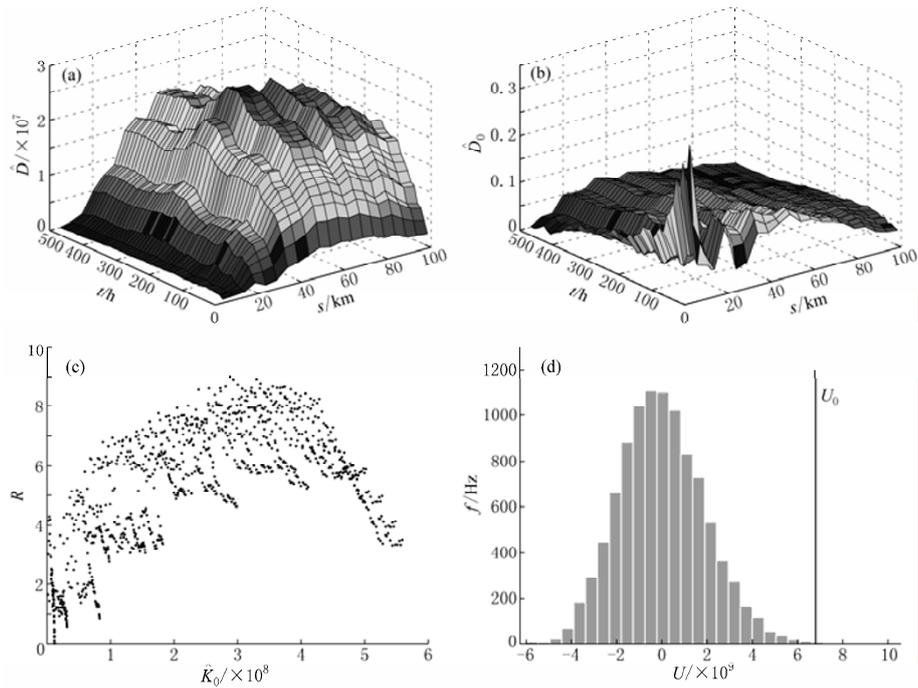


Figure 3 Testing of space-time clustering for the data of Wenchuan aftershocks ($M \geq 4.0$). (a) Estimated D plot; (b) Estimated D_0 plot; (c) Residual plot; (d) Monte-Carlo simulations. A set of 9999 simulations is performed. U_0 ranks the 17th. The distance scales are chosen from 5 km to 100 km with step 5 km; the time scales are chosen from 10 h to 500 h with step 10 h.

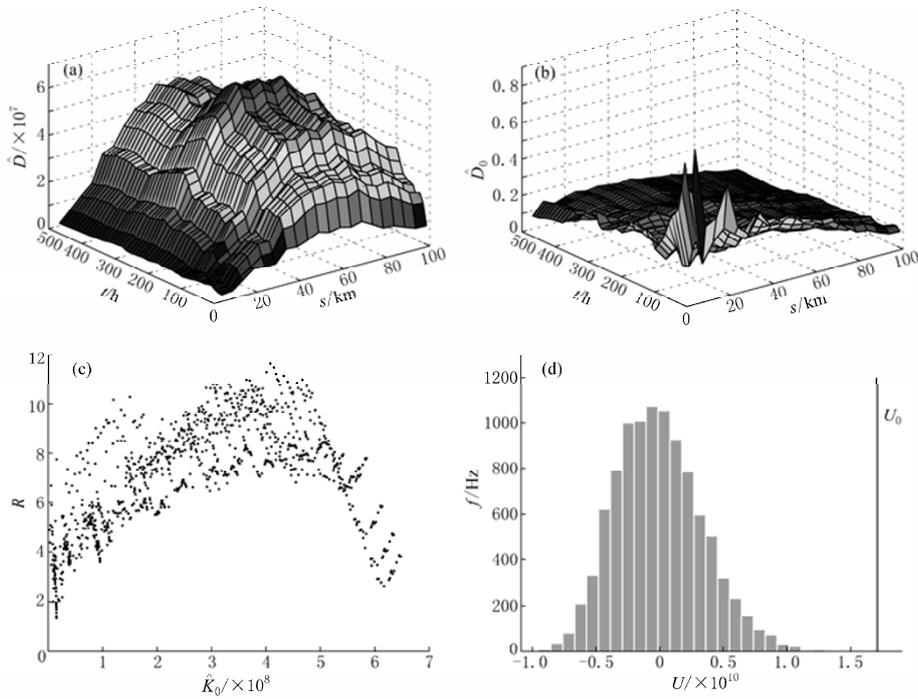


Figure 4 Testing of space-time clustering for the data of Wenchuan aftershocks ($M \geq 4.5$). (a) Estimated D plot; (b) Estimated D_0 plot; (c) Residual plot; (d) Monte-Carlo simulations. A set of 9999 simulations is performed. U_0 ranks the 3rd. The distance scales are chosen from 5 km to 100 km with step 5 km; the time scales are chosen from 10 h to 500 h with step 10 h.

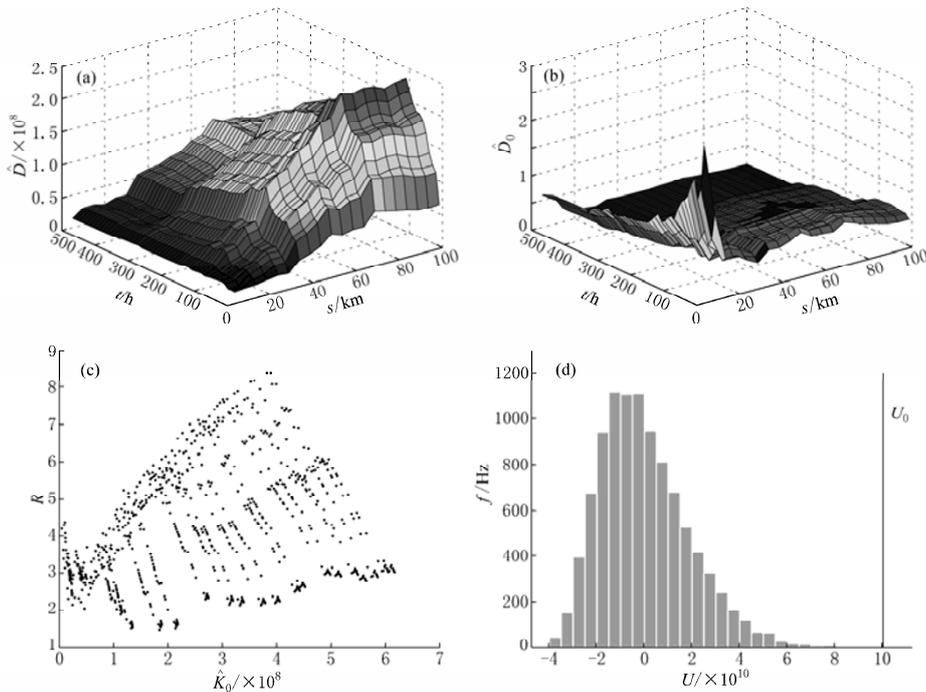


Figure 5 Testing of space-time clustering for the data of Wenchuan aftershocks ($M \geq 5.0$). (a) Estimated D plot; (b) Estimated D_0 plot; (c) Residual plot; (d) Monte-Carlo simulations. A set of 9 999 simulations is performed. U_0 ranks the 2nd. The distance scales are chosen from 5 km to 100 km with step 5 km; the time scales are chosen from 10 h to 500 h with step 10 h.

the data are 294, 127 and 43, respectively.

3.4 Results and discussion

The results show that under the given spatio-temporal scale the three aftershock sequences all obviously exhibit space-time interaction. And the clustering gradually strengthens under a small spatio-temporal scale. In CSR analysis, the clustering of the three aftershock sequences increases under 60 km, however in the range from 60 km to 100 km the clustering does not strengthen, but slightly descends.

Although the paper only analyzes the three different aftershock sequences, an important conclusion can be drawn regardless of the detailed values of the scales under which the spatio-temporal clustering strengthens: under a smaller spatial scale and under a larger temporal scale, the clustering of weaker aftershocks strengthens; but under a larger spatial scale and under a small temporal scale, the clustering of stronger aftershocks strengthens.

It is necessary to point out that, extending the scale of distance and the scale of time further may cause the error of analytical results for the limit of the number of aftershocks. Thus, we do not recommend a spatio-temporal point pattern analysis via extending the two scales further in the example of Wenchuan aftershocks.

The available literature on the spatio-temporal point pattern analysis does not define how to choose the scale of distance and the scale of time, therefore a trusty method is to choose the two scales according to the data in a more scientific context. It is apparent that a subjective decision of scholars about the two scales is unavoidable. In the paper, the maximal scale of distance is chosen to be 100 km and the maximal scale of time is chosen to be 500 h corresponding to the spatial range 350 km and the temporal range from the mainshock occurrence to the nearest aftershock occurrence 500 h. It is very similar to the rule that the predicted time is confined to the observed time in the prediction of the regression. Generally speaking, the maximal scale of distance in CSR analysis is defined as an half of the length of the shorter side in a rectangle (i.e., the studied region). However, the maximal scale of distance is usually chosen to be less than an half of the length of the shorter side in a rectangle (e.g., one third of the length of the shorter side) to reduce the estimated error further. In a word, in the paper, we choose the two scales according to a principle for reducing the estimated error.

The above analyses demonstrate the aftershock sequences with different magnitudes all exhibit obvious space-time clustering under the given spatial scale and

under the given temporal scale. Although it is possible for us to find the clustering via observing the three-dimensional picture (the occurrence time of aftershocks versus the latitude and longitude of aftershocks), quantifying the clustering is very difficult to handle. Therefore, it is valuable to employ the spatio-temporal point pattern analysis to study the spatio-temporal distribution of aftershocks.

There exists a certain spatio-temporal clustering in the aftershock activities, and different aftershock sequences may exhibit different characteristics of the spatio-temporal distribution. The paper just performs a preliminary study on Wenchuan aftershock sequences, thus further studies need to do to find whether aftershocks occurring in different areas have a common characteristic in the spatio-temporal distribution.

4 Conclusions

Regardless of the occurrence time of aftershocks, the characteristic of spatial clustering related to Wenchuan aftershocks gradually strengthens under 60 km by CSR analysis. Considering the occurrence time of aftershocks, the two Wenchuan aftershock sequences with magnitude $M \geq 4.0$ and $M \geq 4.5$ exhibit space-time clustering on a certain distance scale and on a certain time scale, especially under 60 km and under 260 h. For the aftershock sequence with $M \geq 5.0$, the space-time clustering also exists under the given scales, and it obviously strengthens under 80 km and under 150 h.

Acknowledgements This paper is supported by the Key Project of Chinese National Programs for Fundamental Research and Development (973 program) (2008CB425704). We thank Dr. Tom Minka (the Mit Media Laboratory, USA) and Prof. Peter Diggle (the

Medicine Department of Lancaster University, UK) for their extremely valuable help during the preparation of this paper, and we are also grateful to Prof. Bailian Li (the Ecological Complexity and Modeling Lab. of California University, Riverside, USA) for providing us with the necessary literatures. R, Map, Mapproj and SplanCs are all freely available from the R website (URL <http://www.cran.r-project.org>).

References

- Department of Monitoring and Prediction, China Earthquake Administration (2007). *The Research on Earthquake Sequences in Chinese Mainland*. Seismological Press, Beijing, 1–121 (in Chinese).
- Diggle P J (1983). *Statistical Analysis of Spatial Point Patterns*. Academic Press, London, 72.
- Diggle P J, Chetwynd A G, Häggkvist R and Morris S E (1995). Second-order analysis of space-time clustering. *Stat Methods Med Res* **4**: 124–136.
- Gatrell A C, Bailey T C, Diggle P J and Rowlingson B S (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Trans Inst Br Geogr NS* **21**: 256–274.
- Harris R A (1998). Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard. *J Geophys Res* **103**: 24 347–24 358.
- Ogata Y (1999). Seismicity analysis through point-process modeling: A review. *Pure Appl Geophys* **155**: 471–507.
- Ripley B D (1977). Modelling spatial patterns (with discussion). *J R Statist Soc B* **39**: 172–192.
- Robinson R and Zhou S (2005). Stress interactions within the Tangshan, China, earthquake sequence of 1976. *Bull Seism Soc Amer* **85**(6): 2 501–2 505.
- Rowlingson B S and Diggle P J (1993). SplanCs: Spatial point pattern analysis code in S-plus. *Comput Geosci* **51**: 627–655.
- Shi P J, Ge F, Yang Q P and Wang J G (2009). A new algorithm of the edge correction in the point pattern analysis and its application. *Acta Ecol Sin* **29**: 804–809 (in Chinese with English abstract).
- Stein R S (1999). The role of stress transfer in earthquake occurrence. *Nature* **402**: 605–609.
- Tang M P, Tang S Z, Lei X D, Zhang H R, Hong L X and Feng Y M (2003). Edge correction of Ripley's $K(d)$ function on population spatial pattern analysis. *Acta Ecol Sin* **23**: 1 533–1 538 (in Chinese with English abstract).
- Vere-Jones D (2009). Some models and procedures for space-time point processes. *Environ Ecol Stat* **16**: 173–195.
- Vere-Jones D, Ben-Zion Y and Zúñiga R (2005). Statistical seismology. *Pure Appl Geophys* **162**: 1 023–1 026.
- Zhang J T (1998). Analysis of spatial point pattern for plant species. *Acta Phytoecol Sin* **22**: 344–349 (in Chinese with English abstract).