Assessing Spatial Point Process Models for California Earthquakes Using Weighted K-functions

Alejandro Veen[1] and Frederic Paik Schoenberg[2]

[1] UCLA Department of Statistics
8125 Math Sciences Building
Box 951554
Los Angeles, CA 90095-1554
U.S.A.
veen@stat.ucla.edu

[2] UCLA Department of Statistics
8125 Math Sciences Building
Box 951554
Los Angeles, CA 90095-1554
U.S.A.
frederic@stat.ucla.edu

We propose a weighted analogue of Ripley's K-function for assessing the fit of point process models. The advantage of the proposed measure is that it can be used in situations where the null hypothesis is not a stationary Poisson model. We present its distributional properties for a spatial, two-dimensional inhomogeneous Poisson process and use it to assess the goodness-of-fit of two alternative point process models for the spatial distribution of California earthquakes.

1 Introduction

Ripley's K-function [Rip76], $K(h)$, is a widely used statistic to detect clustering or inhibition in point process data. It is commonly used as a test, where the null hypothesis is that the point process under consideration is a homogeneous Poisson process and the alternative is that the point process exhibits clustering or inhibitory behavior. Much research has been directed towards describing the asymptotic distribution of the K-function (see [Hei88], [Rip88] pp. 28–48, [Sil78]) for simple point process models including the homogeneous Poisson case.

The K-function has also been used in conjunction with point process residual analysis techniques in order to assess more general classes of point process models. For instance, a point process may be rescaled (see [MN86], [Oga88], [Sch99]) or thinned [Sch03] to generate residuals which are approximately homogeneous Poisson, provided the model used to generate the residuals is correct. The K-function can then be applied to the residual process in order to investigate the homogeneity of the residuals, and the result can be interpreted as a test of the goodness-of-fit of the point process model in question. Hence, residual analysis of a point process model involves two steps, the transformation of the data into residuals and a subsequent test for whether the residuals appear to be well approximated by a homogeneous Poisson process.

Of course, other methods for assessing the homogeneity of a point process exist, including tests for monotonicity [Saw75], uniformity (see [DRS84], [LL85], [Law88]), and tests on the second and higher-order properties of the process (see [Bar64], [Dav77], [Hei91]). Likelihood statistics, such as Akaike's Information Criterion (AIC, [Aka74]) and the Bayesian Information Criterion (BIC, [Sch79]) are often used to assess more general classes of models; see e.g. [Oga98] for an application to earthquake occurrence models.

We focus here on Ripley's K-function, and in particular on a modified version of the statistic that may be used directly to test a quite general class of null hypothesis models for the point process under consideration. The aim is to provide a direct test for goodness-of-fit for point process models, without having to assume homogeneity or to transform the points using residual analysis, the latter of which often introduces problems of highly irregular boundaries and large sampling variability when the conditional intensity in question is highly variable (see [Sch03]).

In Sect. 2, we introduce the proposed weighted version of the K-function. The statistic is then used in Sect. 3 to assess goodness-of-fit when applied to competing models for the spatial background rate for California earthquakes. Some concluding remarks are given in Sect. 4.

2 A Weighted K-function

In this Section, we propose a weighted analogue of Ripley's K-function which is similar to the K-function applied to the thinned residuals described in [Sch03]. The proposed estimator has the advantage of eliminating the sampling variability of the thinning procedure, and does not require repetition of the random thinning, but instead may be calculated directly. We begin with a review of Ripley's K-function.

2.1 Ripley's K-function and Variants

Consider a Poisson process of intensity λ on an interval $\mathcal{A}$ of the plane $\mathbf{R}^2$ with finite area A, and let the N points of the process be labelled $\{p_1, p_2, \ldots, p_N\}$.

Ripley's K-function $K(h)$ is typically defined as the average number of other points within h of any given point divided by the overall rate λ, and is most simply estimated via

$$\hat{K}(h) = \hat{\lambda}^{-1} N^{-1} \sum_r \sum_{s \neq s} I(|p_r - p_s| \leq h),$$

where $\hat{\lambda} = N/A$ is an estimate of the overall intensity, $I(\cdot)$ is the indicator function and h is some inter-point distance of interest. In applications, $\hat{K}$ is typically calculated for several different choices of h. For a homogeneous Poisson process, the expectation of $\hat{K}(h)$ is πh^2. Values which are higher than this expectation indicate clustering, while lower values indicate inhibition. However, it should be noted that a point pattern can be clustered at some scale, while it may show inhibition at a different scale. Also, a non-Poisson process can have the same K-function, as $K(h)$ only takes the first two moments into account. An example of such a process can be found in [BS84].

Under the null hypothesis that the point process is homogeneous Poisson, $\hat{K}(h)$ is asymptotically normal:

$$\hat{K}(h) \approx N\left(\pi h^2, \frac{2\pi h^2}{\lambda^2 A}\right),$$

as the area of observation A tends to infinity (see [Cre93] p.642, [Rip88] pp. 28–48, [Hei88], [Sil78]).

Several variations on $\hat{K}(h)$ have been proposed. Many deal with corrections for boundary effects, as found in [Rip76], [OS81], and [Ohs83]. Variance-stabilizing transformations of estimated K-functions which are more easily interpretable have been proposed (see [Bes77]), such as $\hat{L}(h)$ and $\hat{L}(h) - h$:

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}}\,. \tag{1}$$

Confidence bounds for $\hat{L}(h)$ can be derived by transforming the confidence bounds of $\hat{K}(h)$.

2.2 A Weighted Analogue of Ripley's K-function

Suppose that a given planar point process on an interval $\mathcal{A}$ of $\mathbf{R}^2$ of area A may be specified by its *conditional* intensity with respect to some filtration on $\mathcal{A}$, for $(x, y) \in \mathcal{A}$ (see [DV03]). The point process need not be Poisson; in the simple case where the point process is Poisson, however, the conditional intensity and ordinary intensity coincide. Suppose that, under the null hypothesis (H_0), the conditional intensity of the point process is given by $\lambda_0(x, y)$.

We define the *weighted K-function*, used to assess the model $\lambda_0(x, y)$, as

$$K_W(h) = \frac{1}{\lambda_* \hat{E}_{H_0}(N)} \sum_r w_r \sum_{s \neq r} w_s I(|p_r - p_s| \leq h) \tag{2}$$

where $\lambda_* := \inf\{\lambda_0(x,y); (x,y) \in \mathcal{A}\}$ is the infimum of the conditional intensity over the observed region under the null hypothesis, $\hat{E}_{H_0}(N) = \int\int_{(x,y)\in\mathcal{A}} \lambda_0(x,y)\mathrm{d}x\mathrm{d}y$ is an estimate of the expected number of points in $\mathcal{A}$ under the null hypothesis, and $w_r = \lambda_*/\lambda_0(p_r)$, where $\lambda_0(p_r)$ is the conditional intensity at point p_r under H_0.

One can think of the weighted K-function as a combination of Ripley's K-function and the thinning method used for residual analysis in [Sch03]. In [Sch03], $K(h)$ is repeatedly applied to thinned data where the probability of retaining a point is inversely proportional to the conditional intensity at that point. The computation of the weighted K-function K_W uses these retaining probabilities as weights for the points in order to offset the inhomogeneity of the process. By incorporating all pairs of the observed points, rather than only the ones that happen to be retained after an iteration of random thinning, the statistic $K_W(h)$ eliminates the sampling variability due to thinning the process repeatedly.

We conjecture that, provided the conditional intensity λ_0 is sufficiently smooth, $K_W(h)$ will be approximately normal as the area of observation A approaches infinity. Indeed, for the Poisson case where λ_0 is locally approximately constant on blocks of large area relative to the interpoint distance h, we have the following result.

Theorem 1. *Suppose that the observed regions $\mathcal{A}^{(n)}$ of areas $A^{(n)}$ increase in area to infinity such that $\mathcal{A}^{(n)}$ may be broken up into disjoint blocks $\mathcal{A}_1^{(n)}, \mathcal{A}_2^{(n)}, \ldots, \mathcal{A}_{I_n}^{(n)}$ of areas $A_1^{(n)}, \ldots, A_{I_n}^{(n)}$, respectively, where $I_n \to \infty$ as $n \to \infty$ and within each subset $\mathcal{A}_i^{(n)}$, the conditional intensity λ_0 is approximately constant, i.e. $\max_{i=1,\ldots,I_n} \{\sup\{\lambda_0(x,y); (x,y) \in \mathcal{A}_i^{(n)}\} - \inf\{\lambda_0(x,y); (x,y) \in \mathcal{A}_i^{(n)}\}\} \to 0$ as $n \to \infty$. Suppose also that h is small compared to the area of each block, i.e. $\sup_i \pi h^2/A_i^{(n)} \to 0$. Further, assume that the boundaries of the sets $\mathcal{A}_i^{(n)}$ are sufficiently small that, of all pairs of points (p_r, p_s) within distance h, the proportion where p_r and p_s are in different subsets $\mathcal{A}_i^{(n)}$ converges to zero as $n \to \infty$. Let $K_W^{(n)}(h)$ and $\hat{E}_{H_0}(N)^{(n)}$ denote $K_W(h)$ and $\hat{E}_{H_0}(N)$, respectively, calculated on the region $\mathcal{A}^{(n)}$. Then $K_W^{(n)}(h)$ is asymptotically normal as $n \to \infty$:*

$$K_W^{(n)}(h) \stackrel{\sim}{\sim} N\left(\pi h^2, \frac{2\pi h^2 A^{(n)}}{\left(\hat{E}_{H_0}(N)^{(n)}\right)^2}\right).$$

Proof. Following [BP04] we will take advantage of the fact that the inter-point distances $d_{rs} = |p_r - p_s|$ can be treated as independent random variables, as the number of points on a given domain approaches infinity. Furthermore, it is evident that the probability of a randomly chosen point p_s being within h

of an arbitrary point p_r is $\pi h^2/A_i^{(n)}$, provided that $p_r, p_s \in \mathcal{A}_k^{(n)}$. It follows therefore that at stage n, the number $P_i^{(n)}(h)$ of pairs of points in subset $\mathcal{A}_i^{(n)}$ with an inter-point distance not larger than h has an approximate Binomial distribution:

$$P_i^{(n)}(h) \sim B\left(\frac{1}{2}\left(\lambda_i^{(n)} A_i^{(n)}\right)^2, \frac{\pi h^2}{A_i^{(n)}}\right),$$

where $\lambda_i^{(n)}$ is the (approximately constant) conditional intensity within $\mathcal{A}_i^{(n)}$, since $(1/2)(\lambda_i^{(n)} A_i^{(n)})^2$ represents the expected number of pairs in $A_i^{(n)}$ and $\pi h^2/A_i^{(n)}$ is the probability that the inter-point distance of a given pair is not greater than h. Hence, the expectation is $E(P_i^{(n)}(h)) = (1/2)\pi h^2 (\lambda_i^{(n)})^2 A_i^{(n)}$ and noting that $\pi h^2/A_i^{(n)} \approx 0$, the variance is approximately the same: $V(P_i^{(n)}(h)) \approx E(P_i^{(n)}(h)$.

Because $P_i^{(n)}(h), i = 1, \ldots, I_n$ are approximately independent by assumption, it follows that

$$E\left(\sum_{i=1}^{I} P_i^{(n)}(h)/\lambda_i^{(n)}\right) = \frac{1}{2}\pi h^2 \sum_{i=1}^{I_n} \lambda_i^{(n)} A_i^{(n)}$$
$$Var\left(\sum_{i=1}^{I} P_i(h)/\lambda_i^{(n)}\right) = \frac{1}{2}\pi h^2 \sum_{i=1}^{I_n} A_i^{(n)} .$$

Noting that $K_W^{(n)}(h)$ can be written as

$$K_W^{(n)}(h) = \frac{2\sum_{i=1}^{I_n} \frac{\lambda_*^{(n)}}{\lambda_i^{(n)}} P_i^{(n)}(h)}{\lambda_*^{(n)} \hat{E}_{H_0}(N)^{(n)}}$$

it follows that the expectation and variance of $K_W^{(n)}(h)$ are given by

$$E\left(K_W^{(n)}(h)\right) = \pi h^2$$
$$Var\left(K_W^{(n)}(h)\right) = \frac{2\pi h^2 A^{(n)}}{\left(\hat{E}_{H_0}(N)^{(n)}\right)^2},$$

since $\sum_{i=1}^{I_n} \lambda_i^{(n)} A_i^{(n)} = \hat{E}_{H_0}(N)^{(n)}$.

Finally, it follows directly from the Central Limit Theorem (since $I_n \to \infty$) that $K_W^{(n)}(h)$ has an asymptotic normal distribution

$$K_W(h) \sim N\left(\pi h^2, \frac{2\pi h^2 A^{(n)}}{\left(\hat{E}_{H_0}(N)^{(n)}\right)^2}\right).$$

□

A variance-stabilized analogue of $L(h)$ in (1), i.e. a variance-stabilized version of the weighted K-function, could be defined as:

$$L_W(h) = \sqrt{\frac{K_W(h)}{\pi}}.$$

3 Application

The test statistic $K_W(h)$ in (2) is applicable to a very general class of planar point process models. We investigate their application to models for the spatial background rate for the occurrences of Southern California earthquakes.

3.1 Data Set

Data on Southern California earthquakes are compiled by the Southern California Earthquake Center (SCEC). The data include the occurrence times, magnitudes, locations, and often even waveforms and moment tensor solutions, based on recordings at an array of hundreds of seismographic stations located throughout Southern California, including over 50 stations in Los Angeles County alone. The catalog is maintained by the Southern California Seismic Network (SCSN), a cooperative project of the California Institute of Technology and the United States Geological Survey. The data are available to the public; information is provided at http://www.data.scec.org.

We focus here on the spatial locations of a subset of the SCEC data occurring between 01/01/1984 and 06/17/2004 in a rectangular area around Los Angeles, California, between longitudes $-122°$ and $-114°$ and latitudes $32°$ and $37°$ (approximately $733\,km \times 556\,km$). The data set consists of earthquakes with magnitude not smaller than 3.0, of which 6,796 occurred within the given 21.5-year period. The epicentral locations of these earthquakes are shown in Fig. 1.

3.2 Analysis

Spatial background rates are commonly estimated by seismologists by smoothing the larger events only. For instance [Oga98] suggests anisotropic kernel smoothing of large mainshocks in order to estimate the spatial background intensity for all earthquakes. In this application, we investigate various spatial background seismicity rate estimates involving kernel smoothings of only the 2030 earthquakes of magnitude 3.5 and higher, by using $K_W(h)$ to assess their fit to the earthquake data set. The local seismicity at location (x, y) may be estimated using a bivariate kernel smoothing $\mu(x, y)$ of the events of magnitude at least 3.5. Figure 2 shows such a kernel smoothing, using an anisotropic

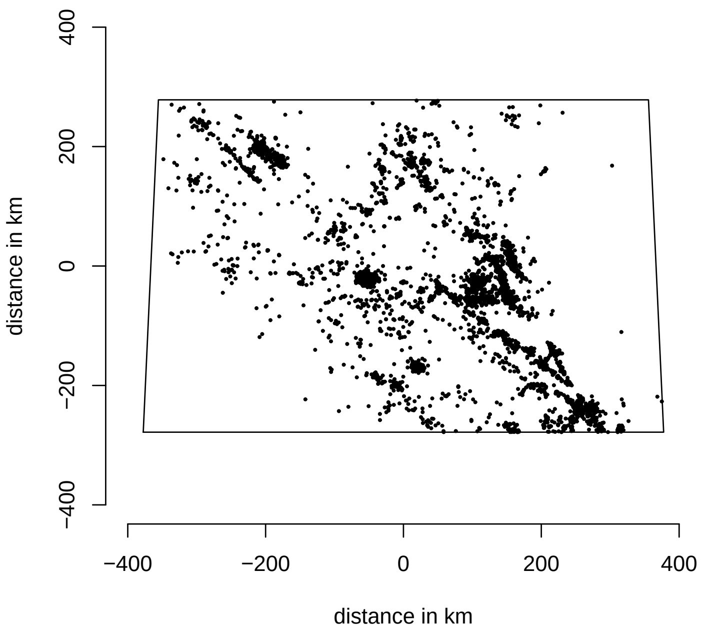

Fig. 1. Earthquakes in Southern California 1984-2004: The data set consists of 6796 earthquakes with magnitude 3.0 or larger

bivariate normal kernel with a bandwidth of 8 km and a correlation of −0.611. That is,

$$\mu(x, y) = \sum_{r=1}^{N} f(x - x_r, y - y_r), \tag{3}$$

where the sum is over all points (x_r, y_r) with magnitude $m_r \geq 3.5$, and f is the bivariate normal density centered at the origin with standard deviation $\sigma_x = \sigma_y = 8\,\mathrm{km}$ and correlation $\rho = -0.611$. This correlation is estimated using the empirical correlation of the values of x_r and y_r, and the bandwidth is selected by inspection. The agreement of Figs. 1 and 2 does not seem grossly unreasonable.

Since such a kernel smoothing only uses the observed seismicity over the last 20 years (a relatively small time period by geological standards), one may wish to allow for the possibility of seismicity in regions where no earthquakes of magnitude 3.5 or higher have recently been observed. One way to do this is by estimating the spatial background intensity via a weighted average of

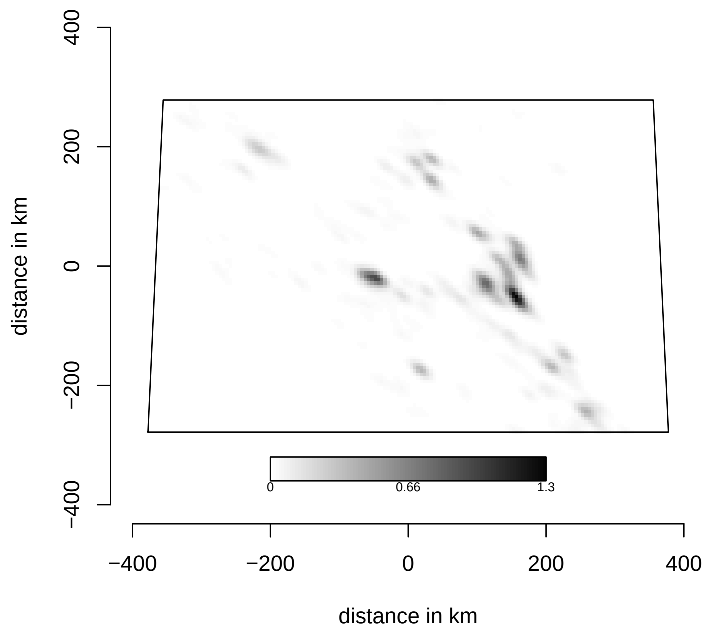

Fig. 2. Kernel smoothing of seismicity in Southern California 1984-2004: An anisotropic bivariate normal kernel with a bandwidth of 8 km ($\rho = -0.611$, $\sigma_x = \sigma_y = 8$ km) is applied to 2030 earthquakes with magnitude not smaller than 3.5

the kernel-smoothed seismicity of magnitude at least 3.5 and a positive constant representing an estimate of the spatial background intensity under the assumption that the process is homogeneous Poisson. Hence we consider the estimate of the form

$$\hat{\lambda}_a(x, y) = a\mu(x, y) + (1 - a)\nu, \tag{4}$$

where $\nu = N/A$ is the estimated conditional intensity for a homogeneous Poisson model and a is some constant with $0 \leq a \leq 1$. Figure 3 shows $K_W(h)$ (2) applied to several spatial intensity estimates, each of the form (4), using different values for the parameter a. For the competing estimates $\hat{\lambda}_a$, a takes on the values 0.95, 0.98, 0.99, 0.9925, 0.995, 0.9975, 0.999, and 0.9999. $K_W(h)$ for the competing models $\hat{\lambda}_a$ can be seen in Fig. 3 where a darker line color indicates a higher value of a. The lower values of a give more weight to the homogeneous background rate than higher values of a. For a value of $a = 0.999$ or higher, $\hat{\lambda}_a$ is virtually identical to the kernel density estimate μ. The dashed

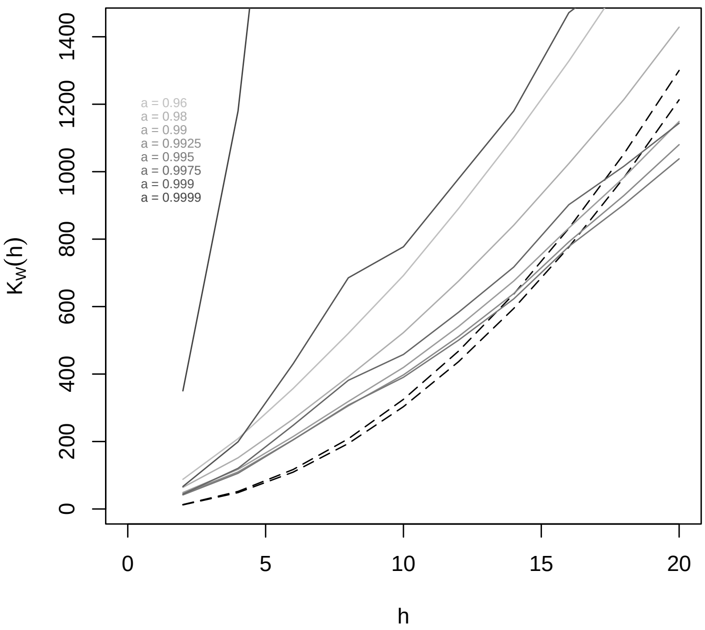

Fig. 3. Weighted K-function for competing models: The weighted K-function $K_W(h)$ is shown for different values of a in the background intensity model $\hat{\lambda}_a$ (4)

curves denote the 95% bounds for $K_W(h)$ based on the result in Theorem 1. Note that the smoothness condition in Theorem 1 is only appropriate if most pairs of points which are within distance h have similar estimated intensities, which is more or less the case for small values of h in this application.

Figure 3 shows that values of $a = 0.999$ or greater fit very poorly to the data. For such values of a, the intensity estimate gives weight almost exclusively to the kernel smoothing, so that pairs of small earthquakes in areas where there were no earthquakes of magnitude greater than or equal to 3.5 have extremely small probability and are hence given enormous weight in the computation of K_W. Similarly, for values of a below $a = 0.99$, the intensity estimate gives too much weight to the homogeneous Poisson component and too little to the kernel smoothing of the large events, so that the resulting model underpredicts the intense clustering in the data occurring around the larger events.

As shown in Fig. 3, for most small values of h the function $K_W(h)$ seems to decrease towards the 95% bounds indicating satisfactory fit for values of a approaching $a = 0.995$ from either direction. This value of a appears to offer better fit than other values of a (and certainly is far better than the conventional $a = 1.0$). However, even for $a = 0.995$, the values of $K_W(h)$ nevertheless far exceed the 95% bounds. Apparently the data set contains significant clustering of the smaller events in locations not covered by the larger events. No mixture of a kernel smoothing of the larger events and a homogeneous Poisson estimate can possibly adequately account for such clustering.

4 Concluding Remarks

The application of K_W to spatial background rate estimates for Southern California seismicity suggests that a superior fit is provided by an estimate that incorporates both a kernel smoothing of the larger events as well as a homogeneous background rate. The function K_W appears to be a quite reasonable goodness-of-fit test for spatial point process models.

In contrast to standard kernel smoothing of the larger events in the catalog, the method of spatial background rate estimation which mixes the kernel estimate with a homogeneous constant rate appears to offer somewhat superior fit to the SCEC dataset. This suggests that spatial background rate estimates in commonly used models for seismic hazard, such as the epidemic-type aftershock sequence (ETAS) model of [Oga98], might possibly be improved in this way as well. Seismologically, the results are consistent with the notion that Southern California earthquakes, though certainly far more likely to occur on known faults, can potentially occur on unknown faults as well, and these faults may be quite uniformly dispersed. The results suggest that a spatial background rate estimate incorporating both of these possibilities could provide improved fit to existing models for seismic hazard. Such a modification may be especially relevant given the occurrences in California of blind (i.e. previously unknown) faults such as the one which ruptured during the Northridge earthquake in 1994, causing at least 33 deaths and 138 injuries as well as extensive public and private property damage [Pee98].

Further study is needed in order to confirm the seismological results suggested herein, for several reasons. First, it remains to be seen whether the features observed here may be reproduced elsewhere or are particular to Southern California. Second, in Theorem 1 and the conjecture preceding it, the observation area is thought to expand to infinity, and the smoothness of the conditional intensity is required. In our application to Southern California earthquakes, it is a bit unclear whether the observed region is sufficiently large and whether the intensity estimates of the form (4) are sufficiently smooth relative to the inter-point distance h to justify applying the results of Theorem 1 with great confidence. Third, in the estimation of the intensities of the form (4), the bandwidth, correlation, and the choice of kernel were not optimally selected, but chosen rather arbitrarily. Another issue worth mentioning is that the earthquakes of magnitude greater than 3.5 were used both in the fitting and in the testing. This is in keeping with common practice in seismology, though in statistical terms this is certainly non-standard. In addition, the problem of boundary effects in the estimation of the weighted K-function has not been addressed in this paper. Instead, we attempted to give a simplified presentation in introducing $K_W(h)$ and its application. It should be noted, however, that exactly the same standard boundary-correction techniques which are used for the ordinary K-function (see Sect. 2.1) can be used for the weighted K-function as well. Fortunately, in our application the fraction of points within distance h of the boundary was so small for all values of h considered as to make such considerations rather negligible.

5 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0306526. We thank Dave Jackson, Yan Kagan for helpful comments, and the Southern California Earthquake Center for its generosity in sharing their data.

References

[Aka74] Akaike, H.: A new look at statistical model identification. IEEE Transactions on Automatic Control, **AU–19**, 716–722 (1974)

[BS84] Baddeley, A.J., Silverman, B.W.: A cautionary example on the use of second-order methods for analyzing point patterns. Biometrics, **40**, 1089–1093 (1984)

[Bar64] Bartlett, M.: The spectral analysis of two-dimensional point processes. Biometrika, **51**, 299–311 (1964)

[Bes77] Besag, J.E.: Comment on "Modelling spatial patterns" by B.D. Ripley. Journal of the Royal Statistical Society, Series B, **39**, 193–195 (1977)

[BP04] Bonetti, M., Pagano, M.: The interpoint distance distribution as a descriptor of point patterns: An application to cluster detection, (submitted, Statistics in Medicine) (1977)

[Cre93] Cressie, N.A.C.: Statistics for spatial data, revised edition. Wiley, New York (1993)

[DV03] Daley, D., Vere-Jones, D.: An Introduction to the Theory of Point Processes, 2nd edition. Springer, New York (2003)

[Dav77] Davies, R.: Testing the hypothesis that a point process is Poisson. Adv. Appl. Probab., **9**, 724–746 (1977)

[DRS84] Dijkstra, J., Rietjens, T., Steutel, F.: A simple test for uniformity. Statistica Neerlandica **38**, 33–44 (1984)

[Hei88] Heinrich, L.: Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes. Statistics, **19**, 87–106 (1988)

[Hei91] Heinrich, L.: Goodness-of-fit tests for the second moment function of a stationary multidimensional Poisson process. Statistics, **22**, 245–278 (1991)

[Law88] Lawson, A.: On tests for spatial trend in a non-homogeneous Poisson process. J. Appl. Stat., **15**, 225–234 (1988)

[LL85] Lisek, B., Lisek, M.: A new method for testing whether a point process is Poisson. Statistics, **16**, 445–450 (1985)

[MN86] Merzbach, E., Nualart. D.: A characterization of the spatial Poisson process and changing time. Ann. Probab. **14**, 1380–1390 (1986)

[OS81] Ohser J., Stoyan D.: On the second-order and orientation analysis of planar stationary point processes. Biom. J. **23**, 523–533 (1981)

[Ohs83] Ohser, J.: On estimators for the reduced second-moment measure of point processes. Math. Operationsf. Statist., Ser. Statistics **14**, 63–71 (1983)

[Oga88] Ogata, Y.: Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. Journal of the American Statistical Association, **83**, 9–27 (1988)

[Oga98] Ogata, Y.: Space-time point-process models for earthquake occurrences. Ann. Inst. Statist. Math., **50**, 379–402 (1998)

[Pee98] Peek-Asa, C., Kraus, J.F., Bourque, L.B., Vimalachandra, D., Yu, J., Abrams, J.: Fatal and hospitalized injuries resulting from the 1994 Northridge earthquake. Int. J. Epidemiology, **27** (3), 459–465 (1998)

[Rip76] Ripley, B.D: The second-order analysis of stationary point processes. J. Appl. Probab., **13**, 255–266 (1976)

[Rip79] Ripley, B.D.: Tests of 'Randomness' for Spatial Point Patterns. Journal of the Royal Statistical Society, Series B, **41**, 368–374 (1979)

[Rip88] Ripley, B.D: Statistical Inference for Spatial Processes. Cambridge University Press, Cambridge (1988)

[Saw75] Saw, J.G.: Tests on intensity of a Poisson process. Communications in Statistics, **4** (8), 777–782 (1975)

[Sch99] Schoenberg, F.: Transforming Spatial Point Processes into Poisson Processes. Stochastic Processes and their Applications, **81**, 155–164 (1999)

[Sch03] Schoenberg, F.P.: Multi-dimensional residual analysis of point process models for earthquake occurrences. Journal of the American Statistical Association, **98**(464), 789–795 (2003)

[Sch79] Schwartz, G.: Estimating the dimension of a model. Annals of Statistics, **6**, 461–464 (1979)

[Sil78] Silverman, B.W.: Distances on circles, toruses and spheres. J. Appl. Probab., **15**, 136–143 (1978)