

# Goodness-of-fit Assessment of Point Process Models Based on $K$ -function Variants

Alejandro Veen      Frederic Paik Schoenberg

April 18, 2006

## Abstract

This work presents an alternative derivation of the asymptotic distribution of Ripley's  $K$ -function for a homogeneous Poisson process and shows how it can be combined with point process residual analysis in order to test for different classes of point process models. This can be done with the mean  $K$ -function of thinned residuals ( $K_M$ ) or a weighted analogue called the weighted or inhomogeneous  $K$ -function ( $K_W$ ). The asymptotic distributions of  $K_M$  and  $K_W$  are derived for inhomogeneous Poisson processes. Both statistics can be used as measures of goodness-of-fit for point process models.

**Keywords:** spatial point processes,  $K$ -function, thinning, model assessment, goodness-of-fit

## 1 Introduction

Ripley's  $K$ -function,  $K(h)$ , (Ripley 1976) 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. Previous authors have described the asymptotic distribution of the  $K$ -function for simple point process models including the homogeneous Poisson case (Silverman 1978, Heinrich 1988, Ripley 1988, pp. 28–48).

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 (Schoenberg 2003). Such model assessment involves two steps. In the first step, the observed point pattern is transformed into a residual process using a null hypothesis model. If the null hypothesis model is correct, the resulting transformed process is homogeneous Poisson. In the second step, the  $K$ -function is applied to the transformed process in order to test whether the residuals appear to be well approximated by a homogeneous Poisson process.

A different approach to model assessment is to modify the  $K$ -function and define a new statistic which incorporates the null hypothesis model. This approach was taken by Baddeley, Møller, and Waagepetersen (2000) as well as Veen and Schoenberg (2005) who define the inhomogeneous or weighted  $K$ -function. This weighted analog of the  $K$ -function provides a direct test for the fit of point process models which does not require the aforementioned two step methodology.

This Chapter will present Ripley's  $K$ -function and derive some of its distributional properties in Section 2 while Section 3 will describe how the goodness-of-fit of point process models can be assessed using the  $K$ -function and the weighted  $K$ -function.

## 2 Ripley's $K$ -function and an Alternative Derivation of its Distribution

### 2.1 Ripley's $K$ -function

Consider a homogeneous Poisson process  $\mathbf{N}$  of intensity  $\lambda$  on a connected subset  $\mathcal{A}$  of the plane  $\mathbf{R}^2$  with finite area  $A$ , and let the  $N$  points of the process be labeled  $\{p_1, p_2, \dots, p_N\}$ . The number of points  $N$  is therefore a Poisson random variable with expectation and variance  $\lambda A$ .

Ripley's  $K$ -function  $K(h)$  is defined as the expected number of further points within distance  $h$  of any given point divided by the overall rate  $\lambda$ . It is most simply estimated via

$$\hat{K}(h) = \frac{1}{\hat{\lambda}N} \sum_r \sum_{s \neq r} \mathbf{1}(|p_r - p_s| \leq h), \quad (1)$$

where  $\hat{\lambda} = N/A$  is an estimate of the overall intensity,  $\mathbf{1}(\cdot)$  is the indicator function, the Euclidian distance is denoted as  $|\cdot|$ , and  $h$  is some inter-point distance of interest. Note that  $\hat{K}(h)$  is not defined for  $N = 0$ . Ripley's  $K$ -function can be written equivalently as

$$\hat{K}(h) = \frac{A}{N^2} \sum_r \sum_{s \neq r} \mathbf{1}(|p_r - p_s| \leq h). \quad (2)$$

In applications, estimates of  $K(h)$  are typically calculated for several different choices of  $h$  and compared to the true value of  $K(h)$  for a homogeneous Poisson process, which is  $\pi h^2$  if effects on the boundary of  $\mathcal{A}$  are ignored or corrected. Values which are higher than this expectation indicate *clustering* of points, while lower values indicate *inhibition* (sometimes referred to as regularity or negative clustering). However, it should be noted that a point pattern can be clustered at certain scales and inhibitory at others (Schoenberg and Bolt (2000), for instance, find such features in seismological point process data). Note also that two distinct point processes may have identical  $K$ -functions, as  $K(h)$  only takes the first two moments into account. An example of such a situation can be found in Baddeley and Silverman (1984) who present a point process which has the same  $K$ -function as a homogeneous Poisson process, yet is a different process. For a homogeneous Poisson process with rate  $\lambda$ ,  $\hat{K}(h)$  is asymptotically Normal:

$$\hat{K}(h) \approx N \left( \pi h^2, \frac{2\pi h^2}{\lambda^2 A} \right), \quad (3)$$

as the area of observation  $A$  tends to infinity (see Ripley (1988) pp. 28–48 and Cressie (1993) p. 642).

Several variations on  $\hat{K}(h)$  have been proposed. Many deal with corrections for boundary effects, as found in Ripley (1976, 1988), Ohser and Stoyan (1981), and Ohser (1983). Variance-stabilizing transformations of estimated  $K$ -functions have been proposed (Besag 1977), such as  $\hat{L}(h)$  and  $\hat{L}(h) - h$  where

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}}. \quad (4)$$

The expectation of this quantity for a homogeneous Poisson process is  $h$ , making it more easily interpretable in plots of  $\hat{L}(h)$  or  $\hat{L}(h) - h$  versus  $h$ . Confidence bounds can be derived by transforming the confidence bounds of  $\hat{K}(h)$  accordingly.

## 2.2 Expectation and Variance of the Modified $K$ -function $\tilde{K}(h)$

As pointed out by Stoyan and Stoyan (2000), the  $K$ -function can also be estimated using an estimator for the squared intensity  $\tilde{\lambda}^2 = N(N-1)/A^2$ . This leads to the following expression:

$$\tilde{K}(h) = \frac{1}{\tilde{\lambda}^2 A} \sum_r \sum_{s \neq r} \mathbf{1}(|p_r - p_s| \leq h), \quad (5)$$

which is defined only for  $N \geq 2$ .  $\tilde{K}(h)$  can be expressed equivalently as

$$\tilde{K}(h) = \frac{2A}{N(N-1)} \sum_m \mathbf{1}_m(h), \quad (6)$$

where  $m \in \{1, 2, \dots, M\}$  is an index for the *pairs* of points of which there are  $M = \binom{N}{2} = \frac{1}{2}N(N-1)$ . The expression  $\mathbf{1}_m(h)$  denotes the indicator function (i.e. Bernoulli random variable) that equals 1 if the Euclidian distance between the two points for pair  $m$  is not larger than  $h$  and zero otherwise. This representation suggests an interpretation in terms of pairs rather than in terms of point counts within distance  $h$ . In this view,  $K(h)$  is proportional to the fraction of pairs which are within distance  $h$ , normalized by the area  $A$ .

Assuming that the inter-point distance  $h$  is sufficiently small in relation to the area of observation  $A$  and further assuming that the boundary of  $\mathcal{A}$  is sufficiently regular, edge effects can be disregarded. In the following, approximations regarding such boundary effects will be denoted as  $\simeq$ . Restricting our attention to observed point patterns which have at least one pair of points ( $N > 1$ ), it can be shown that  $\tilde{K}(h)$  is less biased for an homogeneous Poisson process than the the traditional estimator  $\hat{K}(h)$ . This follows from

$$\begin{aligned} E\left(\tilde{K}(h) \mid N > 1\right) &\simeq E\left(\frac{2A}{N(N-1)} \cdot \frac{1}{2}N(N-1) \cdot \frac{\pi h^2}{A} \mid N > 1\right) \\ &= \pi h^2, \end{aligned}$$

since there are  $\frac{1}{2}N(N-1)$  pairs of points and the expectation of  $\mathbf{1}_m(h)$  is the ratio of the area of a disc with radius  $h$  and the area of observation  $A$ . It follows that  $\hat{K}(h)$ , on the other hand, slightly underestimates the  $K$ -function since the denominator in (2) is  $N^2$  instead of  $N(N-1)$ . However, the asymptotic behavior of  $\hat{K}(h)$  and  $\tilde{K}(h)$  is the same as  $N$  gets large.

As the inter-point distances of two distinct pairs of points are uncorrelated, so are  $\mathbf{1}_m(h)$  and  $\mathbf{1}_{m'}(h)$ . This is also true if the pairs  $m$  and  $m'$  have one point in common. Notice, however, that the inter-point distance of a third pair of points may indeed depend on the inter-point distances of two other pairs. Consider for instance three points  $p_1, p_2, p_3$ ; while the inter-point distances  $|p_1 - p_2|$  and  $|p_1 - p_3|$  are uncorrelated, the distance  $|p_2 - p_3|$  will depend on the previous two distances. Inter-point distances can therefore be regarded as uncorrelated random variables, though their joint distribution is not independent.

In order to find an expression for the variance of  $\tilde{K}(h)$ , it is useful to view the indicator function  $\mathbf{1}_m(h)$  as a Bernoulli random variable with parameter  $p = \frac{\pi h^2}{A}$ . The expectation of the product of Bernoulli variables, such as  $\mathbf{1}_m(h) \cdot \mathbf{1}_{m'}(h)$ , can be written as

$$E(\mathbf{1}_m(h) \cdot \mathbf{1}_{m'}(h)) \simeq \begin{cases} p & \text{for } m = m' \\ p^2 & \text{for } m \neq m', \end{cases}$$

Therefore, the variance of  $\sqrt{A}\tilde{K}(h)$  can be approximated as

$$\begin{aligned} \text{Var}(\sqrt{A}\tilde{K}(h) \mid N > 1) &= E(A\tilde{K}^2(h) \mid N > 1) - E^2(\sqrt{A}\tilde{K}(h) \mid N > 1) \\ &\simeq E\left(\frac{4A^3 \cdot \sum_m \sum_{m'} \mathbf{1}_m(h) \cdot \mathbf{1}_{m'}(h)}{N^2(N-1)^2} \mid N > 1\right) - A\pi^2 h^4 \\ &\simeq \frac{2\pi h^2}{\lambda^2} \left(1 - \frac{\pi h^2}{A}\right) \cdot E\left(\frac{\lambda^2 A^2}{N(N-1)} \mid N > 1\right) \end{aligned} \quad (7)$$

$$\begin{aligned} &\approx \frac{2\pi h^2}{\lambda^2} \cdot E\left(\frac{\lambda^2 A^2}{N(N-1)} \mid N > 1\right) \\ &\approx \frac{2\pi h^2}{\lambda^2}. \end{aligned} \quad (8)$$

Step (7) once again disregards edge effects and follows from the fact that each sum has  $\frac{1}{2}N(N-1)$  terms. The two following approximations hold for a large area of observation  $A$  and the result in (8) is true since the conditional expectation will be close to unity as shown in the following lemma.

**Lemma 2.1** *Let  $\mathbf{N}^{(n)}$  be a sequence of homogeneous Poisson processes with intensity  $\lambda > 0$  defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$  and suppose that  $A^{(n)} \rightarrow \infty$ , as  $n \rightarrow \infty$ . Further, let*

$$X^{(n)} = \frac{N^{(n)}(N^{(n)} - 1)}{\lambda^2 (A^{(n)})^2}$$

and define the bounded measurable function

$$\varphi(x) = \begin{cases} \frac{1}{x} & \text{for } x \geq b, \quad 0 < b < 1 \\ 0 & \text{for } x < b. \end{cases}$$

Then,

$$E\left(\varphi\left(X^{(n)}\right)\right) \xrightarrow{n} 1$$

and

$$E\left(\varphi\left(X^{(n)}\right) \mid X^{(n)} > 0\right) \xrightarrow{n} 1$$

**Proof.**

The proof will use a version of the Helly-Bray theorem which describes the convergence of expectations (see for instance Ferguson (1996) p. 13). The sequence  $X^{(n)}$  converges in law to the constant 1, since

$$\begin{aligned} E\left(X^{(n)}\right) &= 1 \\ \text{Var}\left(X^{(n)}\right) &= \frac{4}{\lambda A^{(n)}} + \frac{2}{\lambda^2 (A^{(n)})^2} \xrightarrow{n} 0, \end{aligned}$$

which makes use of the following expressions for the first four moments for the Poisson distribution:

$$\begin{aligned} E\left(N^{(n)}\right) &= \lambda A^{(n)}, \\ E\left(\left(N^{(n)}\right)^2\right) &= \lambda A^{(n)} \left(1 + \lambda A^{(n)}\right), \\ E\left(\left(N^{(n)}\right)^3\right) &= \lambda A^{(n)} \left(1 + 3\lambda A^{(n)} + \lambda^2 \left(A^{(n)}\right)^2\right), \\ E\left(\left(N^{(n)}\right)^4\right) &= \lambda A^{(n)} \left(1 + 7\lambda A^{(n)} + 6\lambda^2 \left(A^{(n)}\right)^2 + \lambda^3 \left(A^{(n)}\right)^3\right). \end{aligned}$$

The Helly-Bray theorem requires that the function  $\varphi$  be measurable, bounded, and continuous at the limit of  $X^{(n)}$ . Without loss of generality,

let  $b = \frac{1}{2}$ . All these conditions are met since  $\varphi$  is measurable, bounded below by 0, bounded above by  $\frac{1}{b} = 2$ , and since  $\varphi$  is continuous at the limit of  $X^{(n)}$  which is 1. Hence, the expectation of the function of the sequence, converges to the expectation of the function of the limit which is unity:

$$E\left(\varphi\left(X^{(n)}\right)\right) \xrightarrow{n} E(\varphi(1)) = 1.$$

The conditional expectation can be expressed in terms of the unconditional expectation, since

$$E\left(\varphi\left(X^{(n)}\right) \mid X^{(n)} > 0\right) = \frac{E\left(\varphi\left(X^{(n)}\right)\right) - \varphi(0) \cdot \text{prob}(X^{(n)} = 0)}{1 - \text{prob}(X^{(n)} \leq 0)} \xrightarrow{n} 1.$$

This is true, since  $\varphi(0)$  is defined as 0 and  $\text{prob}(X^{(n)} \leq 0) = \text{prob}(X^{(n)} = 0) = \text{prob}(N^{(n)} \leq 1)$  is the probability of observing a point pattern with less than two events. This probability converges to zero in probability as the area of observation approaches infinity.

q.e.d.

All the aforementioned approximations hold for large  $A$  and therefore this section can be summarized in following theorem on the asymptotic expectation and variance of  $\tilde{K}^{(n)}(h)$ .

**Theorem 2.2** *Let  $\mathbf{N}^{(n)}$  be a sequence of homogeneous Poisson processes with intensity  $\lambda > 0$  and  $K$ -functions  $\tilde{K}^{(n)}(h)$ ,  $h > 0$ , defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$ . Suppose that  $h^2/A^{(n)} \rightarrow 0$  and  $A^{(n)} \rightarrow \infty$ , as  $n \rightarrow \infty$ . Further, assume that the boundaries of  $\mathcal{A}^{(n)}$  are sufficiently regular so that the previous approximations regarding boundary effects hold for large  $A^{(n)}$ . For the normalized  $K$ -function*

$$\tilde{\kappa}^{(n)}(h) = \frac{\tilde{K}^{(n)}(h) - \pi h^2}{\sqrt{\frac{2\pi h^2}{\lambda^2 A^{(n)}}}},$$

the asymptotic expectation and variance are

$$E\left(\tilde{\kappa}^{(n)}(h)\right) \xrightarrow{n} 0,$$

and

$$\text{Var}\left(\tilde{\kappa}^{(n)}(h)\right) \xrightarrow{n} 1.$$

This result coincides with the asymptotic expectation and variance of  $\hat{K}(h)$  in (3).

### 2.3 A Caveat Concerning Simulation Studies

As is pointed out by Stoyan and Stoyan (2000), it is crucial to use an *estimate* of  $\lambda$  or  $\lambda^2$  rather than their true values, even if they are known. Situations where the true intensity is known can arise in simulation studies, as it may be tempting to use the true value for the intensity in the definitions of the  $K$ -function estimates (1) and (5). Somewhat surprisingly, using the true value for  $\lambda$  or  $\lambda^2$  will actually inflate the variance of  $\hat{K}(h)$  and  $\tilde{K}(h)$  by a factor of  $(1 + 2\pi h^2 \lambda)$ . Consider for instance the modified estimator  $\check{K}(h)$  which is defined analogously to  $\tilde{K}(h)$  but uses the true value for  $\lambda$ :

$$\check{K}(h) = \frac{2}{\lambda^2 A} \sum_m \mathbf{1}_m(h).$$

The expectation can be derived straightforwardly as

$$\begin{aligned} E(\check{K}(h) \mid N > 1) &\simeq \frac{\pi h^2}{\lambda^2 A^2} E(N(N-1) \mid N > 1) \\ &= \frac{\pi h^2}{1 - \text{prob}(N \leq 1)} \\ &\approx \pi h^2, \end{aligned} \tag{9}$$

which requires expressions for the first and second moments of  $N$  (see Lemma 2.1) and the approximation in (9) works well for large  $A$  so that the probability of obtaining a point pattern with less than two points is negligible.

In order to find the variance of  $\sqrt{A}\check{K}(h)$ , consider

$$\begin{aligned} \text{Var}(\sqrt{A}\check{K}(h) \mid N > 1) &= E(A\check{K}^2(h) \mid N > 1) - E^2(\sqrt{A}\check{K}(h) \mid N > 1) \\ &\simeq E\left(\frac{4 \cdot \sum_m \sum_{m'} \mathbf{1}_m(h) \cdot \mathbf{1}_{m'}(h)}{\lambda^4 A} \mid N > 1\right) - A\pi^2 h^4 \\ &= \frac{4}{\lambda^4 A} E\left(\frac{\pi^2 h^4}{A^2} \frac{1}{4} N^2 (N-1)^2 \right. \\ &\quad \left. + \frac{\pi h^2}{A} \left(1 - \frac{\pi h^2}{A}\right) \frac{1}{2} N(N-1) \mid N > 1\right) \\ &\quad - A\pi^2 h^4 \\ &\approx \frac{2\pi h^2}{\lambda^2} (1 + 2\pi h^2 \lambda), \end{aligned} \tag{10}$$



again assuming in (10) that  $\text{prob}(N \leq 1)$  is negligible as the area of observation approaches infinity. This variance is larger than the variance of  $\tilde{K}(h)$  as shown in (8) by a factor of  $(1 + 2\pi h^2 \lambda)$ . A similarly inflated variance is obtained when using the true  $\lambda$  instead of the estimate  $\hat{\lambda}$  in the definition of  $\tilde{K}(h)$ . Heinrich (1988) derived a similar result in  $\mathbf{R}^d$ .

## 2.4 An Alternative Derivation of the Asymptotic Distribution of the $K$ -function for a Homogeneous Poisson Process

**Theorem 2.3** *Let  $\mathbf{N}^{(n)}$  be a sequence of homogeneous Poisson processes with intensities  $\lambda^{(n)}$  and  $K$ -functions  $\tilde{K}^{(n)}(h_n)$ , defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$  and let  $N^{(n)}$  denote the number of points. Suppose that for each  $n$ , the observed region  $\mathcal{A}^{(n)}$  can be broken up into disjoint subregions  $\mathcal{A}_1^{(n)}, \mathcal{A}_2^{(n)}, \dots, \mathcal{A}_{I_n}^{(n)}$  each having area  $A_i^{(n)} = A^{(n)}/I_n$  and let  $N_i^{(n)}$  be the number of points in area  $i$ . Suppose also that for some scalar  $\lambda_{\min}$ ,  $0 < \lambda_{\min} \leq \lambda^{(n)} < \infty$  for all  $n$ . In addition, suppose that, as  $n \rightarrow \infty$ ,  $I_n \rightarrow \infty$  and  $h_n^2/A_i^{(n)} \rightarrow 0$ . Further, assume that the boundaries of  $\mathcal{A}_i^{(n)}$  are sufficiently regular that the asymptotic expectation and variance of  $\tilde{K}^{(n)}(h_n)$  can be derived as done in Theorem 2.2 and that the number of pairs of points  $(p_r, p_s)$  with  $|p_r - p_s| \leq h_n$  such that  $p_r$  and  $p_s$  are in distinct subregions is small, satisfying  $R^{(n)} := \frac{2A^{(n)}}{N^{(n)}(N^{(n)}-1)} \sum_{p_r, p_s} \mathbf{1}(|p_r - p_s| \leq h_n) \mathbf{1}(i \neq j) \rightarrow 0$  in probability as  $n \rightarrow \infty$ , where the sum is over all  $p_r \in A_i^{(n)}, p_s \in A_j^{(n)}$ . Then the normalized quantity*

$$\tilde{\kappa}^{(n)}(h_n) = \frac{\tilde{K}^{(n)}(h_n) - \pi h_n^2}{\sqrt{\frac{2\pi h_n^2}{(\lambda^{(n)})^2 A^{(n)}}}},$$

is asymptotically Normal as  $n \rightarrow \infty$ :

$$\tilde{\kappa}^{(n)}(h_n) \rightsquigarrow N(0, 1).$$

### Proof.

First note that  $\tilde{K}^{(n)}(h_n)$  can be represented as the arithmetic mean of  $K$ -functions computed individually on each of the subregions  $i = 1, 2, \dots, I_n$ , plus the remainder term  $R^{(n)}$  defined above:

$$\tilde{K}^{(n)}(h_n) = \frac{2A^{(n)}}{N^{(n)}(N^{(n)}-1)} \sum_m \mathbf{1}_m(h_n)$$

$$\begin{aligned}
&= 2A^{(n)} \sum_{i=1}^{I_n} \left\{ \frac{1}{N_i^{(n)}(N_i^{(n)} - 1)} \sum_{m_i} \mathbf{1}_{m_i}(h_n) \right\} + R^{(n)} \\
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \left\{ \frac{2A_i^{(n)}}{N_i^{(n)}(N_i^{(n)} - 1)} \sum_{m_i} \mathbf{1}_{m_i}(h_n) \right\} + R^{(n)} \quad (11)
\end{aligned}$$

$$= \frac{1}{I_n} \sum_{i=1}^{I_n} \left\{ \tilde{K}_i^{(n)}(h_n) \right\} + R^{(n)}, \quad (12)$$

where  $m_i$  indexes the pairs of points in subregion  $i$ . Step (11) follows from  $A_i^{(n)} = A^{(n)}/I_n$  as defined in the theorem and  $R^{(n)}$  converges to zero in probability by assumption. It follows that the distribution of  $\tilde{K}^{(n)}(h_n)$  is equivalent to that of the mean of  $I_n$   $K$ -functions as shown in (12). By the central limit theorem, this mean is asymptotically Normal and the variance of  $\tilde{K}^{(n)}(h_n)$  can be computed as

$$\begin{aligned}
\text{Var} \left( \tilde{K}^{(n)}(h_n) \right) &= \text{Var} \left( \frac{1}{I_n} \sum_{i=1}^{I_n} \tilde{K}_i^{(n)}(h_n) \right) + o(n) \\
&= \frac{1}{I_n^2} \sum_{i=1}^{I_n} \text{Var} \left( \tilde{K}_i^{(n)}(h_n) \right) + o(n) \\
&= \frac{1}{I_n^2} \sum_{i=1}^{I_n} \frac{2\pi h_n^2}{(\lambda^{(n)})^2 A_i^{(n)}} + o(n) \\
&= \frac{2\pi h_n^2}{(\lambda^{(n)})^2 A^{(n)}} + o(n). \quad (13)
\end{aligned}$$

q.e.d.

**Corollary 2.4** *Under the conditions of Theorem 2.3,  $\hat{K}^{(n)}(h^{(n)})$  defined as in (1) is asymptotically Normal. The standardized quantity*

$$\hat{\kappa}^{(n)}(h_n) = \frac{\hat{K}^{(n)}(h_n) - \pi h_n^2}{\sqrt{\frac{2\pi h_n^2}{(\lambda^{(n)})^2 A^{(n)}}}},$$

*approaches the standard Normal distribution as  $n \rightarrow \infty$ :*

$$\hat{\kappa}^{(n)}(h_n) \rightsquigarrow N(0, 1).$$

**Proof.**

Ripley's K-function can be written as  $\hat{K}^{(n)}(h^{(n)}) = \tilde{K}^{(n)}(h^{(n)}) \cdot \frac{N^{(n)}-1}{N^{(n)}}$ ,  $N^{(n)} >$

1. Since  $\frac{N^{(n)}-1}{N^{(n)}} \xrightarrow{n} 1$  in probability,  $\hat{K}^{(n)}(h^{(n)})$  approaches the same asymptotic distribution as  $\tilde{K}^{(n)}(h^{(n)})$  by Slutsky's Theorem (see for instance Ferguson (1996) p. 39).

q.e.d.

### 3 Model Assessment Using the $K$ -function and its Weighted Analog $K_W$

#### 3.1 Residual Analysis of Point Process Models Using Ripley's $K$ -function

The assessment of point process models using residual analysis involves two steps. In the first step, the a hypothesized point process model is used to transform an observed point pattern. The transformation is such that the transformed process – called the residual process or simply the residuals – will be homogeneous Poisson if the hypothesized model is correct, i.e. if the observed data were indeed generated by the null hypothesis model. The second step of this procedure, involves analyzing the residual process. If the residual process is homogeneous Poisson, the hypothesized model seems to describe the data well. Otherwise, the null hypothesis is rejected and a further analysis of the residual process may provide clues as to where the null hypothesis model ought to be improved.

The expression “residual analysis” can be understood in analogy to regression as the transformation converts the observed point pattern (in regression this would be the observed dependent variable) into a new process by somehow compensating for the model in question. The resulting residual point process then represents the inherent randomness in the data as well as those features, which are not captured by the model used for the transformation. In regression analysis the transformation means that the predicted dependent variable is subtracted from its observed value. In the realm of point processes, the observed point pattern can be re-scaled or thinned in order to generate a residual process. To complete the analogy, regression analysis expects the residuals to be “white noise” (mostly defined as an i.i.d. Normal random variables) if the null hypothesis model is correct. The white noise of point processes is the homogeneous Poisson process and similarly the transformation of observed point patterns generates a homogeneous Poisson process if the null hypothesis model is correct.

A well established method of generating residual point processes is re-scaling. This methodology is based on a result by Meyer (1971) and applications of point process residual analysis based on re-scaling can be found in Ogata (1988) Merzbach and Nualart (1986) and Schoenberg (1999).

A second popular method is thinning, which is based on a technique originally designed to simulate inhomogeneous Poisson processes (Lewis and Shedler 1979). This procedure has been generalized by Ogata (1981) for the simulation of more general classes of point process models. A reversal of this methodology allows the transformation of an observed point pattern into a homogeneous Poisson process by deleting certain points of the data set (see for instance Schoenberg (2003) for an application of this method to earthquake occurrence data).

Thinning can be applied to a wide range of point process models. For an inhomogeneous Poisson process with rate  $\lambda(x, y) > 0$  on  $\mathcal{A} \in \mathbf{R}^2$ , a homogeneous Poisson process with rate  $\lambda_*$  can be generated by retaining each point  $(x, y)$  of the observed point pattern with probability  $\lambda_*/\lambda(x, y)$ , where the quantity  $\lambda_*$  ought not be larger than the infimum of the observed rate, i.e.  $0 < \lambda_* \leq \inf_{(x,y) \in \mathcal{A}} \lambda(x, y)$ .

After transforming an observed point pattern either by re-scaling or thinning, the  $K$ -function can be applied to the residual process in order to investigate the homogeneity of the residuals. The result can be interpreted as a test of the goodness-of-fit of the point process model in question. Inference can be performed either by Monte Carlo simulation or by using the results for the asymptotic distribution of the  $K$ -function for a homogeneous Poisson process as presented in the previous section.

Other methods for assessing the homogeneity of a point process exist, including tests for monotonicity (Saw 1975), uniformity (Dijkstra, Rietjens, and Steutel 1984, Lawson 1988, Lisek and Lisek 1985), and tests on the second and higher-order properties of the process (Bartlett 1964, Davies 1977, Heinrich 1991). Likelihood statistics, such as Akaike's Information Criterion (AIC) (Akaike 1974) and the Bayesian Information Criterion (BIC) (Schwartz 1979) are often used to assess more general classes of models including space-time point process models. See for instance Ogata (1998) for an application to earthquake occurrence models.

### 3.2 Point Process Model Assessment Using the Mean $K$ -function of Thinned Residuals $K_M(h)$

Suppose that a given planar point process in a connected subset  $\mathcal{A}$  of  $\mathbf{R}^2$  with finite area  $A$  may be specified by its conditional intensity with respect

to some filtration on  $\mathcal{A}$ , for  $(x, y) \in \mathcal{A}$  (Daley and Vere-Jones 2003). 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 the conditional intensity of the point process is given by  $\lambda(x, y)$ .

The mean  $K$ -function of thinned residuals ( $K_M(h)$ ) is the average of  $T$  ordinary  $K$ -functions, denoted as  $\tilde{K}_t(h)$  with  $t = \{1, 2, \dots, T\}$ , each computed on one of the  $T$  residual processes generated by thinning the observed point pattern:

$$K_M(h) = \frac{1}{T} \sum_{t=1}^T \tilde{K}_t(h).$$

Notice that  $K_M(h)$  could be defined similarly as the average of  $\hat{K}(h)$ .

It is important to realize that  $\tilde{K}_t(h)$  and  $\tilde{K}_\tau(h)$  are not independent of one another for different iterations of thinning ( $t \neq \tau$ ). A derivation of the asymptotic distribution of  $K_M(h)$  will therefore require an expression for this dependence.

Consider a homogeneous Poisson process  $\mathbf{N}$  on connected subsets  $\mathcal{A} \in \mathbf{R}^2$  of area  $A$ , rate  $\lambda \geq \lambda_* > 0$ , and sufficiently well-behaved boundaries as required in the preceding sections. Generate the thinned process  $\mathbf{N}_1$  by retaining each point of  $\mathbf{N}$  with probability  $\lambda_*/\lambda$ . The  $K$ -function of process  $\mathbf{N}_1$ , which has  $N_1$  retained points, will be denoted as  $\tilde{K}_1$  and has the following expectation and variance

$$E(\tilde{K}_1) \simeq \pi h^2 \tag{14}$$

$$Var(\sqrt{A}\tilde{K}_1) \simeq \frac{2\pi h^2}{\lambda_*^2}, \tag{15}$$

since the process  $\mathbf{N}_1$  is homogeneous Poisson with rate  $\lambda_*$  and the expressions for the expectation and variance in Theorem 2.2 can be used.

Similarly, generate a thinned process  $\mathbf{N}_2$  with  $N_2$  points and  $K$ -function  $\tilde{K}_2$  and notice that the expectation and variance of  $\tilde{K}_2$  coincides with the respective expressions for  $\tilde{K}_1$ . Denote the number of points which are in the overlap of the thinned processes  $\mathbf{N}_1$  and  $\mathbf{N}_2$  as  $N_O$ . Further, let  $m_O = \{1, 2, \dots, \frac{1}{2}N_O(N_O - 1)\}$  be an index for the pairs of points in this overlap. The covariance of  $\sqrt{A}\tilde{K}_1$  and  $\sqrt{A}\tilde{K}_2$  can then be computed as

$$\begin{aligned} Cov & \left( \sqrt{A}\tilde{K}_1, \sqrt{A}\tilde{K}_2 \mid N_O > 1 \right) \\ & = E \left( A\tilde{K}_1\tilde{K}_2 \mid N_O > 1 \right) - E \left( \sqrt{A}\tilde{K}_1 \mid N_O > 1 \right) E \left( \sqrt{A}\tilde{K}_2 \mid N_O > 1 \right) \end{aligned}$$

$$\begin{aligned}
&\simeq E \left( \frac{4A^3 \cdot \sum_{m_O} \frac{\pi h^2}{A} \left(1 - \frac{\pi h^2}{A}\right)}{N_1(N_1 - 1)N_2(N_2 - 1)} \middle| N_O > 1 \right) \\
&= E \left( \frac{4A^3 \cdot \sum_{m_O} \frac{\pi h^2}{A} \left(1 - \frac{\pi h^2}{A}\right)}{N_O(N_O - 1)} \frac{N_O(N_O - 1)}{N_1(N_1 - 1)N_2(N_2 - 1)} \middle| N_O > 1 \right) \\
&= \frac{2\pi h^2}{\lambda^2} \left(1 - \frac{\pi h^2}{A}\right) E \left( \frac{\lambda^2 A^2 N_O(N_O - 1)}{N_1(N_1 - 1)N_2(N_2 - 1)} \middle| N_O > 1 \right) \\
&\approx \frac{2\pi h^2}{\lambda^2}, \tag{16}
\end{aligned}$$

where the approximation in step (16) works for large  $A$  and uses the following Lemma 3.1 which proves that the conditional expectation approaches unity in this case. Note that conditioning on  $N_O > 1$  implies that  $N_1, N_2 > 1$ , since there cannot be more points in the overlap than in each of the thinned processes. This guarantees that none of the denominators attain a value of zero. This result for the covariance implies that the correlation between  $K$ -functions of independently thinned homogeneous Poisson processes is

$$\rho(\sqrt{A}\tilde{K}_1, \sqrt{A}\tilde{K}_2) = \frac{\text{Cov}(\sqrt{A}\tilde{K}_1, \sqrt{A}\tilde{K}_2)}{\sqrt{\text{Var}(\sqrt{A}\tilde{K}_1) \cdot \text{Var}(\sqrt{A}\tilde{K}_2)}} = \frac{\lambda_*^2}{\lambda^2}.$$

**Lemma 3.1** *Let  $\mathbf{N}^{(n)}$ , be a sequence of homogeneous Poisson processes with intensity  $\lambda > 0$  defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$  and suppose that  $A^{(n)} \rightarrow \infty$ , as  $n \rightarrow \infty$ . Let  $0 < \lambda_* < \lambda$  and let  $\mathbf{N}_1^{(n)}$  and  $\mathbf{N}_2^{(n)}$  be thinned processes in which each point of  $\mathbf{N}^{(n)}$  is retained with probability  $\lambda_*/\lambda$  as described above. Let  $N_1^{(n)}$  and  $N_2^{(n)}$  denote the number of retained points in the thinned processes  $\mathbf{N}_1^{(n)}$  and  $\mathbf{N}_2^{(n)}$  and let  $N_O^{(n)}$  denote the number of points which are common to both thinned processes. Further, let  $\vec{X}^{(n)}$  be a sequence of vectors defined as*

$$\vec{X}^{(n)} = \begin{pmatrix} X_1^{(n)} \\ X_2^{(n)} \\ X_O^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{N_1^{(n)}(N_1^{(n)}-1)}{\lambda^2(A^{(n)})^2} \\ \frac{N_2^{(n)}(N_2^{(n)}-1)}{\lambda^2(A^{(n)})^2} \\ \frac{N_O^{(n)}(N_O^{(n)}-1)}{\lambda^2(A^{(n)})^2} \end{pmatrix},$$

and define the bounded measurable function  $\varphi(\vec{x})$ ,  $\vec{x} = (x_1, x_2, x_O)$ , such that

$$\varphi(\vec{x}) = \begin{cases} c & \text{for } \frac{x_O}{x_1 x_2} \geq c, \quad 1 < c < \infty \\ 0 & \text{for } x_1 x_2 = 0 \\ \frac{x_O}{x_1 x_2} & \text{otherwise.} \end{cases}$$

Then,

$$E\left(\varphi\left(\vec{X}^{(n)}\right)\right) \xrightarrow{n} 1$$

and

$$E\left(\varphi\left(\vec{X}^{(n)}\right) \mid X_1^{(n)} \cdot X_2^{(n)} > 0\right) \xrightarrow{n} 1$$

**Proof.**

This proof is similar to the proof of Lemma 2.1 as it uses the Helly-Bray theorem, as well. The vector sequence  $\vec{X}^{(n)}$  converges in law to a constant vector

$$\vec{X}^{(n)} \xrightarrow{n} \vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_O \end{pmatrix} = \begin{pmatrix} \frac{\lambda_*^2}{\lambda^2} \\ \frac{\lambda_*^2}{\lambda^2} \\ \frac{\lambda_*^4}{\lambda^4} \end{pmatrix},$$

which can be shown using a similar argument as in the proof of Lemma 2.1. The conditions of the The Helly-Bray theorem on the function  $\varphi$  are all met, as it is measurable, bounded between zero and  $c$  (without loss of generality, let  $c = 2$ , for instance) and continuous at the limit of the vector sequence which is  $\vec{X}$ . Hence, the expectation of the function of the vector sequence, converges to the expectation of the function of the limit:

$$E\left(\varphi\left(\vec{X}^{(n)}\right)\right) \xrightarrow{n} E\left(\varphi\left(\vec{X}\right)\right) = E\left(\varphi\left(\begin{pmatrix} \frac{\lambda_*^2}{\lambda^2} \\ \frac{\lambda_*^2}{\lambda^2} \\ \frac{\lambda_*^4}{\lambda^4} \end{pmatrix}\right)\right) = 1.$$

For the same reason as in the proof of Lemma 2.1, the conditional expectation will be the same as the unconditional expectation as the probability of observing an overlap with less than two events rapidly converges to zero as the area of observation approaches infinity.

q.e.d.

The result for the covariance of  $K$ -functions  $\tilde{K}_1$  and  $\tilde{K}_2$ , computed on thinned residual processes will be used in the following theorem stating that

the mean  $K$ -function of thinned residuals ( $K_M(h)$ ) is asymptotically Normal for inhomogeneous Poisson processes. This allows the use of  $K_M(h)$  for assessing the goodness-of-fit of point process models in the inhomogeneous case.

**Theorem 3.2** *Let  $\mathbf{N}^{(n)}$  be a sequence of inhomogeneous Poisson processes with intensities  $\lambda^{(n)}$  defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$  and let  $N^{(n)}$  denote the number of points. Suppose that for each  $n$ , the observed region  $\mathcal{A}^{(n)}$  can be broken up into disjoint subregions  $\mathcal{A}_1^{(n)}, \mathcal{A}_2^{(n)}, \dots, \mathcal{A}_{I_n}^{(n)}$  each having area  $A_i^{(n)} = A^{(n)}/I_n$  with  $N_i^{(n)}$  points, and that the intensity  $\lambda_i^{(n)}$  is constant within  $\mathcal{A}_i^{(n)}$ . Let  $\lambda_*^{(n)}$  denote the step  $n$  infimum of the intensity,  $\lambda_*^{(n)} = \inf_{\mathcal{A}^{(n)}} \lambda^{(n)}$ , and let there be a scalar  $\lambda_{\min}$  such that  $0 < \lambda_{\min} \leq \lambda_*^{(n)} \leq \lambda_i^{(n)}$  for all  $i, n$ . Let  $K_M^{(n)}$ , denote the mean  $K$ -function of thinned residuals, defined as the arithmetic mean of  $T_n$   $K$ -functions  $\tilde{K}_t^{(n)}, t = \{1, 2, \dots, T_n\}$ , each computed on the thinned processes  $\mathbf{N}_t^{(n)}$ , generated such that  $\mathbf{N}_t^{(n)}$  is a homogeneous Poisson processes with rate  $\lambda_*^{(n)}$ . Moreover, let  $N_t^{(n)}$  denote the number of points of the thinned process and let  $N_{ti}^{(n)}$  denote the number of points in subregion  $i$  of the thinned process. In addition, suppose that, as  $n \rightarrow \infty$ ,  $I_n \rightarrow \infty$ ,  $T_n \rightarrow \infty$ , and  $h_n^2/A_i^{(n)} \rightarrow 0$ . Further, assume that the boundaries of  $\mathcal{A}_i^{(n)}$  are sufficiently regular that the number of pairs of points  $(p_{tr}, p_{ts})$  with  $|p_{tr} - p_{ts}| \leq h_n$  such that  $p_{tr}$  and  $p_{ts}$  are in distinct subregions is small for each thinned process  $\mathbf{N}_t^{(n)}$ , satisfying  $R_t^{(n)} := \frac{2A^{(n)}}{N_t^{(n)}(N_t^{(n)} - 1)} \sum_{p_{tr}, p_{ts}} \mathbf{1}(|p_{tr} - p_{ts}| \leq h_n) \mathbf{1}(i \neq j) \rightarrow 0$  in probability as  $n \rightarrow \infty$ , where the sum is over all  $p_{tr} \in \mathcal{A}_i^{(n)}, p_{ts} \in \mathcal{A}_j^{(n)}$ . Then the normalized quantity*

$$\kappa_M^{(n)}(h_n) = \frac{\tilde{K}_M^{(n)}(h_n) - \pi h_n^2}{\sqrt{\frac{2\pi h_n^2}{H((\lambda^{(n)})^2)A^{(n)}}}},$$

is asymptotically Normal as  $n \rightarrow \infty$ :

$$\kappa_M^{(n)}(h_n) \approx N(0, 1),$$

where  $H((\lambda^{(n)})^2)$  represents the harmonic mean of the squared intensity within the observed region  $\mathcal{A}^{(n)}$ .

**Proof.**

Using a similar representation of as done in (12),  $K_M^{(n)}$  (short for  $K_M^{(n)}(h_n)$ )



can be written as an arithmetic mean of  $I_n$  local  $K_{Mi}^{(n)}$ -functions each computed as the mean of  $T_n$  local  $\tilde{K}_{ti}^{(n)}$ -functions (short for  $\tilde{K}_{ti}^{(n)}(h_n)$ ) of the thinned process  $\mathbf{N}_t^{(n)}$  restricted to the subregion  $i$  in addition to a sum over the remainder term  $R_t^{(n)}$  defined above:

$$\begin{aligned}
K_M^{(n)} &= \frac{1}{T_n} \sum_{t=1}^{T_n} \tilde{K}_t^{(n)} \\
&= \frac{1}{T_n} \sum_{t=1}^{T_n} \left\{ \frac{1}{I_n} \sum_{i=1}^{I_n} \tilde{K}_{ti}^{(n)} + R_t^{(n)} \right\} \\
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \left\{ \frac{1}{T_n} \sum_{t=1}^{T_n} \tilde{K}_{ti}^{(n)} \right\} + \frac{1}{T_n} \sum_{t=1}^{T_n} R_t^{(n)} \\
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \left\{ K_{Mi}^{(n)} \right\} + \frac{1}{T_n} \sum_{t=1}^{T_n} R_t^{(n)} \\
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \left\{ K_{Mi}^{(n)} \right\} + o(n), \tag{17}
\end{aligned}$$

where the last step follows from the condition of the theorem that the remainder term  $R_t^{(n)}$  converges to zero in probability for each thinned process  $\mathbf{N}_t^{(n)}$ . The next step involves finding expressions for the expectation and variance of the local  $K_{Mi}^{(n)}$ -functions.

$$\begin{aligned}
E\left(K_{Mi}^{(n)}\right) &= E\left(\frac{1}{T_n} \sum_{t=1}^{T_n} \tilde{K}_{ti}^{(n)}\right) \\
&= \frac{1}{T_n} \sum_{t=1}^{T_n} E\left(\tilde{K}_{ti}^{(n)}\right) \\
&= \pi h^2,
\end{aligned}$$

since the expectation of each a local  $\tilde{K}_{ti}^{(n)}$ -function is  $\pi h^2$  for each individual thinning (see (14)). The variance of  $K_{Mi}^{(n)}$  is

$$\begin{aligned}
Var\left(\sqrt{A_i^{(n)}} K_{Mi}^{(n)}\right) &= Var\left(\frac{1}{T_n} \sum_{t=1}^{T_n} \sqrt{A_i^{(n)}} \tilde{K}_{ti}^{(n)}\right) \\
&= \frac{1}{T_n^2} \sum_{t=1}^{T_n} Var\left(\sqrt{A_i^{(n)}} \tilde{K}_{ti}^{(n)}\right)
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{T_n^2} \sum_{t=1}^{T_n} \sum_{\tau=1, \tau \neq t}^{T_n} \text{Cov} \left( \sqrt{A_i^{(n)}} \tilde{K}_{ti}^{(n)}, \sqrt{A_i^{(n)}} \tilde{K}_{\tau i}^{(n)} \right) \\
& = \frac{1}{T_n^2} T_n \frac{2\pi h_n^2}{\left(\lambda_*^{(n)}\right)^2} + \frac{1}{T_n^2} T_n (T_n - 1) \frac{2\pi h_n^2}{\left(\lambda_i^{(n)}\right)^2}
\end{aligned}$$

which follows from the expressions for the variance and covariance of  $K$ -functions computed on thinned point patterns in (15) and (16), respectively. Notice that  $K_{Mi}^{(n)}$  is computed locally, i.e. the on the subregion  $i$  with area  $A_i^{(n)} = A^{(n)}/I_n$ . In the limit, i.e. as the number of thinning iterations  $T_n$  approaches infinity, the variance of  $K_{Mi}^{(n)}$  is therefore

$$\lim_{n \rightarrow \infty} \text{Var} \left( \sqrt{A_i^{(n)}} K_{Mi}^{(n)}(h_n) \right) = \frac{2\pi h_n^2}{\left(\lambda_i^{(n)}\right)^2} < \frac{2\pi h_n^2}{\left(\lambda_{\min}^{(n)}\right)^2},$$

which is bounded above by assumption. This allows applying the Lindeberg-Feller Central Limit Theorem in (17), meaning that  $K_M^{(n)}$  is asymptotically Normal with expectation  $\pi h_n^2$  and asymptotic variance

$$\begin{aligned}
\lim_{n \rightarrow \infty} \text{Var} \left( \sqrt{A^{(n)}} K_M^{(n)}(h_n) \right) & = \lim_{n \rightarrow \infty} \text{Var} \left( \frac{1}{I_n} \sum_{i=1}^{I_n} \sqrt{A^{(n)}} K_{Mi}^{(n)} \right) \\
& = \lim_{n \rightarrow \infty} \text{Var} \left( \frac{1}{I_n} \sum_{i=1}^{I_n} \sqrt{I_n A_i^{(n)}} K_{Mi}^{(n)} \right) \\
& = \lim_{n \rightarrow \infty} \frac{1}{I_n} \sum_{i=1}^{I_n} \frac{2\pi h_n^2}{\left(\lambda_i^{(n)}\right)^2} \\
& = \frac{2\pi h_n^2}{H \left( (\lambda^{(n)})^2 \right)}. \tag{18}
\end{aligned}$$

q.e.d.

Note that  $K_M(h)$  may be used to test a quite general class of null hypothesis models without having to assume homogeneity (like the original  $K$ -function does) and without the need to re-scale the observed point pattern. Re-scaling often introduces problems of highly irregular boundaries and such problems are discussed in Schoenberg (2003).

For the purpose of plotting, the  $K_M(h)$  can be transformed into a quantity  $L_M(h)$  with expectation  $h$ . This is done in analogy to (4) which transforms Ripley's original  $K$ -function into the  $L$ -function.

$$L_M(h) = \sqrt{\frac{K_M(h)}{\pi}}. \quad (19)$$

### 3.3 Point Process Model Assessment Using the Weighted $K$ -function $K_W(h)$

This section presents the weighted  $K$ -function,  $K_W(h)$ , and derives its distributional properties under certain conditions.  $K_W(h)$  is a weighted analogue of Ripley's  $K$ -function, which was first introduced as the *inhomogeneous  $K$ -function* in Baddeley, Møller, and Waagepetersen (2000). It provides a direct test for goodness-of-fit, without having to transform the points using residual analysis. The Mean  $K$ -function of Thinned Residuals  $K_M(h)$  is able to avoid some of the problems encountered when re-scaling an observed point pattern. However, repeated thinning iterations are required to compute  $K_M(h)$  which can introduce a large degree of sampling variability when the conditional intensity in question is highly variable (Schoenberg 2003).

As in the case of  $K_M(h)$ , suppose that a given planar point process in a connected subset  $\mathcal{A}$  of  $\mathbf{R}^2$  with finite area  $A$  may be specified by its conditional intensity with respect to some filtration on  $\mathcal{A}$ , for  $(x, y) \in \mathcal{A}$ . The conditional intensity is given by  $\lambda(x, y)$  and the weighted  $K$ -function, used to assess the model  $\lambda_0(x, y)$ , may be defined as

$$K_W(h) = \frac{1}{\lambda_*^2 A} \sum_r w_r \sum_{s \neq r} w_s \mathbf{1}(|p_r - p_s| \leq h) \quad (20)$$

where  $\lambda_* := \inf\{\lambda_0(x, y); (x, y) \in \mathcal{A}\}$  is the infimum of the conditional intensity over the observed region for the model to be assessed and  $w_r = \lambda_*/\lambda_0(p_r)$ , where  $\lambda_0(p_r)$  is the modelled conditional intensity at point  $p_r$ .

The weighted  $K$ -function integrates the two steps needed to compute  $K_M(h)$  into one step.  $K_M(h)$  is computed by estimating Ripley's  $K$ -function repeatedly on thinned data and then taking the average. The computation of the weighted  $K$ -function on the other hand uses the retaining probabilities of the thinning procedure as weights for the points in order to offset the inhomogeneity of the process. By incorporating all pairs of the observed point pattern, 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 in any finite collection of random thinnings.

Provided the conditional intensity  $\lambda$  is sufficiently smooth, it can be conjectured that  $K_W(h)$  will be asymptotically Normal as the area of observation  $A$  approaches infinity. Indeed, for the Poisson case where  $\lambda$  is locally constant on distinct subregions whose areas  $A_i^{(n)}$  are large relative to the inter-point distance  $h_n$ , the following result can be obtained.

**Theorem 3.3** *Let  $\mathbf{N}^{(n)}$  be a sequence of inhomogeneous Poisson processes with intensities  $\lambda^{(n)}$  and weighted  $K$ -functions  $K_W^{(n)}$ , defined on connected subsets  $\mathcal{A}^{(n)} \subset \mathbf{R}^2$  of finite areas  $A^{(n)}$ . Suppose that for each  $n$ , the observed region  $\mathcal{A}^{(n)}$  can be broken up into disjoint subregions  $\mathcal{A}_1^{(n)}, \mathcal{A}_2^{(n)}, \dots, \mathcal{A}_{I_n}^{(n)}$  each having area  $A_i^{(n)} = A^{(n)}/I_n$ , and that the intensity  $\lambda_i^{(n)}$  is constant within  $\mathcal{A}_i^{(n)}$ . Suppose also that for some scalar  $\lambda_{\min}$ ,  $0 < \lambda_{\min} \leq \lambda_i^{(n)} < \infty$  for all  $i, n$ . In addition, suppose that, as  $n \rightarrow \infty$ ,  $I_n \rightarrow \infty$  and  $h_n^2/A_i^{(n)} \rightarrow 0$ . Further, assume that the boundaries of  $\mathcal{A}_i^{(n)}$  are sufficiently regular that the number of pairs of points  $(p_r, p_s)$  with  $|p_r - p_s| \leq h_n$  such that  $p_r$  and  $p_s$  are in distinct subregions is small, satisfying  $R^{(n)} := \frac{1}{A^{(n)}} \sum_{p_r, p_s} \frac{\mathbf{1}(|p_r - p_s| \leq h_n) \mathbf{1}(i \neq j)}{\lambda_i^{(n)} \lambda_j^{(n)}} \rightarrow 0$  in probability as  $n \rightarrow \infty$ , where the sum is over all  $p_r \in \mathcal{A}_i^{(n)}, p_s \in \mathcal{A}_j^{(n)}$ . Then the normalized quantity*

$$\kappa_W^{(n)}(h_n) = \frac{\tilde{K}_W^{(n)}(h_n) - \pi h_n^2}{\sqrt{\frac{2\pi h_n^2}{H((\lambda^{(n)})^2)A^{(n)}}}},$$

is asymptotically Normal as  $n \rightarrow \infty$ :

$$\kappa_W^{(n)}(h_n) \approx N(0, 1),$$

where  $H((\lambda^{(n)})^2)$  represents the harmonic mean of the squared intensity within the observed region  $\mathcal{A}^{(n)}$ .

**Proof.**

We first show that  $K_W^{(n)}(h_n)$  can be represented as the arithmetic mean of  $K$ -functions computed individually on each of the squares  $i = 1, 2, \dots, I_n$ , plus the remainder term  $R^{(n)}$  defined above:

$$\begin{aligned} K_W^{(n)}(h_n) &= \frac{1}{\lambda_*^2 A^{(n)}} \sum_r w_r \sum_{s \neq r} w_s \mathbf{1}(|p_r - p_s| \leq h_n) \quad (21) \\ &= \frac{1}{\lambda_*^2 A^{(n)}} \sum_{i=1}^{I_n} \frac{\lambda_*^2}{(\hat{\lambda}_i^{(n)})^2} \sum_{r_i} \sum_{s_i \neq r_i} \mathbf{1}(|p_{r_i} - p_{s_i}| \leq h_n) + R^{(n)} \quad (22) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \frac{1}{(\hat{\lambda}_i^{(n)})^2 A_i^{(n)}} \sum_{r_i} \sum_{s_i \neq r_i} \mathbf{1}(|p_{r_i} - p_{s_i}| \leq h_n) + R^{(n)} \\
&= \frac{1}{I_n} \sum_{i=1}^{I_n} \hat{K}_i^{(n)}(h_n) + R^{(n)} \tag{23}
\end{aligned}$$

Since the intensity  $\lambda_i^{(n)}$  is constant on each square  $\mathcal{A}_i^{(n)}$ , the weights  $w_r, w_s$  assigned to a pair of points in  $\mathcal{A}_i^{(n)}$  within distance  $h_n$  are each  $\lambda_*/\hat{\lambda}_i^{(n)}$ , which is used in going from (21) to (22). Thus, since  $R^{(n)}$  converges to zero in probability by assumption, the distribution of the weighted  $K$ -function is equivalent to that of the mean of the  $I_n$  ordinary  $K$ -functions in (23).

Under the conditions of the theorem,  $\hat{K}_i^{(n)}$  is asymptotically Normal from (Ripley 1988), and since the point process on  $\mathcal{A}_i^{(n)}$  is homogeneous Poisson with rate  $\lambda_i^{(n)} \geq \lambda_{\min} > 0$ , the variance of  $\hat{K}_i^{(n)}$  is bounded above by the variance of a homogeneous Poisson process on  $\mathcal{A}_i^{(n)}$  with rate  $\lambda_{\min}$ . This

implies that the collection of random variables  $\left\{ \frac{\hat{K}_i^{(n)}(h_n) - \pi h_n^2}{I_n \sqrt{\text{Var}(\hat{K}_i^{(n)}(h_n))}} \right\}$  satisfies the Lindeberg condition (see e.g. Durrett (1991) p. 98), and therefore the mean  $\frac{1}{I_n} \sum_{i=1}^{I_n} \hat{K}_i^{(n)}(h_n)$  is asymptotically Normal. The variance of  $K_W^{(n)}(h)$  is  $\text{Var} \left( \frac{1}{I_n} \sum_{i=1}^{I_n} \hat{K}_i^{(n)}(h) \right) + o(n)$ , which can be computed as

$$\begin{aligned}
\text{Var} \left( \frac{1}{I_n} \sum_{i=1}^{I_n} \hat{K}_i^{(n)}(h) \right) &= \frac{1}{I_n^2} \sum_{i=1}^{I_n} \text{Var} \left( \hat{K}_i^{(n)}(h) \right) \\
&= \frac{1}{I_n^2} \sum_{i=1}^{I_n} \frac{2\pi h^2}{(\lambda_i^{(n)})^2 A_i^{(n)}} \\
&= \frac{2\pi h^2}{A^{(n)} H((\lambda^{(n)})^2)}, \tag{24}
\end{aligned}$$

where (24) follows from the fact that  $A_i^{(n)} = A^{(n)}/I_n$ .

q.e.d.

Note that a variance-stabilized version of the weighted  $K$ -function can be defined in analogy with (4), namely:

$$L_W(h) = \sqrt{\frac{K_W(h)}{\pi}}. \tag{25}$$

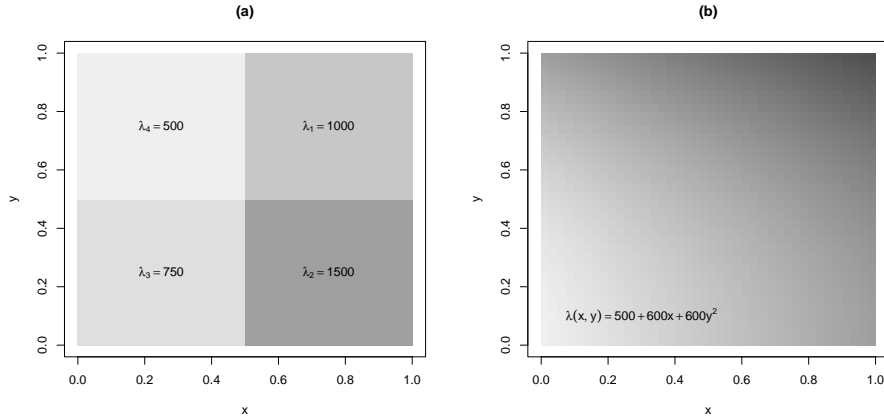


Figure 1: **Intensity models used in the simulation study.** Model (a) has a constant intensity in each quadrant, but it varies across the quadrants. Model (b) on the other hand varies smoothly following a parabolic plane defined as  $\lambda(x, y) = 500 + 600x + 600y^2$ .

### 3.4 Simulation Study Comparing $K_M(h)$ and $K_W(h)$

In this simulation study, we will compare the empirical distributions of  $K_M(h)$  and  $K_W(h)$  to the theoretical results of the previous sections (notice that the theoretical distributions are the same for  $K_W(h)$  and for  $K_M(h)$ ). This comparison will be done for two different inhomogeneous Poisson processes on the unit square. In Model (a), the intensity is constant on each of the four quadrants of the unit square, but varies across the quadrants. In Model (b), the intensity will vary smoothly according to a parabolic plane. The specification of the models is described in Figure 1.

In order to obtain empirical confidence bounds for the distributions of  $K_W(h)$  and  $K_M(h)$ , 1000 point patterns are simulated using Model (a) and similarly 1000 point patterns are simulated in accordance with Model (b). For both models,  $K_W(h)$  and  $K_M(h)$  are then computed for each of the 1000 generated data sets and the empirical 2.5 and 97.5 percentiles are used for computing the empirical confidence bounds. The calculation of  $K_M(h)$  requires multiple iterations and is based on  $T = 400$  thinnings for each of the simulated point patterns.

Figure 2 shows that the theoretical results of the theorems on the distributions of  $K_M(h)$  and  $K_W(h)$  provide good predictions of the simulated

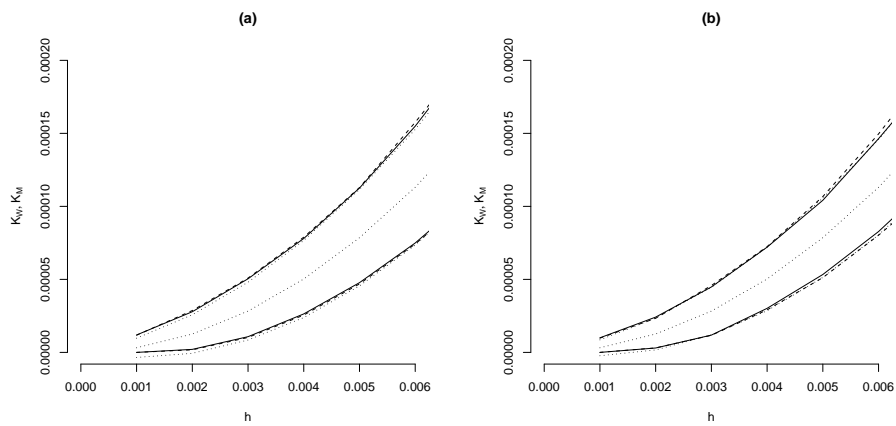


Figure 2: **Empirical 95% confidence bounds of  $K_W(h)$  and  $K_M(h)$  for small values of  $h$ .** For small values of  $h$ , the theoretical bounds are very close to the simulated bounds. The bounds based on the theorem work well for model (a) as well as for model (b) and for both  $K_M(h)$  and  $K_W(h)$ . The expectation as well as the theoretical bounds are shown as dotted lines, the simulated bounds for  $K_M(h)$  are solid and the ones for  $K_W(h)$  are dashed.

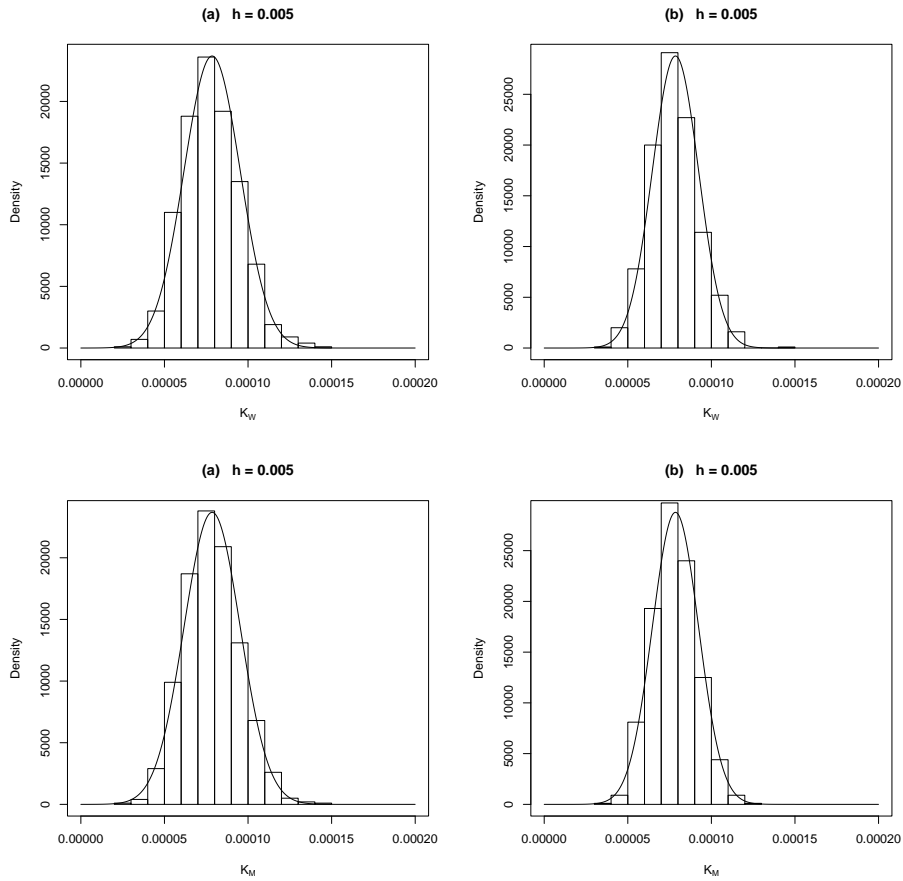


Figure 3: **Histograms of  $K_W(h)$  and  $K_M(h)$  for  $h=0.005$ .** For small values of  $h$ , the theoretical bounds are very close to the simulated bounds (this is true for both models). The solid line is the expected distribution based on the results of the theorems.



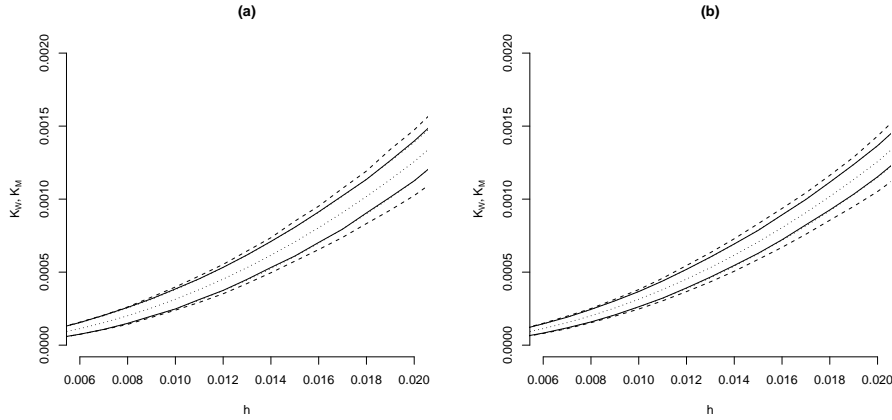


Figure 4: **Empirical 95% confidence bounds of  $K_W(h)$  and  $K_M(h)$  for larger values of  $h$ .** For larger values of  $h$ , the simulated bounds of  $K_M(h)$  are still well in line with the theoretical bounds. However, the simulated bounds of  $K_W(h)$  start to be larger than predicted theoretically, indicating that the conditions of the theorem require a small inter-point distance  $h$  (this is true for both, model (a) and (b)). The expectation as well as the theoretical bounds are shown as dotted lines, the simulated bounds for  $K_M(h)$  are solid and the ones for  $K_W(h)$  are dashed.

confidence bounds for small values of  $h$  (see also the histograms for  $h = 0.005$  in Figure 3).

The theorem for  $K_M(h)$  also holds for larger values of  $h$ , though this is not the case for  $K_W(h)$  (see Figure 4 and the histograms in Figure 5 for  $h = 0.02$ ). The theorem for  $K_W(h)$  makes the crucial assumption that the intensity be approximately constant on each disc with area  $\pi h^2$ . If the radius  $h$  is too large, the intensity will vary on each disc which violates the conditions of the theorem. While the theorem on the distribution of  $K_W(h)$  is more restrictive and is only valid for very small  $h$ , the use of  $K_W(h)$  has the practical advantage that it can be computed directly and does not require repeated sampling. Any test involving  $K_M(h)$ , on the other hand, will always include some degree of sampling variability, that will reduce the power of the test.

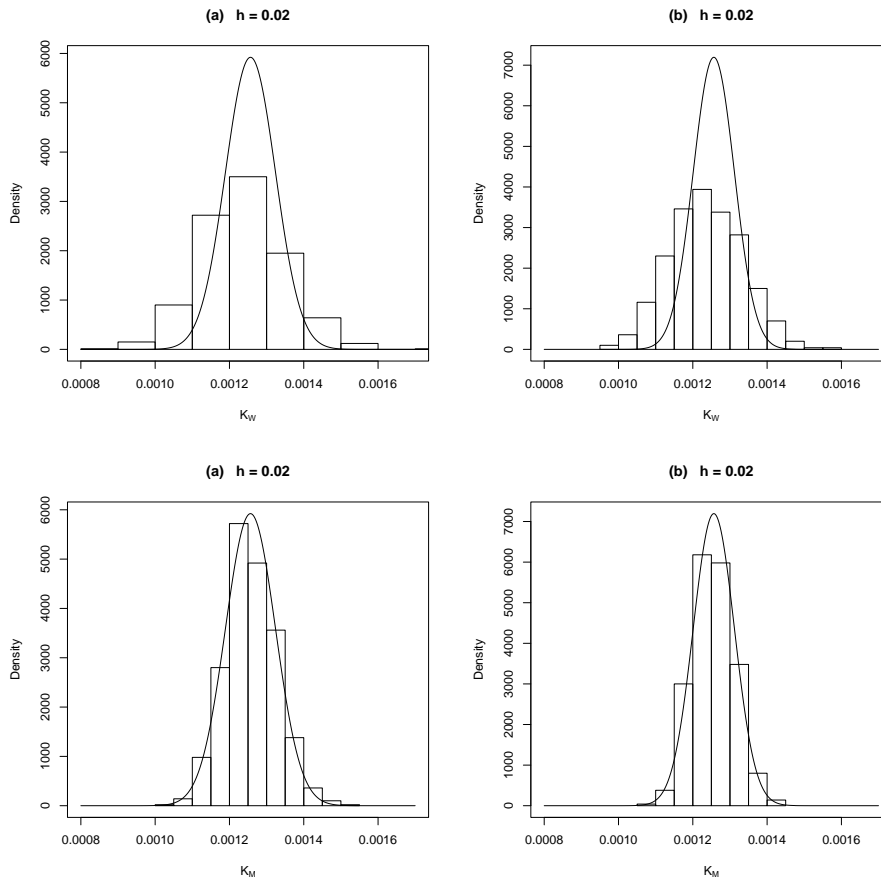


Figure 5: **Histograms of  $K_W(h)$  and  $K_M(h)$  for  $h=0.02$ .** For larger values of  $h$ , the theoretical bounds of  $K_M(h)$  are very close to the simulated bounds (again this is true for both models). However, the variance of  $K_W(h)$  is larger than predicted by the theorem (solid line).

## References

- AKAIKE, H. (1974): “A new look at statistical model identification,” *IEEE Transactions on Automatic Control*, AU-19, 716–722.
- BADDELEY, A. J., J. MØLLER, AND R. WAAGEPETERSEN (2000): “Non- and semi-parametric estimation of interaction in inhomogeneous point patterns,” *Statistica Neerlandica*, 54(3), 329–350.
- BADDELEY, A. J., AND B. W. SILVERMAN (1984): “A cautionary example on the use of second-order methods for analyzing point patterns,” *Biometrics*, 40, 1089–1093.
- BARTLETT, M. (1964): “The spectral analysis of two-dimensional point processes,” *Biometrika*, 51, 299–311.
- BESAG, J. E. (1977): “Comment on “Modelling spatial patterns” by B.D. Ripley,” *Journal of the Royal Statistical Society, Series B*, 39, 193–195.
- CRESSIE, N. A. C. (1993): *Statistics for spatial data*. Wiley, New York, revised edn.
- DALEY, D., AND D. VERE-JONES (2003): *An Introduction to the Theory of Point Processes*. Springer, New York, second edn.
- DAVIES, R. (1977): “Testing the hypothesis that a point process is Poisson,” *Advances in Applied Probability*, 9, 724–746.
- DIJKSTRA, J., T. RIETJENS, AND F. STEUTEL (1984): “A simple test for uniformity,” *Statistica Neerlandica*, 38, 33–44.
- DURRET, R. (1991): *Probability: Theory and Examples*. Wadsworth, Belmont, CA.
- FERGUSON, T. S. (1996): *A Course in Large Sample Theory*. Chapman & Hall, London.
- HEINRICH, L. (1988): “Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes,” *Statistics*, 19, 87–106.
- (1991): “Goodness-of-fit tests for the second moment function of a stationary multidimensional Poisson process,” *Statistics*, 22, 245–278.

- LAWSON, A. (1988): “On tests for spatial trend in a non-homogeneous Poisson process,” *Journal of Applied Statistics*, 15, 225–234.
- LEWIS, P. A. W., AND G. S. SHEDLER (1979): “Simulation of Nonhomogeneous Poisson Processes by Thinning,” *Naval Research Logistics Quarterly*, 26, 403–413.
- LISEK, B., AND M. LISEK (1985): “A new method for testing whether a point process is Poisson,” *Statistics*, 16, 445–450.
- MERZBACH, E., AND D. NUALART (1986): “A characterization of the spatial Poisson process and changing time,” *Annals of probability*, 14, 1380–1390.
- MEYER, P.-A. (1971): “Démonstration simplifiée d’un théorème de Knight,” *Séminaire de probabilités de Strasbourg*, 5, 191–195.
- OGATA, Y. (1981): “On Lewis’ Simulation Method for Point Processes,” *IEEE Transactions on Information Theory*, 27, 23–31.
- (1988): “Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes,” *Journal of the American Statistical Association*, 83, 9–27.
- (1998): “Space-time point-process models for earthquake occurrences,” *Annals of the Institute of Statistical Mathematics*, 50, 379–402.
- OHSER, J. (1983): “On estimators for the reduced second-moment measure of point processes,” *Mathematische Operationsforschung und Statistik, Ser. Statistics*, 14, 63–71.
- OHSER, J., AND D. STOYAN (1981): “On the second-order and orientation analysis of planar stationary point processes,” *Biometrical Journal*, 23, 523–533.
- RIPLEY, B. D. (1976): “The second-order analysis of stationary point processes,” *Journal of Applied Probability*, 13, 255–266.
- (1988): *Statistical Inference for Spatial Processes*. Cambridge University Press.
- SAW, J. G. (1975): “Tests on intensity of a Poisson process,” *Communications in Statistics*, 4 (8), 777–782.
- SCHOENBERG, F. (1999): “Transforming Spatial Point Processes into Poisson Processes,” *Stochastic Processes and their Applications*, 81, 155–164.

- SCHOENBERG, F., AND B. BOLT (2000): “Short-term exciting, long-term correcting models for earthquake catalogs,” *Bulletin of the Seismological Society of America*, 90(4), 849–858.
- SCHOENBERG, F. P. (2003): “Multi-dimensional residual analysis of point process models for earthquake occurrences,” *Journal of the American Statistical Association*, 98, 789–795.
- SCHWARTZ, G. (1979): “Estimating the dimension of a model,” *Annals of Statistics*, 6, 461–464.
- SILVERMAN, B. W. (1978): “Distances on circles, toruses and spheres,” *Journal of Applied Probability*, 15, 136–143.
- STOYAN, D., AND H. STOYAN (2000): “Improving Ratio Estimators of Second Order Point Process Characteristics,” *Scandinavian Journal of Statistics*, 27, 641–656.
- VEEN, A., AND F. P. SCHOENBERG (2005): “Assessing Spatial Point Process Models Using Weighted  $K$ -functions: Analysis of California Earthquakes,” in *Case Studies in Spatial Point Process Modeling*, ed. by A. Baddeley, P. Gregori, J. Mateu, R. Stoica, and D. Stoyan, vol. 185 of *Lecture Notes in Statistics*, pp. 293–306. Springer.