Stoch Environ Res Risk Assess manuscript No.
(will be inserted by the editor)

Some properties of local weighted second-order statistics for spatio-temporal point processes

Giada Adelfio, Marianna Siino, Jorge Mateu and Francisco J. Rodríguez-Cortés

Abstract Diagnostics of goodness-of-fit in the theory of point processes are often considered through the transformation of data into residuals as a result of a thinning or a rescaling procedure. We alternatively consider here second-order statistics coming from weighted measures. Motivated by Adelfio and Schoenberg (2009) for the temporal and spatial cases, we consider an extension to the spatio-temporal context in addition to focussing on local characteristics. In particular, our proposed method assesses goodness-of-fit of spatio-temporal models by using local weighted second-order statistics, computed after weighting the contribution of each observed point by the inverse of the conditional intensity function that identifies the process. Weighted second-order statistics directly apply to data without assuming homogeneity nor transforming the data into residuals, eliminating thus the sampling variability due to the use of a transforming procedure. We provide some characterisations and show a number of simulation studies.

Keywords K-function · Local properties · Residual analysis · Second-order characteristics · Spatio-temporal point patterns

1 Introduction

Spatial, temporal, and spatio-temporal point processes, and in particular Poisson processes, are stochastic processes that are largely used to describe and model the distribution of a plethora of real phenomena, such as, and to name just a few, seismic

Corresponding author: Giada Adelfio, E-mail: giada.adelfio@unipa.it

Giada Adelfio, Marianna Siino
Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università degli Studi di Palermo, Palermo, Italy

Jorge Mateu
Departament of Mathematics, University Jaume I, Castellón, Spain

Francisco J. Rodríguez-Cortés
Escuela de Estadística, Universidad Nacional de Colombia, Medellín, Colombia

events, epidemic diffusions, the distribution of pathologies as a function of pollution sources, traffic accidents on roads, tornado events, etc.

When a model is fitted to a set of random points, observed in a given subset of $\mathbb{R}^d$, $d \geq 1$, diagnostic measures are necessary to assess the goodness-of-fit and to evaluate the ability of that model to describe the random point pattern behaviour. Although for temporal point processes several diagnostic tools have been already introduced, for the multidimensional case the literature is quite recent and somehow scarce. The main problem when dealing with residual analysis for point processes is to find a correct definition of residuals, since the one used in dependence models can not be used for point processes.

A widely used approach considers a stationary Poisson residual process by randomly rescaling (Meyer 1971; Schoenberg 1999) or thinning (Schoenberg 2003), and investigates whether the second-order properties of the observed residuals are consistent with those of the stationary Poisson process, as in Ogata (1988).

An alternative approach is to define a weighted second-order statistic, where essentially to each observed point a weight inversely proportional to the conditional intensity at that point is given. This method was adopted by Veen and Schoenberg (2005) in constructing a weighted version of the spatial K-function of (Ripley 1977; Veen 2006), although the spatial weighted analogue of Ripley's K-function which was first introduced by Baddeley et al. (2000).

There are often two steps involved in the diagnostic of goodness-of-fit in the theory of point processes: the transformation of data into residuals (i.e. the result of a thinning or a rescaling procedure Schoenberg (2003)), and the use of second-order statistics-based tests to assess the consistency of the residuals with the homogeneous Poisson process. For instance, an estimate of the autocorrelation function of residuals could indicate the amount of dependence of data which is not described by the fitted model.

We note that Schoenberg (2003) shows how this type of rescaling and transformation for diagnostics could provide irregular boundaries, rendering difficult any comparison with a homogeneous Poisson process, considered as the benchmark model.

However, in general, assessing the consistency of observed points with homogeneous Poisson processes often constitutes only a starting stage of a more complex analysis, necessary to build and fit a more realistic process. Considering a more general point process model rather than the stationary Poisson is often more complicated, but Zhuang (2006) considered this case for second-order residuals for various general space-time branching processes such as the Epidemic-Type Aftershock Sequence (ETAS) model.

A possible answer to considering more general process models comes from considering weighted measures, and we focus here on this alternative procedure, as proposed in Adelfio and Schoenberg (2009) for the temporal and spatial cases. Motivated by this paper, we consider here an extension of their proposal to the spatio-temporal context in addition to focussing on local characteristics. In particular, our proposed method assesses goodness-of-fit of spatio-temporal models by using local weighted second-order statistics, computed after weighting the contribution of each observed point by the inverse of the conditional intensity function that identifies the process.

Weighted second-order statistics directly apply to data without assuming homogeneity nor transforming the data into residuals, eliminating thus the sampling variability due to the use of a transforming procedure. The probability used in the thinning method to retain points are replaced here by weights in order to offset the inhomogeneity of the process, with the advantage to include all the observed points rather than only the ones retained after the application of a random thinning. Moreover, this method can be applied to processes of any dimension, provided the statistics can be computed, and allows such second-order statistics to play a primary role in the diagnostic procedure, so that features such as clustering and inhibition can easily be interpreted.

More in detail, we prove in this paper that some weighted second-order statistics behave as the corresponding non-weighted ones of a homogeneous Poisson process and, therefore, these statistics can be used as a diagnostic tool for a general set of models defined on $\mathbb{R}^d$.

In addition to the weighting scheme, we deviate from the existing methods in that rather than using global characteristics, we introduce local tools, considering individual contributions of a global estimator as a measure of clustering (Anselin 1995). Generally, as explained in Anselin (1995), the individual contributions to a global statistic can be used to identify outlying components measuring the influence of each contribution to the global statistic. Getis and Franklin (1987) introduce a local version of the K-function to show that trees exhibit different heterogeneity when examined at different scales of analysis. The notion of individual functions for L and G statistics has been introduced in Stoyan and Stoyan (1994). Mateu et al. (2010) showed that the local product density function (Cressie and Collins 2001) is more sensitive to identify different local structures and unusual points than the local K-function. In Siino et al. (2017) the authors extend local indicators of spatial association to the spatio-temporal context (LISTA functions) based on the product density, and these local functions have been used to define a proper statistical test for clustering detection.

In this paper, we extend the spatial weighting approach of Veen and Schoenberg (2005) to spatio-temporal local second-order statistics, proving that the weighted second-order statistics behave as the corresponding non-weighted ones of a homogeneous Poisson process; departures suggest the unsuitability of the conditional intensity function used in the weighting scheme. As for the regression analysis the residuals are expected to be white noise if the null hypothesis model is correct, the white noise of point processes is the homogeneous Poisson process; thus, similarly, the transformation of observed point patterns generates a homogeneous Poisson process if the null hypothesis model is correct.

The plan of the paper is the following. Section 2 presents some basics of spatial and spatio-temporal point processes with an emphasis on second-order characteristics. We analyse by simulation the convergence of the estimated K-function to the normal distribution. The construction of a weighted spatial process and its second-order properties come in Section 3. Here we also show some properties of the convergence to a normal distribution of a weighted K-function. Section 4 presents our new proposal focusing on a local space-time weighted version of the K-function. We show some theoretical properties and a simulation exercise is developed in Section 5. The paper ends with some final conclusions and a discussion.

2 Point processes and second-order properties

A spatial point process is a stochastic process with realisations consisting of a finite or a countably infinite set of points in the plane. If we additionally consider the temporal occurrences for each spatially located event, we then have a spatio-temporal point process. Throughout this paper, we use N for both a spatial and a spatio-temporal point process, when no confusion arises. Point processes are introduced here by a mathematical approach that uses the definition of a counting measure on a set $X \subseteq \mathbb{R}^d, d \geq 1$, with positive values in $\mathbb{Z}_+$: for each Borel set B this $\mathbb{Z}_+$-valued random measure gives the number of events falling in B.

This section reviews some basic concepts related to point processes, reported to introduce the notation used throughout the paper. For further elaboration and references, we refer the reader to Daley and Vere-Jones (2003).

Definition 1 Point process (Cressie 1991)
Let $(\Omega, \mathscr{A}, P)$ be a probability space and Φ a collection of locally finite counting measures on $X \subset \mathbb{R}^d$. Define $\mathscr{X}$ as the Borel σ-algebra of X and let $\mathscr{N}$ be the smallest σ-algebra on Φ, generated by sets of the form $\{\phi \in \Phi : \phi(B) = n\}$ for all $B \in \mathscr{X}$. A point process N on X is a measurable mapping of $(\Omega, \mathscr{A})$ into $(\Phi, \mathscr{N})$. A point process defined over $(\Omega, \mathscr{A}, P)$ induces a probability measure $\Pi_N(Y) = P(N \in Y), \forall Y \in \mathscr{N}$.

Given a point process N defined on the space $(X, \mathscr{X})$ and a Borel set B, the number of points $N(B)$ in B is a random variable with first moment defined by

$$\mu_N(B) = E[N(B)] = \int_\Phi \phi(B)\Pi_N(d\phi)$$

that is a measure on $(X, \mathscr{X})$. The measure μ_N is called the mean measure or first moment measure of N (Cressie 1991). The second moment measure of N is given by

$$\mu_N^{(2)}(B_1 \times B_2) = E[N(B_1)N(B_k)] = \int_\Phi \phi(B_1)\phi(B_k)\Pi_N(d\phi),$$

with $B_1, B_2 \in \mathscr{X}$. If it is finite in $\mathscr{X}^{(2)}$ the process is said to be of second-order.

Let $d\boldsymbol{s}$ and $d\boldsymbol{u}$ be small regions located at $\boldsymbol{s}$ and $\boldsymbol{u} \in X$, and let $\ell(\cdot)$ be the Lebesgue measure. The first-order intensity is defined by

$$\eta(\boldsymbol{s}) = \lim_{\ell(d\boldsymbol{s}) \to 0} \frac{\mu_N(d\boldsymbol{s})}{\ell(d\boldsymbol{s})},$$

and the second-order intensity is given by

$$\eta_2(\boldsymbol{s}, \boldsymbol{u}) = \lim_{\substack{\ell(d\boldsymbol{s}) \to 0 \\ \ell(d\boldsymbol{u}) \to 0}} \frac{\mu_N^{(2)}(d\boldsymbol{s} \times d\boldsymbol{u})}{\ell(d\boldsymbol{s})\ell(d\boldsymbol{u})}.$$

Consider now that N is a point process on a spatio-temporal domain $X = \mathbb{R}^2 \times \mathbb{R}_+$, whose realisations are events in the form of $(\mathbf{z}, t)$. A fundamental tool is given by the conditional intensity function, defined as

$$\lambda(\mathbf{z}, t|\mathscr{H}_t) = \lim_{d\mathbf{z}, dt \to 0} \frac{E[N([\mathbf{z}, \mathbf{z} + d\mathbf{z}) \times [t, t + dt)|\mathscr{H}_t)]}{\ell(d\mathbf{z}dt)}, \tag{1}$$

where $\mathscr{H}_t$ is the space-time occurrence history of the process up to time t, i.e. the σ-algebra of events occurring at times up to but not including t; $dt, d\mathbf{z}$ are time and space increments, respectively, and $E[N([\mathbf{z},\mathbf{z}+d\mathbf{z})\times[t,t+dt)|\mathscr{H}_t)]$ is the history-dependent expected value of occurrence in the volume $\{[\mathbf{z},\mathbf{z}+d\mathbf{z})\times[t,t+dt)\}$. The conditional intensity function is a function of the point history and it is itself a stochastic process depending on the past up to time t. Assuming the limit in (1) exists for each point $(\mathbf{z},t)$ in the space-time domain, and that the point process is simple, then the conditional intensity uniquely characterises the finite-dimensional distributions of N (Daley and Vere-Jones 2003). If the conditional intensity function is independent of the past history, but dependent only on the current time and the spatial locations, (1) identifies an inhomogeneous Poisson process. A constant conditional intensity characterises a stationary Poisson process.

2.1 Second-order statistics for spatial and spatio-temporal point processes

The K-function (Ripley 1976) is a measure of the distribution of the inter-point distances and captures the spatial dependence of a point process. It is defined as the expected number of further events within distance δ of any given point divided by the overall rate λ, i.e.

$$K(\delta)=\lambda^{-1}E[\text{n. of extra events within distance } \delta \text{ of an arbitrary event}],\ \delta\geq 0. \quad (2)$$

Let N be a point process defined on $A\subseteq\mathbb{R}^2$ and let $\{\mathbf{z}_1,\ldots,\mathbf{z}_n\}$ a realisation of the process on A. A simple estimator of $K(\delta)$ is defined by

$$\hat{K}(\delta)=\hat{\lambda}^{-1}\sum_i\sum_{i\neq j}I(|\mathbf{z}_i-\mathbf{z}_j|\leq\delta)/n,$$

where $\hat{\lambda}$ is the estimator of the overall intensity given by $n/\ell(A)$, i.e. the observed number of events per unit area $\ell(A)$ and $I(\cdot)$ is the indicator function. Stoyan and Stoyan (2000) used $\tilde{\lambda}^2=\frac{n(n-1)}{\ell(A)^2}$ as an estimator of the squared intensity function, and thus

$$\tilde{K}(\delta)=\frac{2\ell(A)}{n(n-1)}\sum_i^n\sum_{j>i}^n I(|\mathbf{z}_i-\mathbf{z}_j|\leq\delta). \quad (3)$$

The properties of the K-function are well understood. Assume N is a homogeneous Poisson process with rate λ with values in the subset A of $\mathbb{R}^2$ with finite area $\ell(A)$, and let the distances δ between the $\binom{n}{2}$ pairs of points be small compared with $\ell(A)$. Then, as the area $\ell(A)$ tends to infinity (Cressie (1991), p.642)

$$\hat{K}(\delta)\xrightarrow{d}N\left(\pi\delta^2,\frac{2\pi\delta^2}{\lambda^2\ell(A)}\right) \quad (4)$$

For data analysis, the variance stabilised K-function, called the L-function, is generally used. The sample version of the L-function is defined as

$$\hat{L}(\delta)=\left(\hat{K}(\delta)/\pi\right)^{1/2}.$$

For approximately homogeneous data, the L-function has expected value δ and its variance is approximately constant in δ.

Moving towards local features, for a point $\mathbf{z}_i$ of the spatial point pattern N, Getis (1984) proposed

$$\hat{L}_i(\delta) = \frac{\ell(A)}{\pi(n-1)} \sum_{j \neq i}^{n} I(|\mathbf{z}_i - \mathbf{z}_j| \leq \delta) = \left(\hat{K}_i(\delta)/\pi\right)^{1/2}$$

where

$$\hat{K}_i(\delta) = \frac{1}{\hat{\lambda}^2 \ell(A)} \sum_{j \neq i}^{n} I(|\mathbf{z}_i - \mathbf{z}_j| \leq \delta) \tag{5}$$

is the local version of the K-function. Getis (1984) proved that $E(\hat{L}_i(\delta)) = \delta$.

Second-order properties are also used to analyse the spatio-temporal structure of a point process. In particular, the space-time inhomogeneous pair correlation function and K-function can be used as measures of spatio-temporal clustering/regularity and of spatio-temporal interaction (Gabriel and Diggle 2009; Møller and Ghorbani 2012).

Let $\{z_i, t_i\}_{i=1}^{n}$ be a realisation of a spatio-temporal point process observed on $A \times T$ within the spatio-temporal domain $\mathbb{R}^2 \times \mathbb{R}_+$. The simplest expression of an estimator of the spatio-temporal K-function is given by

$$\hat{K}(\delta, r) = \frac{1}{|A \times T|} \sum_{i}^{n} \sum_{j > i}^{n} I(|\mathbf{z}_i - \mathbf{z}_j| \leq \delta, |t_i - t_j| \leq r). \tag{6}$$

such that, if the process is Poisson, then (Gabriel and Diggle 2009)

$$E(\hat{K}(\delta, r)) = \pi \delta^2 r. \tag{7}$$

We can also consider the spatio-temporal version of (5), for the generic event $\mathbf{s}_i = (\mathbf{z}_i, t_i)$ as follows

$$\hat{K}_i(\delta, r) = \frac{1}{\hat{\lambda}^2 \ell(A) T} \sum_{j \neq i}^{n} I_{ij}(\delta, r) \tag{8}$$

with $I_{ij}(\delta, r) = I(|\mathbf{z}_i - \mathbf{z}_j| \leq \delta, |t_i - t_j| < r), r \in [0, T]$. If the process is Poisson it is easy to prove that, taking conditional expectation with respect to $\mathscr{H}_j$,

$$E\left(\int_{\mathbb{R}^2 \times \mathbb{R}} I_{ij}(\delta, r) dN_j\right) = E \int_{\mathbb{R}^2 \times \mathbb{R}} I_{ij}(\delta, r) \lambda d\mathbf{s}_j = \lambda^2 \delta^2 r \pi \ell(A) T$$

and therefore,

$$E(\hat{K}_i(\delta, r)) = \pi \delta^2 r. \tag{9}$$

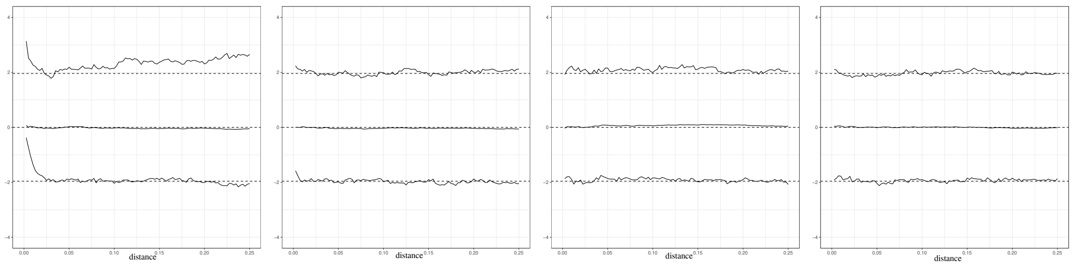
In Section 4, a weighted version of equation (8) is proposed, in order to use the local K-function as a diagnostic tool for finding local lack of fitting of the estimated intensity function.

2.1.1 Some experiments to assess the normality condition in (4)

Under the homogeneity of the Poisson process, the converge theorem in (4) occurs tending the area of the spatial region to infinity ($\ell(A) \to \infty$). In this section we show a small simulation study to investigate for which range of distances and window dimensions the asymptotic normal distribution for $\hat{K}(\delta)$ is reliable. We set the first-order intensity function as $\lambda = 100$, and $A = [0,a] \times [0,a]$ is the spatial squared region, with a taking the values $a = \{1,5,10,20\}$; thus the corresponding expected number of points per each a is $n = \{100, 2500, 10000, 40000\}$.

Given λ, and for each value of a, 100 random point patterns are generated from a homogeneous Poisson process. For each point pattern the K-function is computed using Ripley's isotropic edge-correction over a vector of distances ranging from 0 to 0.25. Each estimated curve is standardised with respect to its mean ($\pi\delta^2$) and asymptotic estimated variance ($2\pi\delta^2/\tilde{\lambda}^2\ell(A)$) (see eq. (4)), where $\tilde{\lambda}^2 = n(n-1)/l(A)^2$.

In Figure 1, the mean, and the 2.75% and 97.5% percentiles for the 100 standardised estimated K-functions of point patterns generated from a homogeneous Poisson process with intensity $\lambda = 100$ in the squared window $A = [0,a] \times [0,a]$ are shown. As expected, the standardised curves follow a standard normal distribution for most of the considered distances. It is clear that the approximation to the normal distribution is better when increasing the window size. In addition, for a selected set of distances $\delta \in \{0.005, 0.01, 0.02, 0.05, 0.10, 0.15, 0.2, 0.25\}$ the Shapiro-Wilks test is computed to check the normality assumption (see Table 1 and Figure 2). When $a = 1$ the normality assumption is not accepted for all the selected distances, see the histograms in Figures 2a and 2e. However, increasing the window size, and so the number of points in A, the normality clearly holds.

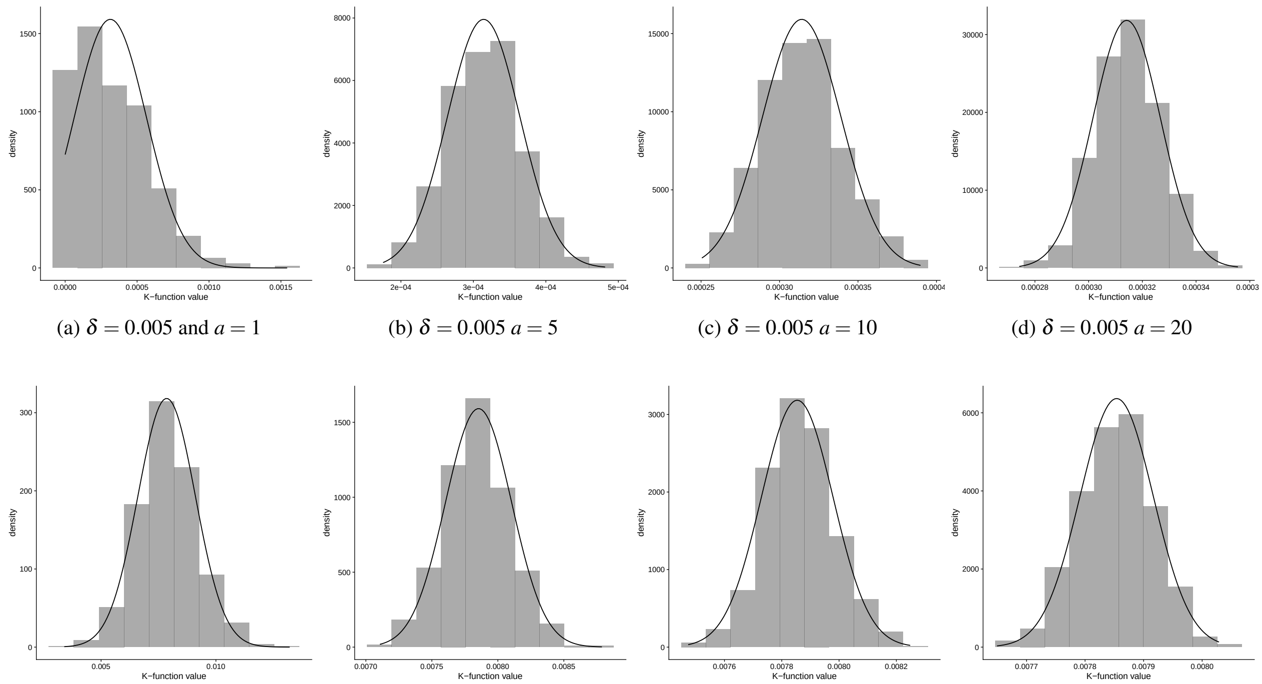


(a) $a = 1$ (b) $a = 5$ (c) $a = 10$ (d) $a = 20$

Fig. 1: Mean, 2.75% and 97.5% percentiles for the 100 standardised estimated K-functions of point patterns generated from a homogeneous Poisson process with intensity $\lambda = 100$ in the squared window $A = [0,a] \times [0,a]$. The horizontal dotted lines indicate the theoretical 2.75% percentile, mean and 97.5% percentile of the standard normal distribution.

Table 1: p-values of the Shapiro-Wilks test to check the normality distribution of the estimated value of the K-function for a fixed distance, δ. The K-function is computed over 100 simulated point patterns from a homogeneous Poisson process with intensity $\lambda = 100$ and area of the squared window a^2.

	a			
δ	1	5	10	20
0.005	0.000	0.000	0.054	0.068
0.01	0.000	0.487	0.000	0.984
0.02	0.000	0.072	0.574	0.568
0.05	0.069	0.635	0.380	0.345
0.1	0.004	0.450	0.253	0.851
0.15	0.000	0.301	0.379	0.143
0.2	0.000	0.072	0.230	0.381
0.25	0.000	0.122	0.075	0.729



(a) $\delta = 0.005$ and $a = 1$ (b) $\delta = 0.005$ $a = 5$ (c) $\delta = 0.005$ $a = 10$ (d) $\delta = 0.005$ $a = 20$

(e) $\delta = 0.01$ $a = 1$ (f) $\delta = 0.01$ $a = 5$ (g) $\delta = 0.01$ $a = 10$ (h) $\delta = 0.01$ $a = 20$

Fig. 2: Histograms of the sample distribution of the K-function from the simulated point patterns in Figure 1 and Table 1, for a fixed distance δ and window size $A = [0,a] \times [0,a]$.

3 The weighted spatial process and its second-order properties

Let N be a point process defined on $X \subseteq \mathbb{R}^d, d \geq 1$. For any point $\mathbf{s}$ in X, let $\lambda(\mathbf{s}|\mathscr{F})$ be the conditional intensity function of the process with respect to some filtration $\mathscr{F}$ on X, for simplicity denoted by $\lambda(\mathbf{s})$. Then

$$\lambda(\mathbf{s})\delta \approx E[N(\mathbf{s},\mathbf{s}+\delta)|\mathscr{F}].$$

Since $\lambda(\mathbf{s})$, which is related to the probability that a point occurs near $\mathbf{s}$, depends on the information on which the conditioning is based, it is a random process; we assume that $\lambda(\mathbf{s})$ is positive and bounded away from zero.

Let N_w be a real-valued random measure such that for any set $S \subseteq X$, $N_w(S) = \int_S \frac{1}{\lambda_*(\mathbf{s})} dN$ holds, with $\frac{1}{\lambda_*(\mathbf{s})} = \frac{\lambda_{inf}}{\lambda(\mathbf{s})}$, and assuming the existence of the positive constant $\lambda_{inf} \leq \inf\{\lambda(\mathbf{s}); \mathbf{s} \in S\}$.

Adelfio and Schoenberg (2009) formally show that the main second-order statistics of $N_w(\cdot)$ behave similarly to those of a homogeneous Poisson process. Here, we first recall some results relative to the main features of N_w that will be used in proofs of subsequent results.

Rewriting the weighted process as $N_w(S) = \frac{N(S)}{\lambda_*(\mathbf{s})}$, following results in Zhuang (2006) and Adelfio and Schoenberg (2009), we have that

$$\lim_{\delta\to 0} \frac{E[N_w(\mathbf{s}_i,\mathbf{s}_i+\delta)]}{\delta} = \lambda_{inf},$$

and the mixed moment takes the form

$$\lim_{\delta\to 0} \frac{E[N_w(\mathbf{s}_i,\mathbf{s}_i+\delta)N_w(\mathbf{s}_j,\mathbf{s}_j+\delta)]}{\delta^2} = \lambda_{inf}^2.$$

Thus the covariance becomes

$$\text{cov}[N_w(\mathbf{s}_i,\mathbf{s}_i+\delta), N_w(\mathbf{s}_j,\mathbf{s}_j+\delta)] = 0 \tag{10}$$

3.1 The spatial weighted K-function

Consider now that N is defined on A, a subregion of $\mathbb{R}^2$ with area $\ell(A)$, and let $I(\cdot)$ be the indicator function.

Let $\lambda_0(\mathbf{z})$ be the *true* intensity function of the model that describes the observed process on A, with $\mathbf{z} \in A$, here denoted as the null hypothesis. Then, the estimator of the weighted K-function (Baddeley et al. 2000) is defined as

$$\hat{K}^w(\delta) = \frac{1}{\lambda_{\text{inf}}^2 \ell(A)} \sum_i^n \omega_i \sum_{j\neq i}^n \omega_j I(|\mathbf{z}_i - \mathbf{z}_j| \leq \delta),$$

where $\lambda_{\text{inf}} = \inf\{\lambda_0(\mathbf{z}); \mathbf{z} \in A\}$ is the minimum of the conditional intensity over the observed region under the null hypothesis and, for each k, $\omega_k = \lambda_{\text{inf}}/\lambda_0(\mathbf{z}_k)$, with $\lambda_0(\mathbf{z}_k)$ the conditional intensity at the point $\mathbf{z}_k$ of A under H_0 and $\delta > 0$.

Combining the point process residual analysis techniques and the use of the K-function as a diagnostic tool applied to residual processes, Veen and Schoenberg (2005) provided theorems on the distributional properties of the K-function and its weighted variant, assuming homogeneous and inhomogeneous Poisson processes.

Formally, let ${}^{(m)}N$, $m = 1, 2, \ldots, M$ be a sequence of inhomogeneous Poisson processes with intensities ${}^{(m)}\lambda$ and weighted K-functions ${}^{(m)}K^w(\delta)$, defined on the subsets ${}^{(m)}A \in \mathbb{R}^2$ of areas ${}^{(m)}\ell(A)$. Using a central limit theorem, and for $m \to \infty$, Veen and Schoenberg (2005) provided the following convergence result for the sequence of the estimators of the weighted K-function:

$$ {}^{(m)}\hat{K}^w(\delta) \xrightarrow{d} N\left(\pi\delta^2, \frac{2\pi\delta^2}{{}^{(m)}\ell(A)H(({}^{(m)}\lambda)^2)}\right), \tag{11} $$

where $H(({}^{(m)}\lambda)^2)$ is the harmonic mean of the squared intensity in the observed region ${}^{(m)}A$.

3.1.1 Some experiments to assess the normality condition in (11)

As stated above, given an inhomogeneous Poisson process, the converge theorem in (11) holds asymptotically when the observed region is broken up into a disjoint number of subregions ${}^{(m)}A \to +\infty$ with sufficiently regular boundaries such that the pairs of points $(\mathbf{z}_i, \mathbf{z}_j)$ with $|\mathbf{z}_i - \mathbf{z}_j| < h_n$ are in distinct subregions. We report a small simulation experiment to figure out for which range of distances and number of subregions the asymptotic normal distribution for $\hat{K}^w(\delta)$ holds.

We consider the following inhomogeneous first-order intensity function $\lambda(\mathbf{z}) = \exp(\beta_0 + \beta_1 \mathbf{z})$, where $\beta_0 = \{5, 7, 8\}$ and $\beta_1 = 3$ in the unit square spatial window. Changing the value of β_0 the number of points in the spatial window increases. Given $\lambda(\mathbf{z})$ for each value of β_0, 100 random point patterns are generated from the corresponding inhomogeneous Poisson process. For each point pattern the weighed K-function is computed using Ripley's isotropic edge-correction, with weights based on the estimated intensity, and over a range of distances from 0 to 0.015. Considering m the number of subregions in which the region A is divided, we explored the cases $m = \{2^2, 4^2, 6^2, 10^2, 100^2\}$. For a selected set of distances $\delta \in \{0.005, 0.008, 0.01, 0.014\}$ the Shapiro-Wilks test is computed to check the normality assumption, and the results are reported in Table 2. As expected, it is clear that as m increases the normality distribution in (11) holds for any distance δ. The effect of parameter β_0 is evident. We need a larger number of subregions, m, to get normality when β_0 increases, i.e. when the number of points increases. Finally, small distances favour normality at smaller m. Thus, the combination of parameters m, δ and the form of the intensity that controls the number of points is crucial in understanding the convergence theorem in (11).

		β_0		
δ	m	5	7	8
0.005	2	0	0	0
	4	0.096	0	0
	6	0.205	0.004	0
	10	0.301	0.655	0.011
	100	0.289	0.444	0.439
0.008	2	0	0	0
	4	0.014	0	0
	6	0.103	0	0

	10	0.254	0.436	0
	100	0.307	0.877	0.759
0.01	2	0	0	0
	4	0.001	0	0
	6	0.355	0	0
	10	0.417	0.005	0
	100	0.433	0.445	0.72
0.014	2	0	0	0
	4	0	0	0
	6	0.033	0	0
	10	0.144	0.041	0
	100	0.317	0.156	0.538

Table 2: p-values of the Shapiro-Wilks test to check the normality distribution of the estimated value of the inhomogeneous K-function for a fixed distance, δ. The K-function is computed using 100 point patterns simulated from an inhomogeneous Poisson process with intensity equals to $\lambda(\mathbf{z}) = exp(\beta_0 + \beta_1 \mathbf{z})$. The weights in K^w are obtained estimating $\hat{\lambda}$ from the data. m^2 indicates the number of subregions in which the window is divided.

4 The spatio-temporal local weighted K-function: the new proposal

Let N be a point process on a spatio-temporal domain $X = \mathbb{R}^2 \times \mathbb{R}_+$, whose realisations are events in the form of $\mathbf{s} = (\mathbf{z}, t)$. The weighted version of the spatio-temporal K-function, as defined in (6), is given by (Gabriel and Diggle 2009)

$$\hat{K}^w(\delta, r) = \frac{1}{\lambda_{\inf}^2 \ell(A) T} \sum_i \sum_{j \geq i}^n \frac{I_{ij}(\delta, r)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)} \tag{12}$$

with expectation

$$\begin{aligned} &E\left(\int_{\mathbb{R}^2 \times \mathbb{R}_+} \frac{I_{ij}(\delta, r)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)} dN(\mathbf{s}_i) dN(\mathbf{s}_j)\right) \\ &\text{taking conditional expectations on } \mathscr{H}_i \text{ and on } \mathscr{H}_j \\ &= E\left(\int_{\mathbb{R}^2 \times \mathbb{R}_+} \frac{I_{ij}(\delta, r)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)} \lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j) d\ell(\mathbf{s}_i) d\ell(\mathbf{s}_j)\right) \\ &= E\left(\int_{\mathbb{R}^2 \times \mathbb{R}_+} 1_{\{0 < |\mathbf{s}_i - \mathbf{s}_j| < (\delta, r)\}} d\ell(\mathbf{s}_i) d\ell(\mathbf{s}_j)\right) \\ &= (\ell(A) \times T)\pi r \delta^2, \end{aligned} \tag{13}$$

where $\ell(\cdot)$ denotes the Lebesgue measure. Therefore,

$$E\left[\hat{K}^w(\delta, r)\right] = \pi \delta^2 r \tag{14}$$

that is the same as the expectation of $\hat{K}(\delta,r)$ in (7), when the intensity used for the weighting is the true generator model.

The weighted version of the estimator of the local spatio-temporal K-function defined in (8) is given by

$$\hat{K}_i^w(\delta,r) = \frac{1}{\lambda_{\text{inf}}^2 \ell(A)T} \sum_{j\neq i}^{n} \frac{I_{ij}(\delta,r)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)}. \tag{15}$$

We note that the expectation of $\hat{K}_i^w(\delta,r)$ is the same as the expectation of $\hat{K}_i(\delta,r)$ in (8), when the intensity used for the weighting is the true generator model. Indeed, taking the conditional expectation with respect to $\mathscr{H}_j$

$$\begin{aligned}\lambda_{inf}^2 E\left(\int_{\mathbb{R}^2\times\mathbb{R}_+} \frac{I_{ij}(\delta,r)}{\lambda(\boldsymbol{s}_i)\lambda(\boldsymbol{s}_j)} dN_j\right) &= \frac{\lambda_{inf}^2}{\lambda(\boldsymbol{s}_i)} E \int_{\mathbb{R}^2\times\mathbb{R}_+} \frac{1}{\lambda(\boldsymbol{s}_j)} I_{ij}(\delta,r)\lambda(\boldsymbol{s}_j) d\boldsymbol{s}_j \\ &= \frac{\lambda_{inf}^2}{\lambda(\boldsymbol{s}_i)} \delta^2 r\pi\ell(A)T\lambda(\boldsymbol{s}_i),\end{aligned}$$

since varying the index j while keeping fixed i, the integral depends on the intensity of points around the point $\mathbf{s}_i$. Therefore, after some algebra, the expectation of (15) is given by

$$E(\hat{K}_i^w(\delta,r)) = \pi\delta^2 r, \tag{16}$$

which equals that of the non-weighted version in (9).

This result shows that the local spatio-temporal K-function estimator for a general point process, weighted by the *true* intensity function, has the same expectation of the local spatio-temporal K-function under the Poisson case. Therefore, we can use this weighted estimator as a diagnostic tool for general space-time point processes. Indeed, this result develops a new approach for diagnostics in a local scale, useful also for general (not necessary Poisson) point processes with a generic conditional intensity function $\lambda(\cdot)$, since the local weighted K-function behaves as the corresponding non-weighted one of a homogeneous Poisson process. Thus departures from this expected value directly suggest the unsuitability of the conditional intensity function used in the weighting scheme. In other words, and in an inferential context, if the estimated intensity function used for weighting in (15) is the *true* one, then the local weighted K-function should behave as the corresponding non-weighted function of a homogeneous Poisson process, that corresponds to the reference model.

For completeness, we show in the next section several results on the asymptotic equivalence of distributions of these residual and weighted estimates using martingale techniques.

4.1 Martingale characterisation for the local weighted K-function

A key theorem for the convergence of martingales, rephrased from Hall and Heyde (1980), is here introduced to further prove some converge features of the proposed statistic in equation (15).

Theorem 1 ***Central limit theorem for martingales***
If the martingale difference stochastic process $\{X_n, \mathscr{H}_n\}_{n=1}^{\infty}$, with $\mathscr{H}_n = \sigma(X_1, X_2, \ldots, X_n)$ and $E[X_i|\mathscr{H}_{i-1}] = 0,\ i = 2, 3, \ldots$, satisfies the following conditions:

1. *Lindeberg condition: $E[(X_i)^2] < \infty,\ i = 1, 2, \ldots$ such that for any $\varepsilon > 0$*
$$\lim_{n\to\infty}\left(\frac{1}{S_n^2}\sum_{i=1}^{n} E\left[X_i^2 I_{\{|X_i|>\varepsilon S_n\}}\right]\right) = 0,$$
with $S_n^2 = var[\sum_{i=1}^{n} X_i] \to \infty$ as $n \to \infty$.
2. *$E[E[X_i^2|\mathscr{H}_{i-1}]] = \sigma_i^2,\ i = 2, 3, \ldots$*

Then $\frac{1}{S_n}\sum_{i=1}^{n} X_i \xrightarrow{d} N(0,1)$.

Let us now define, for any i, $I_{ij}^{w}(\delta, r) = \frac{I_{ij}(\delta,r)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)}$ and $\int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_i} dN_j$ as the number of pairs of points with elements occurring both up to τ_i, that is the last point less than or equal to t_i such that, focussing on the time dimension, no points are in $(\tau_i, \tau_i + r)$, i.e., $\tau_i = \sup\{\tau : N(\tau) = 1, N(\tau, \tau + r) = 0, \tau + \delta < t_i\}$, letting $\tau_i = 0$ if no such τ exists. In other words, τ_i is the left end-point of the last gap prior to t_i of size at least r.

Define the process $Z(\mathbf{s}_i) = \int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_i} dN(j) - \lambda_{inf}^2 \pi\delta^2 r\ell(A)\tau_i$, for any i, assuming that such kind of gaps exists. The existence of a gap of size r corresponding to each t_i guarantees that, conditioning on $Z(\mathbf{s}_{i-1})$, no pair of points within distance r crosses $\mathbf{s}_{i-1}$ and therefore the knowledge about the past up to any point $\mathbf{s}_{i-1}$ does not give any information about $\mathbf{s}_i$. Note that for each i, τ_i is an $\mathscr{H}_{t_i}$-stopping time, and thus $Z(\mathbf{s}_i)$ is measurable (see e.g. Corollary A3.4.VIII on p.430 of Daley and Vere-Jones (2003)).

Theorem 2 ***Martingale characterisation***
Let N be a spatio-temporal point process with conditional intensity function $\lambda(\boldsymbol{s}|\mathscr{H})$, positive and bounded away from zero. Let $Z(\boldsymbol{s}_i) = \int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_i} dN_j - \lambda_{inf}^2 \pi\delta^2 r\ell(A)\tau_i$ for any $i = 1, \ldots, n$, $\delta, r > 0$, with $\int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_i} dN_j$ the number of pairs of points with elements occurring both up to τ_i and $\tau_i = \sup\{\tau : N(\tau) = 1, N(\tau, \tau + r) = 0, \tau + r < t_i\}$.

Then $Z(\boldsymbol{s}_i)$ is a martingale with respect to a filtration $\mathscr{H}_i$, i.e.,
$$E[Z(\boldsymbol{s}_i) - Z(\boldsymbol{s}_{i-1})|Z(\boldsymbol{s}_{i-1})] = 0. \tag{17}$$

Proof
$$\begin{aligned}
&E[Z(\boldsymbol{s}_i) - Z(\boldsymbol{s}_{i-1})|Z(\boldsymbol{s}_{i-1})]\\
&= E\Bigg[\left(\int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_i} dN_j - \int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{\tau_{i-1}} dN_j\right)\\
&- \lambda_{inf}^2 \pi\delta^2 r\,(\tau_i - \tau_{i-1})\,\ell(A)\Big|Z(\boldsymbol{s}_{i-1})\Bigg]\\
&= E\left[\int_{\mathbb{R}^2\times\mathbb{R}_+} I_{ij}^{w}(\delta, r)|_{(\tau_i-\tau_{i-1})} dN_j - \lambda_{inf}^2 \pi\delta^2 r\ell(A)\,(\tau_i - \tau_{i-1})\,\Big|Z(\boldsymbol{s}_{i-1})\right]\\
&= 0.
\end{aligned}$$

Therefore, starting form the martingale characterisation of its components, and extending results of Theorem 1 as in Adelfio and Schoenberg (2009), it is possible to state the asymptotic normal distribution of $\hat{K}_i^w(\delta, r)$, according to the following corollary.

Corollary 1 *Let N be a spatio-temporal point process in $\mathbb{R}^2 \times \mathbb{R}_+$ with conditional intensity function $\lambda(\mathbf{s}|\mathscr{H})$, positive and bounded away from zero, such that $E[N^2(d\mathbf{s}_i)|\mathscr{H}_i]$ $\forall i$ is bounded, and there exists an $\alpha > 0$ such that $E[N^{2+\alpha}(d\mathbf{s}_i)|\mathscr{H}_i]$ $\forall i$ is also bounded. Moreover, considering the temporal dimension, assume that for any i there exists $\tau_i = \sup\{\tau : N(\tau) = 1, N(\tau, \tau + r) = 0, \tau + r < t_i\}$, $r > 0$.*

Then the local weighted spatio-temporal K-function defined as in (15) is asymptotically normally distributed.

5 Model checking: simulation studies

The local weighted K-function is obtained by weighting the process by the inverse of the conditional intensity function of the generator model of the data. Since in simulations the weights are obtained by considering the real intensity function, the weighted statistics are expected to behave as the ones of a homogeneous Poisson process. Indeed, if departures from a such behaviour were observed, then the data would be supposed to come from a model identified by a conditional intensity function different from the one used in the weighting procedure. We consider three different scenarios, in which the K-function is weighted by (a) the real intensity function, known in simulations, (b) the estimated intensity by maximum likelihood, that is the intensity function with MLE estimates of unknown parameters, and (c) the intensity function estimated for a homogeneous Poisson process (wrong model). For each one of these scenarios, we consider both global (Section 5.1) and local (Section 5.2) spatio-temporal weighted K-functions. In Section 5.3 we consider an application to real data.

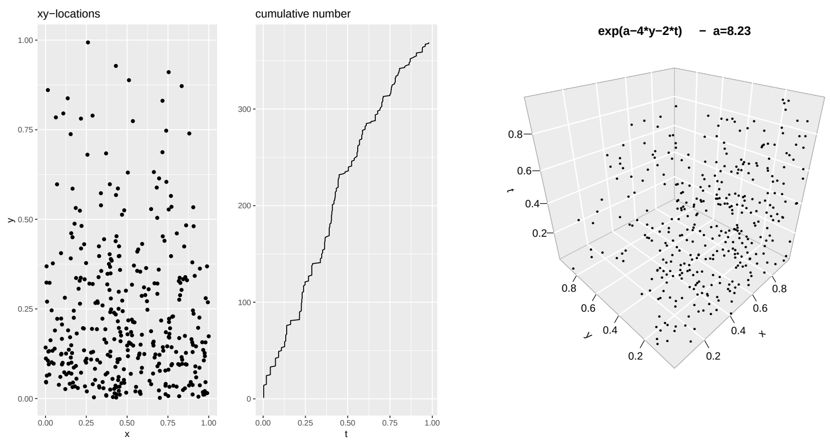
5.1 Global space-time diagnostics

Consider first a spatio-temporal inhomogeneous Poisson process with intensity function

$$\lambda(x, y, t) = \exp(a - 4y - 2t)$$

with $a = 8.25$. Figure 3 shows representations of the points in space, space-time and the accumulative number of points in time. Figure 4 reports the difference between the observed values of the global weighted spatio-temporal K-function $\hat{K}^w(\delta, r)$ reported in eq. (12) and the expected ones in eq. (14), weighting by, respectively, the true intensity (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right). We note that the χ^2 statistic, computed over a fixed grid that covers the region of interest, evaluates this distance. The same interpretation is possible for the corresponding χ^2 statistic values reported in subsection 5.2. We obtain $\chi^2 = 0.617$ (under the true intensity), $\chi^2 = 0.614$ (under the MLE intensity), and $\chi^2 = 33.654$ (under the constant intensity). It is evident that this difference is significantly larger in the latter case corresponding to the wrong model. This indicates that when weighting the spatio-temporal K-function by the true or estimated intensity its expected value resembles to that of a Poisson process, but if we use a wrong model for the intensity, the expected value is far from the Poisson one.

By looking at Figure 4, it is interesting to note that the weighted K-function obtained when the true model is used for weighting and the one computed considering the estimated intensity function are very similar, suggesting that the effect on the second-order statistics of estimating the true model is quite negligible. Similar conclusions can be reported from Figures 5, 7 and 8 that justify the practical use of these methods, when the true model is estimated rather than known.



(a) (b)

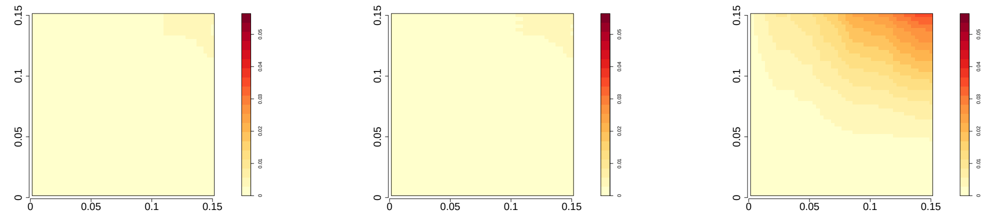
Fig. 3: A spatio-temporal inhomogeneous Poisson process with intensity function $\lambda(x,y,t) = \exp(a-4y-2t)$ with $a = 8.25$.

Repeating the above procedure for 100 simulations of a spatio-temporal inhomogeneous Poisson process with the previous intensity function, we can evaluate the distribution of the χ^2 statistic based on the differences between the weighted K-functions and the theoretical expected values. The results are shown in Figure 5, and highlight the bad result when in the weighting scheme the wrong homogeneous Poisson intensity is used.

To provide more precise information on the errors associated to the test, we have completed a simulation study for the inhomogeneous Poisson case using a global statistic. In particular, we propose the following hypothesis test:

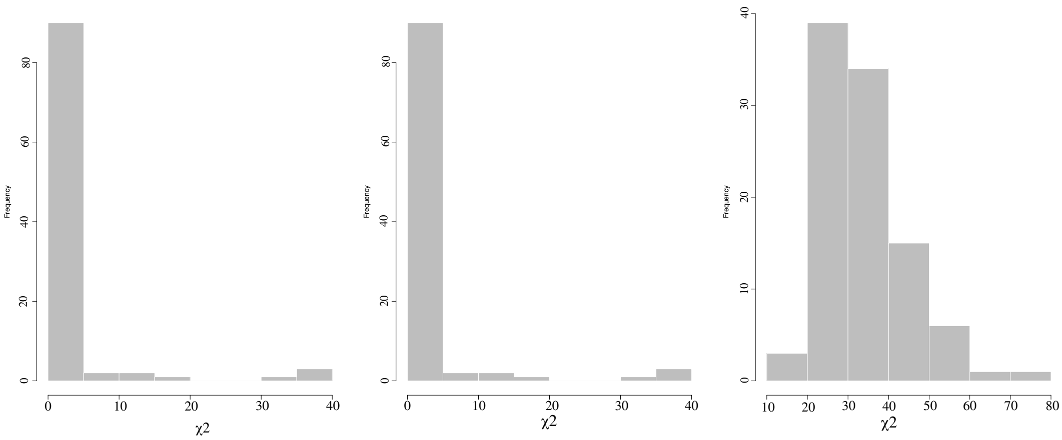
$$\begin{cases} H_0 : \lambda(\cdot) = \lambda_w(\cdot) \\ H_1 : \lambda(\cdot) \neq \lambda_w(\cdot) \end{cases}$$

with $\lambda(\cdot)$ the conditional intensity function that characterises the generating process, and $\lambda_w(\cdot)$ the conditional intensity function used in the weighted scheme. As described above, the χ^2 statistic, measuring the distance between observed and expected values



(a) (b) (c)

Fig. 4: Difference between the global weighted spatio-temporal K-function (see eq. 12) and the expected one (see eq. 14) weighting by, respectively, the true intensity (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right).



(a) (b) (c)

Fig. 5: Distribution of the χ^2 statistic computed for assessing the differences between the weighted K-functions and the theoretical expected value, calculated from 100 simulations of a spatio-temporal inhomogeneous Poisson process weighting by, respectively, the true intensity (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right).

of the local K-functions, is used to evaluate this hypothesis test. The test statistic is computed over a fixed grid that covers the region of interest. Figure 5 shows the distribution of the test statistic under the null hypothesis (Figures 5a,b) and under the alternative (Figure 5c) for the Poisson case. Therefore, to obtain significance levels under the null hypothesis and power values under the alternative we run 50 replicates of 100 simulations each for an inhomogeneous Poisson process under the null and alternative hypothesis (see Table 3). We note that the statistical test reports quite exact

Table 3: Significance levels and power of the test based on the χ^2 statistic computed for a global statistic, under the null (columns (a) and (b)) and the alternative hypothesis (column (c)) for an inhomogeneous Poisson process with intensity function defined in Sect 5.1

(a)	(b)	(c)
0.054	0.053	0.945

significance levels (very close to 0.05) and a power almost 0.95, confirming that we can use our testing procedure as a goodness-of-fit analysis for spatio-temporal point processes. It is also interesting to note that columns (a) and (b) report basically the same value, which confirms what we have noticed and reported in the comments of Fig. 4; the estimation of the true model seems to have a negligible effect on the second order statistics, since the weighted K-function obtained when the true model is used for weighting and the one computed considering the estimated intensity function have very similar behaviour.

Note that the same study could be done for the other cases (Poisson local, ETAS global/local) analysed thereafter in this subsection and in Subsection 5.2, but we have only reported the Poisson case to save space, and also because these are quite time-consuming computations.

For completeness, we also consider a more complex model for the data. In particular, we simulate point patterns from a general ETAS model (Ogata 1988; Adelfio and Chiodi 2015a), with conditional intensity function

$$\lambda_{\boldsymbol{\theta}}(\mathbf{z},t|\mathscr{H}_t) = \mu f(\mathbf{z}) + \sum_{t_j<t} \frac{k_0 \, \exp\left(\alpha(m_j - m_0)\right)}{(t-t_j+c)^p} \left\{(\mathbf{z}-\mathbf{z}_j)^2 + d\right\}^{-q} \tag{18}$$

where $f(\mathbf{z})$ is the spatial density of the background/long term component, stationary in time. The aftershock induced component is the product of the density of aftershocks in time, i.e. the Omori law representing the occurrence rate of aftershocks at time t, following the earthquake of time t_j and magnitude m_j, and the density of aftershocks in space. In particular, m_j is the magnitude of the j-th event and m_0 the threshold magnitude, i.e. the lower bound for which earthquakes with higher values of magnitude are surely recorded in the catalogue. The rest of elements of (18) are as follows: α measures the influence on the relative weight of each sequence, k_0 is a normalising constant, c and p are characteristic parameters of the seismic activity of the given region (p is useful for characterising the pattern of seismicity, indicating the decay rate of aftershocks in time), and finally d and q are two parameters related to the spatial influence of the mainshock.

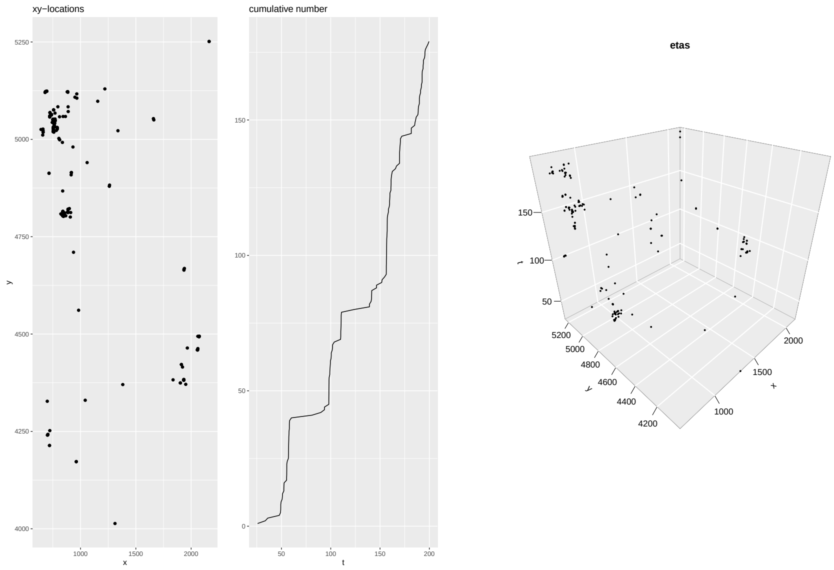
The induced spatial distribution conditioned to the magnitude of the generating event is of the form

$$\ell(\mathbf{z}-\mathbf{z}_j|m_j) = \left\{\frac{(\mathbf{z}-\mathbf{z}_j)^2}{e^{\gamma\,(m_j-m_0)}} + d\right\}^{-q} \tag{19}$$

relating the occurrence rate of aftershocks to the mainshock magnitude m_j, through the parameter γ that, as α, measures the influence on the relative weight of each sequence. In this paper we assume, without loss of generality, that $\gamma = 0$ (see Chiodi and Adelfio (2017) for a more complete formulation of the ETAS model).

The ML estimation of the background and the triggered intensity components of an Epidemic type model has been performed by the FLP approach (Chiodi and Adelfio 2011, 2017), which is a nonparametric estimation procedure based on the subsequent increments of the log-likelihood obtained adding one observation at a time, to account for the information and dependence of the previous observations on the next one.

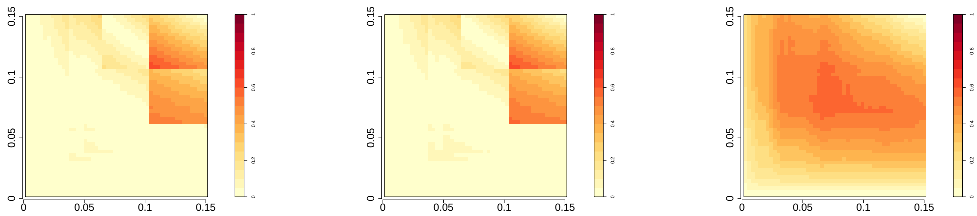
Figure 6 reports a simulated spatio-temporal point pattern using an ETAS model with parameters $\boldsymbol{\theta} = (\mu, k_0, c, p, \alpha, d, q)$ =$(0.079, 0.004, 0.013, 1.2, 0.5, 0.424, 1.165)$. This figure shows representations of the points in space, space-time and the accumulative number of points in time.



(a) (b)

Fig. 6: Simulated spatio-temporal point pattern using an ETAS model with parameters $\boldsymbol{\theta} = (\mu, k_0, c, p, \alpha, d, q)$ =$(0.079, 0.004, 0.013, 1.2, 0.5, 0.424, 1.165)$.

Figure 7 reports the difference between the global weighted spatio-temporal K-function (see eq. 12) and the expected one (see eq. 14) weighting by, respectively, the true intensity from the ETAS model (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right). Again, the difference was evaluated by a χ^2 statistic to provide some more evidence, obtaining $\chi^2 = 68.14$ (under the true intensity), $\chi^2 = 73.27$ (under the MLE intensity), and $\chi^2 = 586.38$ (under the constant intensity). It is evident that this difference is significantly larger in the latter case corresponding to the wrong model. This indicates that when weighting the spatio-temporal K-function by the true or estimated intensity its expected value resembles to that of a Poisson process, but if we use a wrong model for the intensity, the expected value is far from the Poisson one. It is important to note that this result stands the same independently of the process governing the point pattern.



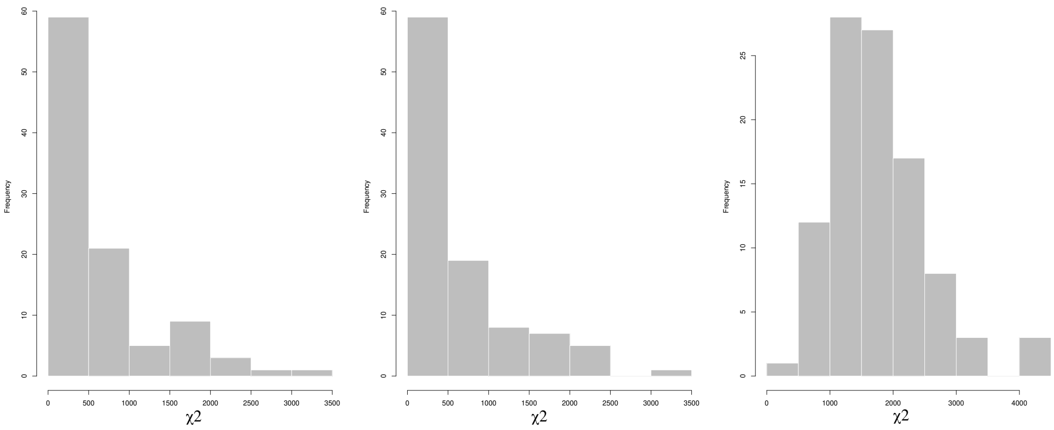
(a) (b) (c)

Fig. 7: Difference between the global weighted spatio-temporal K-function (see eq. 12) and the expected one (see eq. 14) weighting by, respectively, the true intensity from the ETAS model(left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right).

Finally, repeating again the above procedure for 100 simulations of a spatio-temporal ETAS process with the previous parameter specifications, we evaluate the distribution of the χ^2 statistic based on the differences between the weighted K-functions and the theoretical expected values. The results are shown in Figure 8, and highlight the bad result when in the weighting scheme the wrong homogeneous Poisson intensity is used.

5.2 Local space-time diagnostics

We now focus on the properties of the local weighted spatio-temporal K-function reported in Section 4, and in particular we focus on the theoretical expected value provided in equation 16. We assess these properties through simulations. We first consider an inhomogeneous Poisson process with the same previous intensity $\lambda(x,y,t) = \exp(a-4y-2t)$ with $a = 8.25$, and estimated the local weighted spatio-temporal K-function as in (15), using the three weighting schemes as above (true, estimated and constant intensities). Note that as these are local characteristics, we do have a surface per individual point of the point pattern. Thus, we can follow the same procedure as in Section 5.1 but for each point, computing the corresponding χ^2 statistic for the difference between the estimated local weighted spatio-temporal K-function reported in eq. (15) and the corresponding theoretical expected value in eq. (16), over a fixed grid that covers the region of interest.



(a) (b) (c)

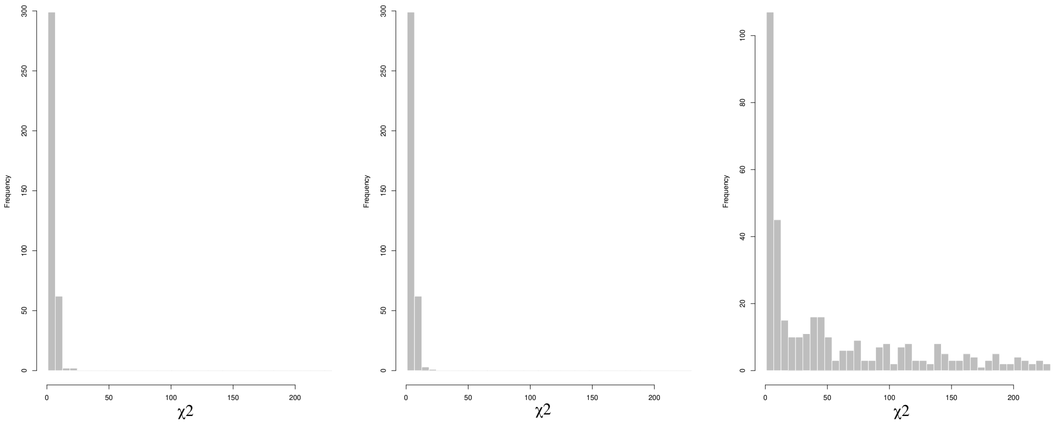
Fig. 8: Distribution of the χ^2 statistic computed for assessing the differences between the weighted K-functions and the theoretical expected value, calculated from 100 simulations of an ETAS process.

Figure 9 shows the distribution of the χ^2 statistics for each point of a spatio-temporal inhomogeneous Poisson process with the above intensity function.

If an individual point of the pattern would show a different spatio-temporal surface (e.g. because it is an outlier) its corresponding χ^2 statistic would reflect that difference and would show a larger value. Also, we note in Figure 9 for the local space-time K-function that the χ^2 statistic takes much larger values under a constant intensity weighting, detecting thus the wrong model. To highlight this behaviour, we selected those four individual points in the point pattern that provided the worst, i.e. the largest, χ^2 statistic for the simulated pattern. Figures 10-12 show the surface of the difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of these selected points, under the three different weighting schemes.

To avoid random fluctuations in the estimation of the K-function coming from just one realisation, we repeated the above procedure for 100 simulations of a spatio-temporal inhomogeneous Poisson process with the previous intensity function. We then evaluated the distribution of the χ^2 statistic based on the differences between the local weighted K-functions and the theoretical expected values for each individual point and for each realisation, using the three weighting schemes. The results are shown in Figure 13, and it basically reports the same general behaviour as in Figure 9 where it is highlighted the bad results when in the weighting scheme the wrong homogeneous Poisson intensity is used.

We again considered an ETAS spatio-temporal model, as in Section 5.1 to analyse if the results stand the same for a more complex model. We first considered only one simulation of a point pattern from a general ETAS model and evaluated the χ^2



(a) (b) (c)

Fig. 9: Distribution of the χ^2 statistic computed for assessing the differences between the local weighted K-functions and the theoretical expected value, calculated for each individual point from a simulation of an inhomogeneous Poisson process, and for the three weighting schemes.

statistic for the difference between every local surface of the weighted spatio-temporal K-function and its theoretical expected value (see Figure 14).

We selected those two individual points in the point pattern that provided the worst, i.e. the largest, χ^2 statistic for the simulated pattern. Figures 15-17 show the surface of the difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of these selected points, under the three different weighting schemes.

Finally, we repeated the above scenario for 100 simulations of an ETAS model, obtaining the distribution of the χ^2 statistic evaluated for the difference between empirical and expected surfaces for each individual in each realisation (see Figure 18).

5.3 Data application: seismic activity in Italy

In this section, we show an application to real data concerning the description of the seismic activity of the Italian area, from 2005 to 2013. Italy is a very active area and, in the considered period, several strong events have occurred, causing deaths and relevant damages to both modern and historical building (e.g. the Reggio Emilia sequence in 2012 (north of Italy), L'Aquila earthquake in 2009 (in the center of Italy)). Different kind of activity is observed around the Etna Volcano, where a quite continuous activity is recorded.

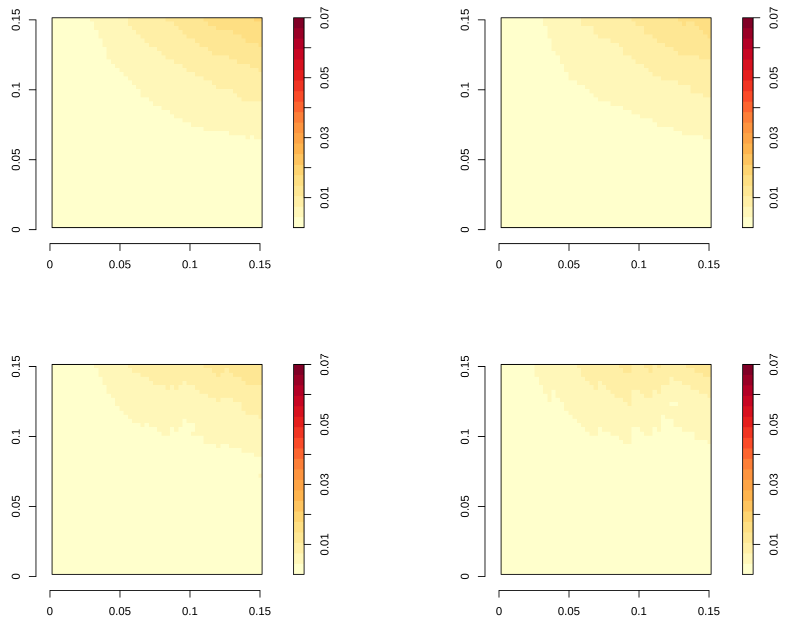


Fig. 10: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of four individual points of the inhomogeneous Poisson pattern that showed the largest χ^2 statistic, and when the weighting is through the true intensity.

The seismicity of Italy consists mainly of aftershock sequences and more seldom of isolated events. Starting from this catalog, we fit a space-time ETAS model, estimating the triggered intensity function and the background component using the FLP approach (etasFLP approach Chiodi and Adelfio (2017)). The observed seismicity together with the spatial diagnostic results are reported in Fig. 19. This diagnostic is based on a residual measure obtained by the comparison of the estimated and theoretical intensity computed over a fixed grid (see Adelfio and Chiodi (2015b) for more details).

From previous studies (Adelfio and Chiodi 2015b,a) we have observed that, although the estimated ETAS model for the whole area is generally satisfying mostly for the background seismicity, some characteristics are not well caught in the description of the space-time triggered part. Indeed, from these studies, we noted that the FLP improves the fitting of the background, that in general appears very appropriate for data, both in time and space, while some lack is still evident for the offspring component, mostly in the spatial domain. This would suggest the importance of modifying the ETAS model also in the induced component.

The diagnostics in **??** (b) of the estimated background and triggered intensity functions for Italy show high residual values in correspondence of some focal areas of

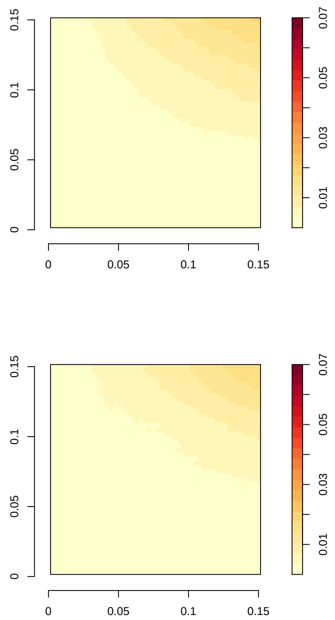


Fig. 11: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of four individual points of the inhomogeneous Poisson pattern that showed the largest χ^2 statistic, and when the weighting is through the estimated intensity.

the Italian seismicity, e.g. L'Aquila and Reggio Emilia, suggesting that a definition of an intensity function more coherent with some known tectonic structures is necessary for describing a such complex structure.

Performing a local diagnostic, as proposed in this paper, we localise the worst individual surfaces obtained by the comparison between the estimated local weighted spatio-temporal K-function and the theoretical expected value (see Fig. 20).

In Figure 21 we also report the map of Italy together with the observed events (in black) and the evidence of those points (in red) corresponding to the surfaces of Fig. 20.

As expected, the events for which the bad fitting is more evident are the earthquakes occurred in the main sequences of the observed region, confirming that a more complex model and a more flexible model for both the components of the ETAS model seems crucial in the presence of such inhomogeneous patterns.

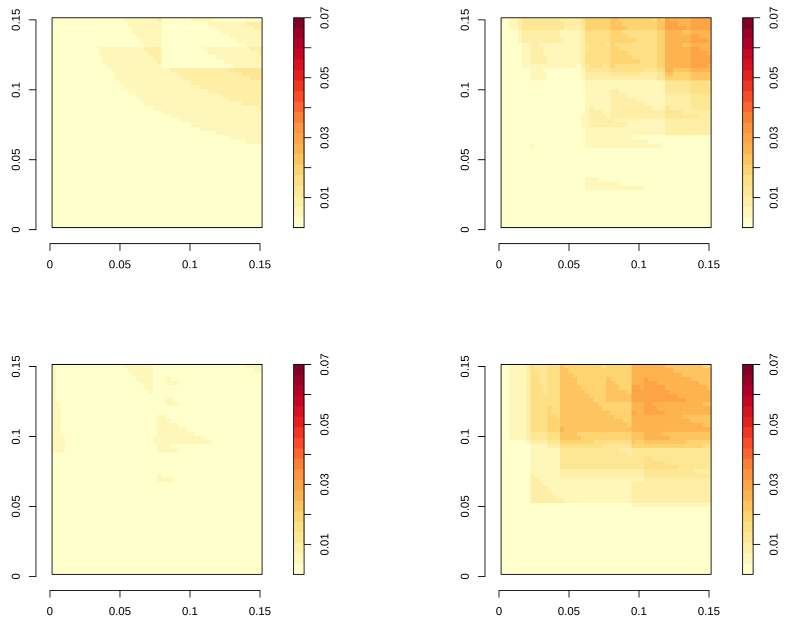
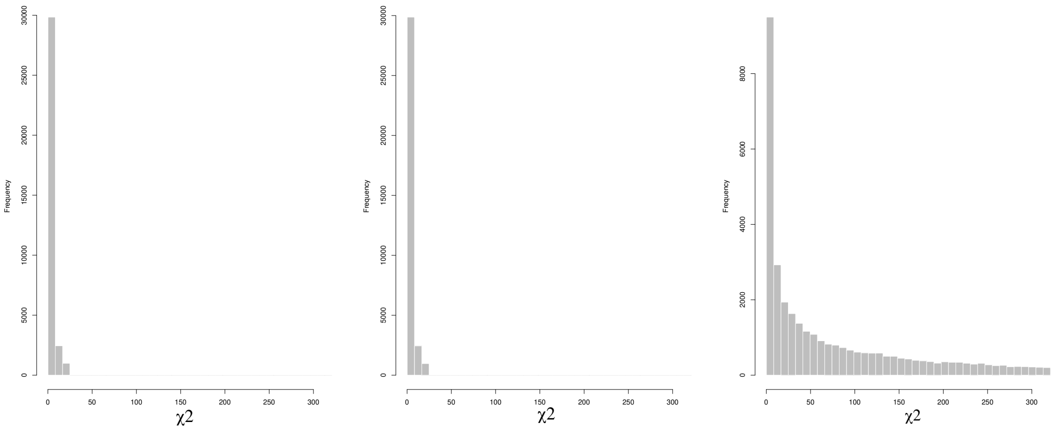


Fig. 12: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of four individual points of the inhomogeneous Poisson pattern that showed the largest χ^2 statistic, and when the weighting is through the constant intensity.

6 Conclusions and discussion

Alternatively to the more classical diagnostic methods based on transformation of the data into residuals as a result of a thinning or a rescaling procedure, we here have presented weighted second-order measures in the spatio-temporal domain. One attractive feature of weighted second-order statistics is that they directly apply to data without assuming homogeneity nor transforming the data into residuals, eliminating thus the sampling variability due to the use of a transforming procedure. We have provided theoretical arguments reinforced by a number of simulation studies. We have proved and also have shown by simulations that our methods do not rely on any particular model assumption on the data, and thus they can be applied for whatever is the generator model of the process. We believe that our method can be used routinely in the goodness-of-fit analysis for spatio-temporal point processes.

Investigations of convergence rates and of the sample sizes required to reach the asymptotic regime are not discussed here, and they are considered important issues for future work. Moreover, we have verified that a positive correlation seems to exist between the χ^2 statistics and the value of the conditional intensity function $\lambda(\cdot)$ for

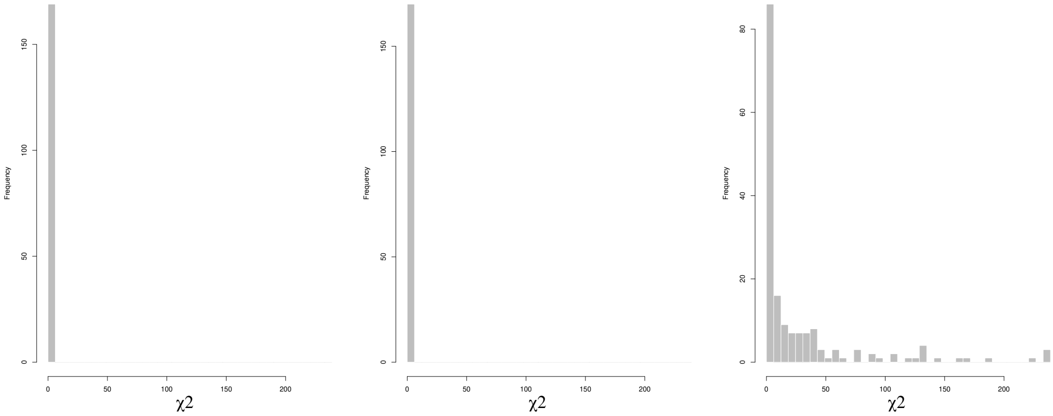


(a) (b) (c)

Fig. 13: Distribution of the χ^2 statistic computed for assessing the differences between the local weighted K-functions and the theoretical expected value, calculated from 100 simulations of a spatio-temporal inhomogeneous Poisson process weighting by, respectively, the true intensity (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right).

each point, both for the true and the estimated cases. In our opinion this point would require more investigation and further reflection to be extended for deeper comments. Other second-order measures, such as the pair-correlation, can be also considered as an alternative to the K-function. Finally, extensions to the marked case are very welcome in this context.

Acknowledgements This paper has been partially supported by the national grant of the Italian Ministry of Education University and Research (MIUR) for the PRIN-2015 program, ‘Complex space-time modelling and functional analysis for probabilistic forecast of seismic events’.



(a) (b) (c)

Fig. 14: Distribution of the χ^2 statistic computed for assessing the differences between the local weighted K-functions and the theoretical expected value, calculated for each individual point from a simulation of an ETAS process, and for the three weighting schemes.

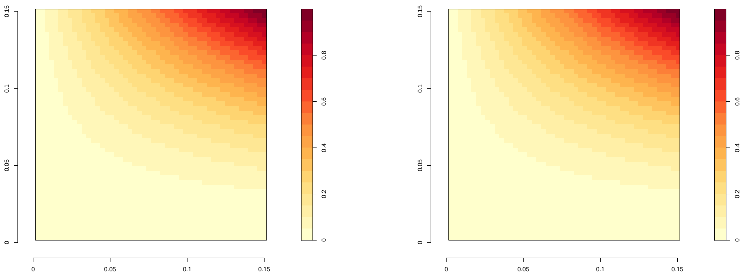


Fig. 15: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of two individual points of the ETAS pattern that showed the largest χ^2 statistic, and when the weighting is through the true intensity.

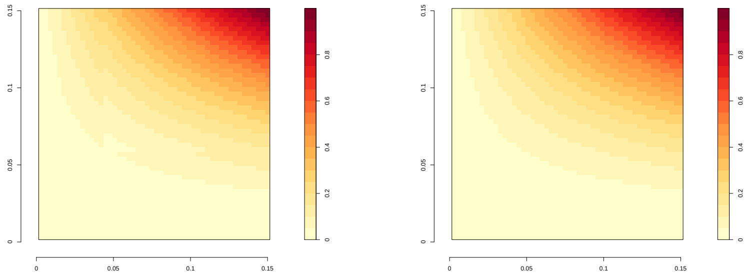


Fig. 16: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of two individual points of the ETAS pattern that showed the largest χ^2 statistic, and when the weighting is through the estimated intensity.

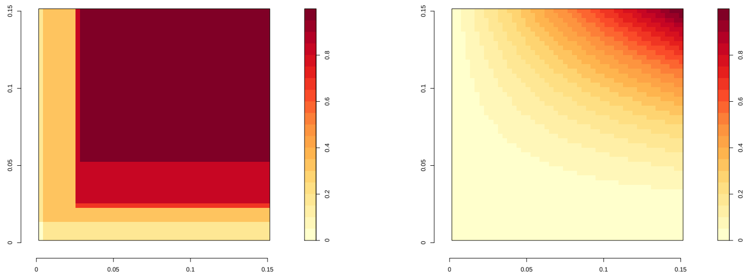
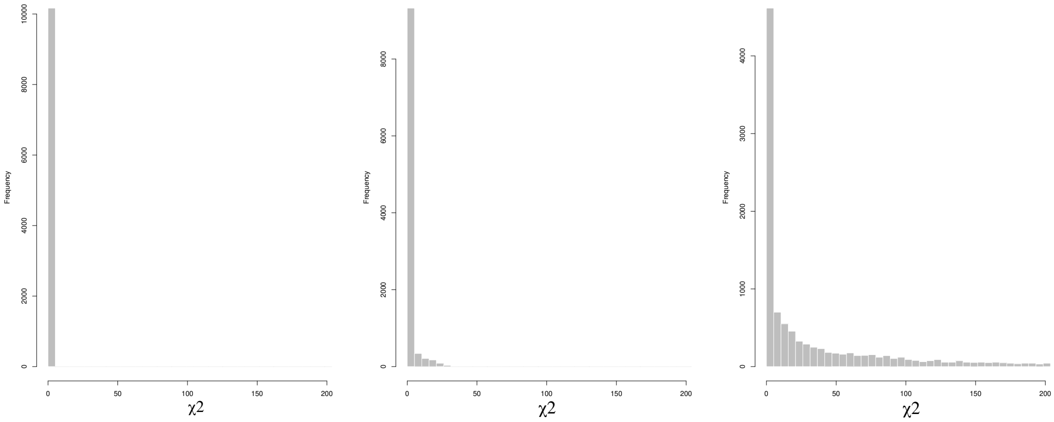
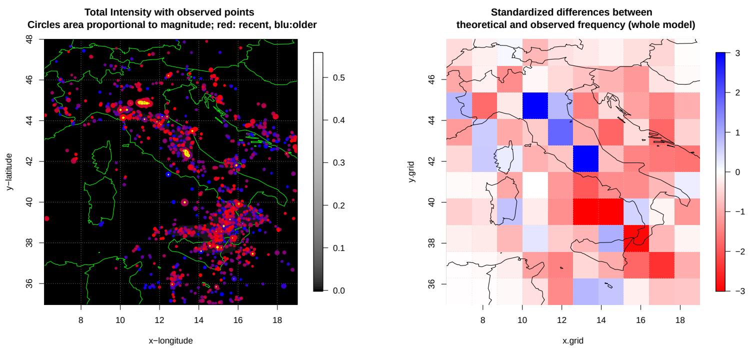


Fig. 17: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for each one of two individual points of the ETAS pattern that showed the largest χ^2 statistic, and when the weighting is through the constant intensity.



(a) (b) (c)

Fig. 18: Distribution of the χ^2 statistic computed for assessing the differences between the local weighted K-functions and the theoretical expected value, calculated from 100 simulations of a spatio-temporal ETAS model weighting by, respectively, the true intensity (left), the MLE intensity (middle), and a constant intensity, which is a wrong model (right).



(a) (b)

Fig. 19: Observed seismicity in Italy from 2005 to 2013, with $m > 3$ (a); Spatial diagnostics for the estimated space-time ETAS model by the etasFLP approach (b)

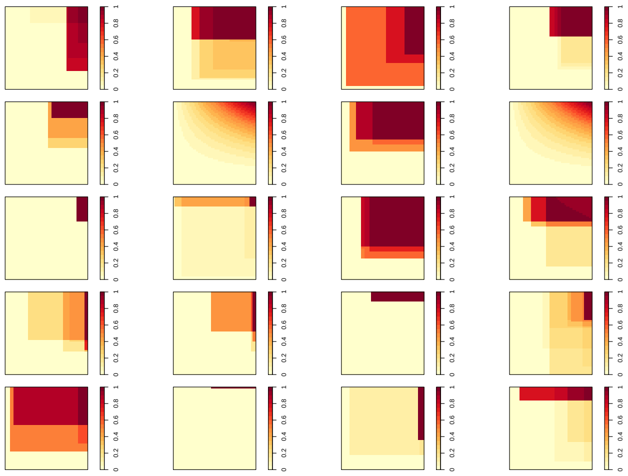


Fig. 20: Surface of difference between the estimated local weighted spatio-temporal K-function and the theoretical expected value for individual points of the ETAS pattern that showed the largest χ^2 statistic, and when the weighting is through the estimated ETAS intensity.

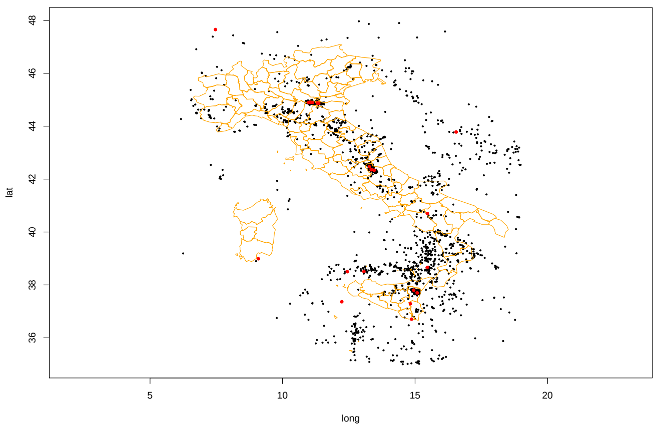


Fig. 21: Individuation of the 'bad' points

References

Adelfio, G. and M. Chiodi (2015a). Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stochastic Environmental Research and Risk Assessment 29*, 443–450.

Adelfio, G. and M. Chiodi (2015b). Flp estimation of semi-parametric models for space-time point processes and diagnostic tools. *Spatial Statistics 14*, 119–132.

Adelfio, G. and F. P. Schoenberg (2009). Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Annals of the Institute of Statistical Mathematics 61*(4), 929–948.

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis 27*(2), 93–115.

Baddeley, A., J. Møller, and R. Waagepetersen (2000). Non and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica 54*(3), 329–350.

Chiodi, M. and G. Adelfio (2011). Forward likelihood-based predictive approach for space-time processes. *Environmetrics 22*(6), 749–757.

Chiodi, M. and G. Adelfio (2017). Mixed non-parametric and parametric estimation techniques in R package etasFLP for earthquakes' description. *Journal of Statistical Software 76*(3), 1–28.

Cressie, N. (1991). *Statistics for spatial data*. Wiley series in probability and mathematical statistics.

Cressie, N. and L. B. Collins (2001). Analysis of spatial point patterns using bundles of product density lisa functions. *Journal of Agricultural, Biological, and Environmental Statistics 6*(1), 118–135.

Daley, D. J. and D. Vere-Jones (2003). *An introduction to the theory of point processes* (second ed.). New York: Springer-Verlag.

Gabriel, E. and P. J. Diggle (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica 63*(1), 43–51.

Getis, A. (1984). Interaction modeling using second-order analysis. *Environment and Planning A 16*, 173–183.

Getis, A. and J. Franklin (1987). Second-order neighborhood analysis of mapped point patterns. *Ecology 68*(3), 473–477.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and Its Application*. New York: Academic Press.

Mateu, J., G. Lorenzo, and E. Porcu (2010). Features detection in spatial point processes via multivariate techniques. *Environmetrics 21*(3-4), 400–414.

Meyer, P. (1971). Démonstration simplifée d'un théorème de knight. *Lecture Notes in Mathematics 191*, 191–195.

Møller, J. and M. Ghorbani (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica 66*(4), 472–491.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association 83*(401), 9–27.

Ripley, B. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B 39*, 172–212.

Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability 13*(2), 255–266.

Schoenberg, F. P. (1999). Transforming spatial point processes into poisson processes. *Stochastic Processes and their Applications 81*(2), 155–164.

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

Siino, M., F. J. Rodríguez-Cortés, J. Mateu, and G. Adelfio (2017). Testing for local structure in spatiotemporal point pattern data. *Environmetrics*.

Stoyan, D. and H. Stoyan (1994). *Fractals, Random Shapes and Point Fields*. Chichester: Wiley.

Stoyan, D. and H. Stoyan (2000). Improving ratio estimators of second-order point process characteristics. *Scandinavian Journal of Statistics 27*, 641–656.

Veen, A. (2006). *Some Methods of Assessing and Estimating Point Processes Models for Earthquake Occurrences*. Ph. D. thesis, UCLA.

Veen, A. and F. P. Schoenberg (2005). Assessing spatial point process models using weighted k-functions: Analysis of california earthquakes. *UC's eScholarship Repository*.

Zhuang, J. (2006). Second-order residual anaysis of spatio-timporal point processes and applications in model evaluation. *Journal of the Royal Statistical Society, Series B 68*(4), 635–653.