

Testing independence between two nonhomogeneous point processes in time

ARTICLE HISTORY

Compiled March 25, 2020

ABSTRACT

Point processes are often used to model the occurrence times of different phenomena, such as earthquakes, heatwaves or spike trains. Many of those modelling problems require to study the independence between two point processes, that in many cases are nonhomogeneous. This work develops three families of tests to assess the independence between two nonhomogeneous point processes in time. They can be applied to different types of processes (Poisson processes, cluster Poisson processes and other more general point processes), and all together they cover a wide range of situations appearing in real problems. The first family includes two tests for Poisson processes. The second family, with also two tests, is based on the close point distance, and the third one includes two nonparametric tests based on a version for nonhomogeneous processes in time of two cross dependence spatial functions. The tests of the two last families are based on bootstrap and computational methods.

An extensive simulation study of the size and power of the tests is carried out and, according to the results, some practical rules to select the most appropriate test in different cases, are provided. The proposed tests are demonstrated on a real data application about the occurrence of extreme heat events in three Spanish locations.

KEYWORDS

Nonhomogeneous point process; Poisson process; test of independence; bootstrap test

Word count (everything except appendix): ≈ 9600

1. Introduction

A point process in time (PP in short) is a random collection of points located on \mathbb{R}^+ , which represent the occurrence times of an event. Many real problems involve not one, but two or more PPs, for example, the timing of trades and mid-quote changes in stock exchange, the occurrence of climate extreme events in different locations, or the synchrony detection in spike train analysis. The study of dependence between the processes is important since independent processes can be modelled separately without any loss of information, while the modelling of dependent PPs is more complicated. For example, the study of independence between the processes in multi-site datasets will allow us to determine if sites can be modelled separately, or more complex spatial models have to be used.

Tests to assess independence between PPs are also a useful tool in statistical modelling problems involving covariates. In effect, a common source of dependence is the influence on several processes of the same or dependent variables, see for example [1] or the extreme heat event analysis in Section 5. A usual approach to model this dependence is to allow the parameters of the marginal models to be a function of those

covariates. Then, to analyse if the dependence is well captured by the covariates, the null hypothesis of independent marginal models given the covariates has to be checked. Independence is also a common assumption in statistical models, such as the common Poisson shock processes [2], so that independence tests are required to validate those models.

The analysis of independence between PPs has been mainly studied in the framework of spatial processes. The well-known K and J functions quantify the interpoint dependence in a PP, and their cross versions measure the dependence between two PPs. They were first defined in the homogeneous case and extended to nonstationary multivariate point processes [3,4]. There are some tests of independence for marked spatial point processes [5,6]. Allard et al. [7] proposed a test to analyse local independence in spatial PPs based on isotropy properties. In most cases, the spatial tools cannot be straightforwardly applied to PPs in time, although some underlying ideas can be adapted to the time framework, as we will see later.

The study of independence tests between PPs is also important in neurosciences, to assess if two spike trains are dependent [8]. Tuleau-Malot et al. [9] proposed an asymptotic test for Poisson processes, and Albert et al. [10] a bootstrap and permutation test. Although this type of data are often nonstationary the previous tests are developed for homogeneous processes. Another limitation is that they require a sample of m i.i.d trials of the processes, and although this type of samples is common in the study of spike trains, it is not in many other fields.

Not many tools for the study of dependence between nonhomogenous (NH) processes have been developed, although homogeneity is seldom a reasonable assumption in real problems. Dutilleul [11] proposed a graphical tool based on a Diggle's randomization but, as far as we know, only Abaurrea et al. [12] developed a formal tests to asses independence in a NH framework, with a study of its size and power.

The aim of this work is to propose different statistical tests to study the independence between two NH point processes in time, in different situations. Given the difficulty of characterising the distribution of the statistic in this type of tests, bootstrap and computational methods are used to obtain the p-values, as it is often done in spatial PPs [13–15]. In this framework, we develop three families of tests which can be applied to different types of point processes with a non random intensity (Poisson processes, homogeneous and NH point processes, and others). Although the size and power of the tests cannot be determined analytically, an extensive simulation study is conducted to assess their performance [16,17]. This study also allows us to stablish practical rules to select the most appropriate test in each situation.

The outline of the paper is as follows. Section 2 summarizes some properties of PPs. Three families of independence tests are developed in Section 3, and their size and power is thoroughly analysed in Section 4. Section 5 shows an application to the study of the independence between the occurrence of extreme heat events, and conclusions are discussed in Section 6.

2. Point processes in time

A point process in time N can be described in three equivalent ways: by the sequence of occurrence times in \mathbb{R}^+ , T_1, T_2, \dots, T_n , by the interpoint distances, $\tau_i = T_i - T_{i-1}$ for $i = 1, \dots, n$, with T_0 usually equal to 0, or by the counting measures defined by a set of random variables $N(A)$ representing the number of points in A , for each $A \in \mathbb{R}^+$. The usual notation is $N(t) = N((0, t])$, and $N(t) = \sum_{i \geq 1} I_{(T_i \leq t)}$, where I_B is the indicator

variable of the set B . The intensity measure of the process, Λ , gives the expected number of points, so that $\Lambda((0, t]) = E(N(t))$. Its derivative function, provided it exists, is the intensity function $\lambda(t) = \frac{\partial \Lambda((0, t])}{\partial t}$. If $\lambda(t)$ is constant, the process is homogeneous, and NH otherwise. The archetypal PP is the Poisson process, where $N(A) \sim \text{Poisson}(\mu_A)$ with $\mu_A = \int_A \lambda(t) dt$, and, $N(A_1), \dots, N(A_k)$ are independent r.v. if $A_i \cap A_j = \emptyset \ \forall i \neq j$.

Herein, we will consider vectors of two PPs (N_x, N_y) observed in the same time interval $(0, T]$, where t_1, \dots, t_{n_x} and s_1, \dots, s_{n_y} denote the n_x and n_y points in each process. In general, N_x and N_y are not independent.

3. Tests of independence between point processes

Independence is a complex hypothesis, since both the marginal distribution of the processes and their interaction are involved. A common approach to construct this type of tests is to fix the observed marginal structure [15,18]. All the tests in this work use this approach, and are constructed by keeping fixed the observed process N_x .

It is noteworthy that in all the suggested tests the null hypothesis is the independence between processes N_x and N_y , and the alternative is the existence of any type of random dependence between them. Some examples of common types of dependence will be described in Section 4.2.1. Each family uses different statistics and requires different assumptions. The first family, denoted POISSON, assumes that N_y is a Poisson process, and the distribution of the statistics is characterized under that assumption. The other two families, CLOSE and CROSS, are based on the close point distance and cross dependence functions, respectively, and their p-values are obtained using bootstrap and computational methods.

3.1. POISSON family

The POISSON family of tests requires N_y to be a Poisson process with known intensity $\lambda_y(t)$. We keep fixed the observed process N_x , and under independence between N_x and N_y , given that a point t_i has occurred, the distribution of N_y , a $\text{Poisson}(\lambda_y(t))$ process, does not change. Consequently, the distribution of Y_i , the number of points in N_y in an interval I_i of length l_i around t_i , is $\text{Poisson}(\mu_i)$ with $\mu_i = \int_{I_i} \lambda_y(t) dt$, for $i = 1, \dots, n_x$. For simplicity, disjoint centred intervals of maximum length l around t_i are considered. Two tests based on this property are proposed.

Poisson test. Considering disjoint intervals around each t_i and under the null, the test statistic $Y = \sum_{i=1}^{n_x} Y_i \sim \text{Poisson}(\mu)$ with $\mu = \sum_{i=1}^{n_x} \mu_i$. The p-value is

$$pv = 2 \min \left[\left(P(Y < y_o) + \frac{1}{2} P(Y = y_o) \right), \left(P(Y > y_o) + \frac{1}{2} P(Y = y_o) \right) \right],$$

where y_o is the total number of points in N_y in the considered intervals. The p-value of a discrete statistic only guarantees that the size is lower than the significance level, but a study of the size of this test is carried out in Section 4.1.

Normal test. Under independence, and using disjoint intervals, Y_i are independent Poisson variables with $E(Y_i) = \text{Var}(Y_i) = \mu_i$. Defining $Y_i^* = (Y_i - \mu_i) / \sqrt{\mu_i}$, we obtain $E(Y_i^*) = 0$ and $\text{Var}(Y_i^*) = 1$, but still they are not identically distributed. However, those variables satisfy the Lyapunov condition, that states that given independent r.v.

with null expectation and finite variance σ_i^2 , if for some $\delta > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n E(|Y_i^*|^{2+\delta}) = 0, \quad (1)$$

with $s_n^2 = \sum_{i=1}^n \sigma_i^2$, then the central limit theorem (CLT) applies. In effect, the variables Y_i^* satisfy condition (1) with $\delta = 2$, given that $s_n^2 = n$ and $E[(Y_i - \mu_i)^4] = \mu_i(1 + 3\mu_i)$, since $Y_i \sim \text{Poisson}(\mu_i)$, so that,

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n E \left[\left(\frac{Y_i - \mu_i}{\sqrt{\mu_i}} \right)^4 \right] = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \frac{\mu_i + 3\mu_i^2}{\mu_i^2} \leq \lim_{n \rightarrow \infty} \frac{1}{n} \left(3 + \frac{1}{\min_{1 \leq i \leq n} \mu_i} \right) = 0.$$

Then, applying the CLT,

$$O_{n_x} = \frac{1}{\sqrt{n_x}} \sum_{i=1}^{n_x} \frac{Y_i - \mu_i}{\sqrt{\mu_i}}$$

is asymptotically $N(0, 1)$, and a test based on O_{n_x} can be constructed. The p-value is defined as,

$$pv = 2P(Z > |\bar{y}_o|)$$

with $Z \sim N(0, 1)$ and $\bar{y}_o = \frac{1}{\sqrt{n_x}} \sum_{i=1}^{n_x} \frac{y_i - \mu_i}{\sqrt{\mu_i}}$.

Given the asymptotic distribution of the test, a simulation analysis is performed in Appendix A to study the Normal approximation of O_{n_x} in terms of the values of n_x and μ_i .

3.2. CLOSE family

The CLOSE family is based on the close point distance. The first test can be applied to any process with a constant intensity (homogeneous processes). The second can be used with NH processes but requires N_y to follow a known probability model with a generation algorithm (parametric model herein). The concept of close points was introduced by Abaurrea et al. [12]: s_j in N_y is a close point of t_i in N_x , if the intervals to their previous points, s_{j-1} and t_{i-1} , overlap, see Figure 1. The test aims to compare the behaviour of the sets of close points in the observed PPs and in two independent PPs with the same marginal distribution. The behaviour of the set is summarized by the distances between t_i and its close points, $dx_{ij} = |t_i - s_j|$, and their mean value \bar{dx}_i .

A test based on a theoretical distribution is difficult to obtain, but a computational statistical approach is used to construct a summary statistic and a p-value. The implementation of this approach requires a sample of R vectors of processes $N_r^* = (N_x, N_{y,r}^*)$, where N_x is the observed first process, and $N_{y,r}^*$ is a generated process with the same distribution as N_y and independent of N_x ; $N_{y,r}^*$ must be mutually independent.

Using (N_x, N_y) and the R pairs $(N_x, N_{y,r}^*)$, dx_i and R mean distances \bar{dx}_{ir} from the independent processes are calculated for each $t_i \in N_x$. Then, p_i , the percentile rank divided by 100 of each observed \bar{dx}_i in its corresponding sample of distances under

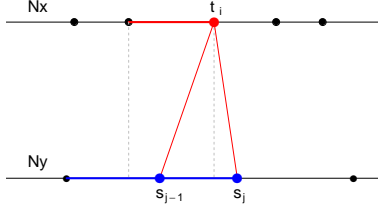


Figure 1. Definition of close points: s_{j-1} and s_j are close points of t_i .

the null, are calculated, and a sample p_1, \dots, p_{n_x} is obtained. Hence, the problem of testing independence is reduced to assess if p_1, \dots, p_{n_x} is a uniform but possibly correlated sample. In effect, under the null, $p_i \sim \text{Uniform}(0, 1)$ because they come from percentiles ranks of i.i.d observations, but p_1, \dots, p_{n_x} may be correlated because the sets of close points of nearby points in N_x may not be disjoint. To study the uniform character of the sample, the statistic D_n of the Kolmogorov-Smirnov test is considered. Since the distribution of this statistic in a correlated sample is unknown, the empirical distribution is obtained by a computational approach based again on the pairs $(N_x, N_{y,r}^*)$. Then, the only issue to apply this approach is to generate pairs $N_r^* = (N_x, N_{y,r}^*)$, and this can be done in different ways.

Lotwick-Silverman test (LoS). Processes $N_{y,r}^*$ are generated by the LoS approach [18], which does not require any assumption, not even to know the marginal intensities, but it is only valid for homogeneous processes. In this approach, the time interval $(0, T]$ is wrapped onto a circumference by identifying the opposite sides. Then, fixing N_x , a new $N_{y,r}^*$ is generated by translating N_y a random uniform amount over the circumference. This breaks any dependence between the processes and keeps the marginal distributions, provided they do not change over time. Consequently, it is only valid for processes with a constant intensity.

Parametric bootstrap test (PaB). The test by Abaurrea et al. [12] requires N_y to follow a parametric model, that will be used to generate independent $N_{y,r}^*$. If the model parameters are known, it is a Monte Carlo approach, but if they have to be estimated, it is a parametric bootstrap.

3.3. CROSS family

The CROSS family is based on statistics inspired by the spatial cross K and J functions [4,19], adapted to PPs in time. These tests can be applied to any pair of point processes, the only assumption is that $\lambda_x(t)$ and $\lambda_y(t)$ are known. The p-values of the two proposed tests are calculated using a LoS approach.

3.3.1. Test statistic based on the cross K-function

In the case of two homogeneous PPs, N_x and N_y , the spatial cross function $K_{xy}(r)$ is defined as the expected value of the number of points in N_y within a distance r of a randomly chosen point in N_x , divided by λ_y . This function can be adjusted to NH

processes in the following way

$$K_{xy}(r) = \frac{1}{|B|} E \left[\sum_{t_i \in N_x \cap B} \sum_{s_j \in N_y} \frac{I(|t_i - s_j| \leq r)}{\lambda_x(t_i) \lambda_y(s_j)} \right] \quad r \geq 0 \quad (2)$$

where $|B|$ is the area of $B \in \mathcal{B}_0$, the class of bounded Borel sets in \mathbb{R}^2 [20].

Given the definition in (2), a natural measure of the dependence between two PPs in time, in a time interval of length $2r$, is

$$\tilde{K}_{1,xy}(r) = \frac{1}{T} \sum_{t_i \in N_x \cap (0, T)} \sum_{s_j \in N_y} \frac{I(|t_i - s_j| \leq r)}{\lambda_x(t_i) \lambda_y(s_j)} \quad 0 \leq r < T.$$

$\tilde{K}_{1,xy}(r)$ is adjusted to NH processes by scaling the occurrence of points s_j in the interval $[t_i - r, t_i + r]$ by the intensity at that point $\lambda_y(s_j)$. However, with a highly varying intensity, the mean intensity of N_y in $[t_i - r, t_i + r]$, $\int_{t_i-r}^{t_i+r} \lambda_y(t) dt / 2r$, could be a better scale factor. Hence, we propose a new measure,

$$\tilde{K}_{2,xy}(r) = \frac{1}{|W|} \sum_{t_i \in N_x \cap W} \sum_{s_j \in N_y} \frac{I(|t_i - s_j| \leq r)}{\lambda_x(t_i) (\int_{t_i-r}^{t_i+r} \lambda_y(t) dt / 2r)} \quad 0 \leq r < T$$

Since estimations of $\lambda_x(t)$ and $\lambda_y(t)$ are often needed, and PPs are usually observed in discrete time, a common estimator of the mean intensity is $\sum_{t=t_i-r}^{t_i+r} \hat{\lambda}_y(t) / 2r$.

To summarize the dependence appearing in any value r , a reasonable statistic is,

$$\mathcal{K}_i = \frac{1}{R} \sum_{r=r_1}^{r_R} \frac{\tilde{K}_{i,xy}(r)}{2r} \quad (3)$$

for $i = 1, 2$, and with r_1, \dots, r_R a grid of R values (r-grid). We emphasize that our aim here is only to obtain a statistic able to discriminate between dependent and independent processes, and that it is not needed to completely characterize its distribution under the null, since the p-value will be obtained by a bootstrap method.

3.3.2. Test statistic based on the cross J function

The spatial cross J function $J_{xy}(r)$ measures the dependence between N_x and N_y by comparing the nearest point function $D_{xy}(r)$ and the empty space function $F_y(r)$,

$$J_{xy}(r) = \frac{1 - D_{xy}(r)}{1 - F_y(r)}$$

if $F_y(r) < 1$. $D_{xy}(r)$ is the distribution function of the distances from a point in N_x to the nearest point in N_y , and $F_y(r)$ of the distances from a point in the space to the nearest point in N_y . Under independent PPs, the occurrence of a point in N_y is not influenced by the occurrence of a point in N_x . Then, $D_{xy}(r) = F_y(r)$, and $J_{xy}(r) = 1$ for any r . In a NH framework the J function is still useful if it is adjusted for the

nonhomogeneous intensity as suggested by [4],

$$D_{xy}(r) = 1 - E^{(t_x)} \left[\prod_{s_j \in N_y} \left(1 - \frac{\inf_{t \in R^2} \lambda_y(t)}{\lambda_y(s_j)} I\{s_j \in B(t_x, r)\} \right) \right]$$

where $E^{(t_x)}$ is the expectation calculated with the conditional probability given that a point of N_x occurs in t_x [21] and $B(t_x, r)$ is the closed ball centred at t_x with radius r . Analogously, $F_y(r)$ is,

$$F_y(r) = 1 - E \left[\prod_{s_j \in N_y} \left(1 - \frac{\inf_{t \in R^2} \lambda_y(t)}{\lambda_y(s_j)} I\{s_j \in B(t_x, r)\} \right) \right].$$

The difference between $D_{xy}(r)$ and $F_y(r)$ is that the expectation in $D_{xy}(r)$ is conditioned on the occurrence of a point in N_x .

Using the ideas of the spatial cross J function, we suggest the use of $\tilde{J}_{xy}(r) = \frac{1 - \tilde{D}_{xy}(r)}{1 - \tilde{F}_y(r)}$ as a measure to quantify the dependence between two PPs in time, in intervals of length $2r$, where $\tilde{D}_{xy}(r)$ is defined as,

$$\tilde{D}_{xy}(r) = 1 - \frac{1}{\sum_{t_i \in N_x \cap W_r} \frac{1}{\lambda_x(t_i)}} \sum_{t_i \in N_x \cap W_r} \frac{\prod_{s_j \in N_y \cap (t_i - r, t_i + r)} \left(1 - \frac{\lambda_y^*}{\lambda_y(s_j)} \right)}{\lambda_x(t_i)}$$

where $\lambda_y^* = \inf_{t \in (0, T]} \lambda_y(t)$, index t_i moves in the points in N_x which lie in the space $W_r = [r, T - r]$ and s_j in the points in N_y which lie in $(t_i - r, t_i + r)$. Analogously, $\tilde{F}_y(r)$ is defined as,

$$\tilde{F}_y(r) = 1 - \frac{1}{\#(L \cap W_r)} \sum_{l_k \in L \cap W_r} \left[\prod_{s_j \in N_y \cap (l_k - r, l_k + r)} \left(1 - \frac{\lambda_y^*}{\lambda_y(s_j)} \right) \right]$$

where $L \subseteq (0, T]$ is a finite grid with points l_k (L-grid), and $\#(\cdot)$ the number of points in a set. If $\tilde{D}_{xy}(r)$ and $\tilde{F}_y(r)$ are applied to homogeneous PPs, the empirical distribution functions of the corresponding distances are obtained.

To consider dependence in any value r , and since deviations of $\tilde{J}_{xy}(r)$ from 1 suggest dependence, a reasonable summary statistic is

$$\mathcal{J} = \frac{1}{R} \sum_{r=r_1}^{r_R} |\tilde{J}_{xy}(r) - 1|. \quad (4)$$

3.3.3. A LoS approach to obtain the p-value

It is difficult to determine the distribution under the null of the statistics \mathcal{K}_i , expression (3), and \mathcal{J} , expression (4), but a LoS approach can be used. In this approach, the observation space is wrapped onto a circumference, N_x is kept fixed and N_y is randomly translated over the circumference to break the potential dependence. In NH processes, the problem is that random translations change the marginal distribution of

Table 1. Summary of the characteristics of the proposed independence tests.

Family	Test	Assumptions	Statistic	Distribution
POISSON	Poisson test	N_y Poisson process	$\sum_{i=1}^n Y_i$	Poisson
	Normal test	with known λ_y	$\frac{1}{\sqrt{n_x}} \sum_{j=1}^{n_x} \frac{Y_j - \mu_j}{\sqrt{\mu_j}}$	Asymptotic Normal
CLOSE	LoS test	N_y homogeneous PP	KS statistic D_n	LoS approach
	PaB test	N_y parametric model with known parameters		Parametric bootstrap
CROSS	K_i tests J test	Known λ_x, λ_y	$\frac{1}{R} \sum_{r=r_1}^{r_R} \frac{K_{i,xy}(r)/2r}{ J_{xy}(r) - 1 }$	LoS approach

the processes. However, since $\tilde{K}_{i,xy}$ and \tilde{J}_{xy} are adjusted for the time-varying intensity, this approach is still valid to obtain values of the statistic under the null, provided that the intensity is translated with the process. Using the empirical distribution from those samples, the p-value can be obtained. Lotwick and Silverman [18] and Cronie and van Lieshout [4] applied a similar approach in homogeneous and NH spatial processes, respectively.

The tests based on \mathcal{K}_1 , \mathcal{K}_2 and \mathcal{J} are denoted K_1 , K_2 and J , and the influence of r and L grids on their size and power is studied in Section 4.

The main characteristics of all the proposed tests are summarized in Table 1. All the tests except LoS assume that $\lambda_y(t)$ is known.

4. Size and power analysis

The aim of this section is to perform a simulation study of the size and power of each test in terms of the intensity, the sample size n_x , and in the case of the power, the type and level of dependence. Three intensities are considered in the study: a constant λ^C (homogeneous process), a log-linear intensity that increases in time, $\lambda^I(t) = \exp(a+bt)$, and $\lambda^F(t)$, the fitted intensity to a real dataset with a complex shape, including a non monotonic trend and seasonal behaviour, see Figure 2. In all cases the mean intensity in $(0, T]$, $\bar{\lambda}(t)$, is 0.01 but the variability is quite different, with a range equal to 0, 0.02 and 0.40 respectively. For simplicity, the same intensity is considered in N_x and N_y .

In a simulation study the number of points in $(0, T]$ in a PP, n_x , is random, but the expected sample size $\bar{n}_x = T\bar{\lambda}(t)$ can be fixed. Since $\bar{\lambda}(t) = 0.01$, $T = 5000, 10000$ or 20000 time units are used to obtain $\bar{n}_x = 50, 100$ and occasionally 200. The results for significance levels, $\alpha_1 = 0.05$ and $\alpha_2 = 0.1$, are calculated.

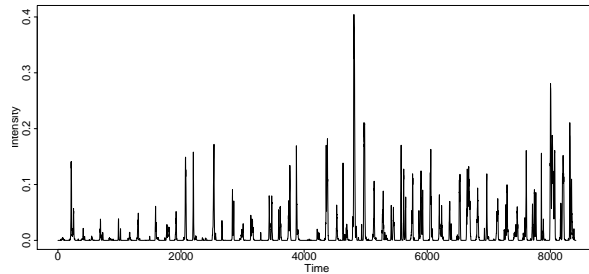


Figure 2. Intensity function $\lambda^F(t)$

Table 2. Size analysis of the POISSON and CLOSE tests, $\alpha_1 = 0.05$ and $\alpha_2 = 0.10$.

Test		Poisson		Normal		PaB		LoS	
Int.	n_x	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
λ^C	50	.05	.10	.05	.10	.05	.11	.06	.11
	100	.05	.10	.05	.10	.05	.10	.06	.11
$\lambda^I(t)$	50	.05	.10	.05	.10	.06	.11	-	-
	100	.05	.10	.05	.10	.06	.10	-	-
$\lambda^F(t)$	50	.05	.10	.05	.10	.06	.11	-	-
	100	.05	.10	.05	.09	.05	.11	-	-

4.1. Size analysis

The simulation study is as follows: 1000 trajectories of two independent Poisson processes, with intensities $\lambda_x(t)$ and $\lambda_y(t)$, are generated in $(0, T]$, and the considered test is applied. The size at an α significance level is estimated as the rate of rejections at that level in the 1000 repetitions.

The CLOSE tests, and the POISSON tests with the minimum l obtained to assume a good Normal approximation, show an estimated size with values equal or very close to the nominal level in all the settings, see Table 2.

In the CROSS family, the influence of the grids on the size is studied first. Four r -grids are considered: $r1$ and $r1b$ cover a small range of values, from 1 to around half of the mean interpoint distance of the PPs, with 10 and 50 points respectively; $r2$ and $r2b$ cover a wider range from 1 to around five times the mean interpoint distance. Two L -grids are considered in the J test: La is a sparse grid with points at a distance 50, or 100 if the number of points in the grid is higher than 100, and Lb is a thick grid with points every 2 time units. These grids are summarized in Table 3 and the results in Table 4. They show that both L -grids lead to a similar size so that La , is used herein. The estimated size of K_2 test is adequate with all the intensities and r -grids, only with $\lambda^I(t)$ and $r2b$ it is a bit high, 0.08 for $\alpha = 0.05$. The estimated size of the J test is satisfactory in homogeneous processes, the size with $\lambda^I(t)$ doubles the nominal value, and with $\lambda^F(t)$, it is close to the nominal value only with the grids covering a wide range of values, $r2$ and $r2b$. The estimated size of the K_1 test is too high in all the NH processes, specially with $\lambda^F(t)$, consequently, only the K_2 test is kept in the analysis.

Table 3. r -grids and L -grids considered in the study of the CROSS family.

r -grid	Sparse	# points	Thick	# points
Short range	$r1=1,6,\dots, 46$	10	$r1b=1,2,\dots, 50$	50
Long range	$r2=1,101,\dots, 451$	10	$r2b=1,11,\dots, 491$	50
L -grid	Sparse		Thick	
T=5000	$La=1,51, 101,\dots, 4951$	100	$Lb=1,3,5,\dots, 5000$	2500
T=10000 or 20000	$La=1,101,\dots, 9901$	100	$Lb=1,3, 5,\dots, 10000$	5000

Table 4. Effect of r and L -grids on the size of the CROSS tests; $\bar{n}_x = 50$ and grids $r1, r1b, r2$ and $r2b$. The size of the J test is calculated with the grid La , for all r -grids, and with Lb for $r1$ and $r2$ (in brackets)

Test		K_1		K_2		J	
Int.	r	α_1	α_2	α_1	α_2	α_1	α_2
λ^C	$r1$.04	.09	.04	.09	.06 (.05)	.12 (.11)
	$r1b$.06	.10	.06	.10	.06	.11
	$r2$.04	.10	.04	.10	.05 (.05)	.10 (.12)
	$r2b$.05	.10	.04	.10	.05	.12
$\lambda^J(t)$	$r1$.07	.13	.06	.12	.09 (.10)	.14 (.15)
	$r1b$.08	.13	.06	.12	.10	.15
	$r2$.06	.13	.05	.11	.09 (.09)	.18 (.18)
	$r2b$.11	.18	.08	.15	.12	.19
$\lambda^F(t)$	$r1$.09	.20	.06	.12	.11 (.10)	.18 (.17)
	$r1b$.11	.20	.06	.12	.11	.18
	$r2$.06	.15	.06	.11	.06 (.06)	.12 (.11)
	$r2b$.09	.15	.05	.10	.05	.09

4.2. Power analysis

The power is estimated as the rate of rejections at an α level in 1000 repetitions. Since the null hypothesis of all the tests is the independence between N_x and N_y , the tests are studied under four different types of alternative hypotheses, which represent common types of dependence structures appearing in real problems, and are described in Section 4.2.1. In each of those cases, high dependence (HD) and a low dependence (LD) levels are considered. Moreover, the power is estimated under the settings determined by the previous values of \bar{n}_x (determined by T) and the three intensities. LD50 denotes the setting with low dependence and $\bar{n}_x = 50$, and analogously LD100, HD50 and HD100.

There does not exist a general definition of dependence between PPs, but to somehow quantify the dependence in the HD and LD levels, the correlation between the number of points in the PPs is calculated in Appendix B. All the results in this section correspond to the power for $\alpha_1 = 0.05$, and are graphically summarized. The numeric values and the power for $\alpha_2 = 0.10$ are shown in Appendix C.1.

4.2.1. Models for dependent PPs

Common Poisson shock process (CPS): It is a vector of two PPs with an underlying Poisson process of shocks, which may yield points only in one or in both marginal processes at the same time [22]. It can be decomposed into three independent processes $N_{(x)}$, $N_{(y)}$ and $N_{(xy)}$, the processes of points occurring only in N_x , only in N_y , and in both of them (simultaneous points), with intensities $\lambda_{(x)}$, $\lambda_{(y)}$ and $\lambda_{(xy)}$. The three processes, and also N_x and N_y , are Poisson processes. N_x and N_y are independent if and only if $\lambda_{(xy)} = 0$. The HD and LD settings in this model correspond to $\lambda_{(x)} = \lambda_{(y)} = \lambda_{(xy)}/2$ and $\lambda_{(x)} = \lambda_{(y)} = 2\lambda_{(xy)}$, respectively.

Multivariate dependent Neyman-Scott process (MNS): A Neyman-Scott (NS) process is a process of clusters of points such that the cluster centres are a Poisson process [23]. We define a multivariate dependent Neyman-Scott process as a vector of NS processes with the same centres. Here, the number of points in each cluster is $Poisson(4)$, and the distribution of the i.i.d. distances of the points in a cluster to its centre, is $Normal(0, \sigma)$. The HD and LD levels are obtained with $\sigma_H = 20$ and $\sigma_L = 100$.

The CPS and the MNS models show a positive short-term dependence generated by unobserved common shocks, which may trigger points in several PPs, but the way the points are triggered is different. This dependence is quite common, for example when a warm front provoke heatwaves in different locations.

Queueing Process (QUE): A queue system models the arrivals and the output times in a waiting line with random serving times. Here, queues $M/M/\infty$ and $M(t)/M/\infty$ are considered to obtain homogeneous and NH processes. In both queues, the inputs and the outputs are dependent Poisson processes [24] and the serving times are *Exponential*(μ). The HD and LD levels are obtained with $\mu_H = 20$ and $\mu_L = 70$. The dependence in this model is a causal point to point relationship, where an event in a process triggers the occurrence of an event in the other. Examples with this type of dependence are the processes of the occurrence times of floods provoked by the occurrence of an event of intense rainfall.

Poisson process with dependent marks (PDM): A Poisson process with discrete marks is a process in which a variable taking values in $\{1, \dots, d\}$ is attached to each point, so that the marginal process j is formed by the points with mark j . The marginal processes are Poisson if and only if the marks are independent [25]. Here, marks from a 2-state Markov chain are used, and HD and LD levels are defined by the transition matrices,

$$M_H = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix} \quad M_L = \begin{pmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{pmatrix}.$$

The PDM model appears when the occurrence of an event in one process boots or blocks the occurrence of an event in the other. For example, the process of the growth of a specie of plant which favours or prevents the growth of another plant in time. It generates medium or long term dependence.

4.2.2. Power of the POISSON family

The POISSON tests require N_y to be a Poisson process, so that only CPS and QUE models can be used. The top plots in Figure 3 summarize the power of the Poisson and the Normal test for $\alpha = 0.05$ under the settings defined by the intensities (λ^C , $\lambda^I(t)$ and $\lambda^F(t)$), the sample size and dependence level (LD50, H50 and LD100), and dependence type (CPS and QUE). The power is obtained with an automatic selection of the interval length l , the median of the interpoint distances in N_y , which leads to a valid Normal approximation. The Normal test overperforms the Poisson test in all cases, so that we will focus on this test herein.

Sample size and dependence level. The power of the Normal test is over 0.9 and 0.72 in HD50 and LD100 respectively, except under QUE with $\lambda^F(t)$ where it is 0.74 and 0.26. The power always increases with the sample size and dependence level, even in the previous worst case, where it is 0.51 in LD200.

Type of dependence and intensity. Under CPS, the power is similar for all the intensities but, under QUE, it is lower with $\lambda^F(t)$, specially in LD.

Effect of l on the power. The bottom plots in Figure 3 show the power of the Normal test in the previous settings with interval lengths $l = 26, 50, 100$. Dependence in the CPS model appears in the simultaneous points, so that it is expected that the power will be higher using short intervals. Then, $l = 14$ is also considered in the settings where the Normal approximation may be assumed, that is LD100 with $\lambda^I(t)$ and all the settings with $\lambda^F(t)$, see Appendix A.

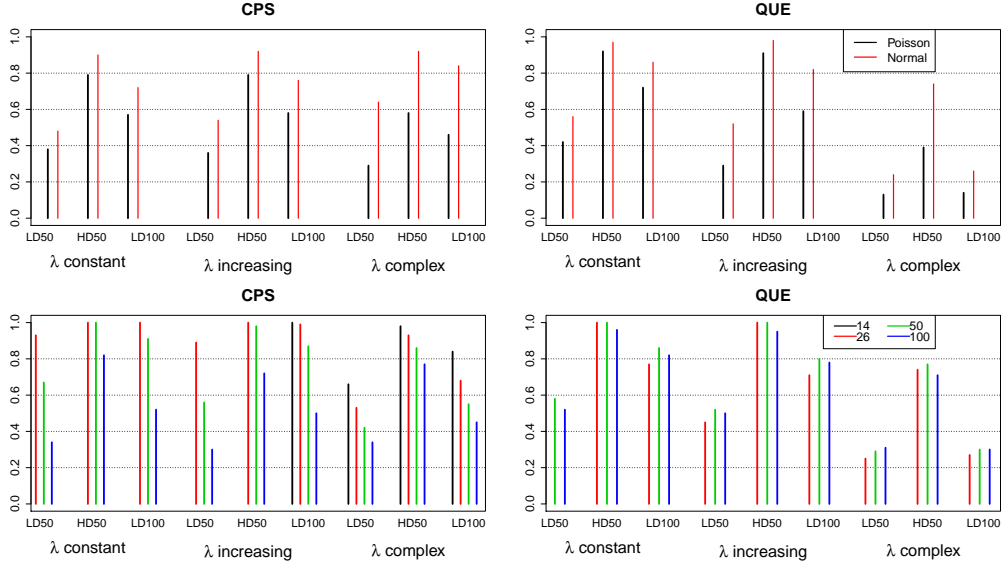


Figure 3. Top: Power of the POISSON tests with $\alpha_1 = 0.05$ under CPS and QUE models, with an automatic selection of l . Bottom: Effect of l on the power of the Normal test.

Under CPS the power is quite sensitive to l , and the best power is always obtained with the shortest interval. The power is over 0.55 up to $l = 50$, in all the settings except LD50 with $\lambda^F(t)$. Under QUE the effect of l is much weaker, with a similar power in all the l values, due to the long-term dependence of this model. It is noteworthy that if a length l adapted to the type of dependence is used, for example a low value under CPS, the power is higher than that obtained with an automatic selection of l .

4.2.3. Power of the CLOSE family

The CLOSE family can be applied to the four dependence models, but the LoS test only to homogeneous PPs. Figure 4 shows the power in the previous settings. The LoS test overperforms slightly the PaB test in LD50, but the improvement is negligible otherwise, with a power over 0.84, except in LD100 with PDM where it is 0.58.

Sample size and dependence level. The power always increases with the sample size and dependence level, and it is higher in HD50 than in LD100.

Type of dependence and intensity. The power with each intensity depends on the type of dependence: under PDM and CPS, it is similar with all the intensities, while under MNS and QUE, it is near 0.9 with λ^C and $\lambda^I(t)$ in HD50 and LD100, while $\lambda^F(t)$ requires bigger samples. The CPS dependence is the most easily detected, with a power around 0.8 in LD50, which increases to almost 1 in HD50 and LD100. The power under PDM is lower, 0.28 in LD50, but increases to 0.8 in HD50 and LD200.

To sum up, with a high enough sample size, a good power is obtained in all cases.

4.2.4. Power of the CROSS family

The K_2 and J tests require to know $\lambda_x(t)$ and $\lambda_y(t)$. Thus, they cannot be applied to MNS model, whose intensities do not have an explicit expression. A study of the power of the J test does not show any evidence of influence of the L -grid, so that La is used.

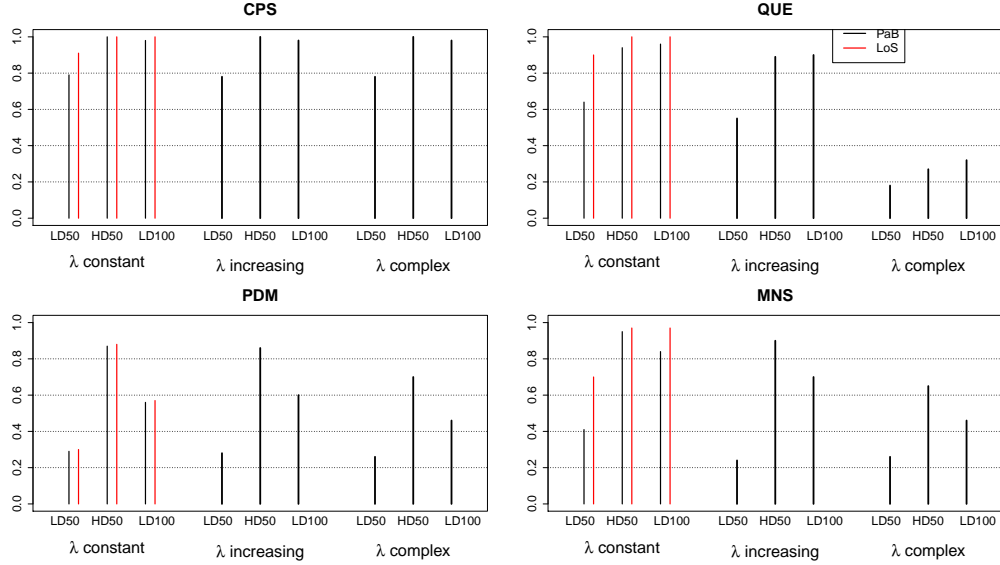


Figure 4. Power of the CLOSE family with $\alpha = 0.05$ under the four dependence models.

First, the effect of the r -grid on the performance of the tests is analyzed. Figure 5 (left) summarizes the power in the most complicated setting, LD50, using the grids $r1, r1b, r2$ and $r2b$. The J test with $\lambda^I(t)$ is not included, since its estimated size was not adequate. The power of K_2 does not show relevant differences among the r -grids, although $r2b$ leads to the highest, or close to the highest, power under all dependence models. Then, $r2b$ is a good option, except with $\lambda^I(t)$, where $r2$ should be preferred since the size is slightly better. The J test is quite sensitive to the r -grid with λ^C and under CPS with $\lambda^F(t)$, where a power below 0.10 is obtained with $r2$ and $r2b$, and around 0.9 with $r1$ and $r1b$. In the other settings, the power is low with all the r -grids. Moreover, only $r2b$ shows a correct size with $\lambda^F(t)$. Figure 5 (right) summarizes the power with the selected r -grids, and the following conclusions are obtained.

Sample size and dependence level. The power increases with the sample size and dependence level in all cases, except for the J test under PDM with $\lambda^F(t)$. In that setting, the test K_2 also improves slowly with the sample size.

Type of dependence and intensity. The power decreases with the complexity of the intensity, and the highest power is always observed with λ^C , closely followed by $\lambda^I(t)$. The power under the dependence models is quite heterogeneous. The K_2 test detects all the dependence structures and specially CPS with a high power, except PDM with a NH intensity. The J test is even more heterogeneous: its power is high with λ^C , except under PDM in LD50, but the test does not detect any dependence under CPS and PDM with $\lambda^F(t)$.

To sum up, the K_2 test clearly outperforms the J test, since its power is similar under CPS and QUE with λ^C and much higher otherwise. The J test is not recommended in NH processes.

4.3. Comparison of the POISSON, CLOSE and CROSS families

This section compares the test with the best general performance in each family: the Normal, the PaB and the K_2 tests. The solid lines in Figure 6 summarize the power

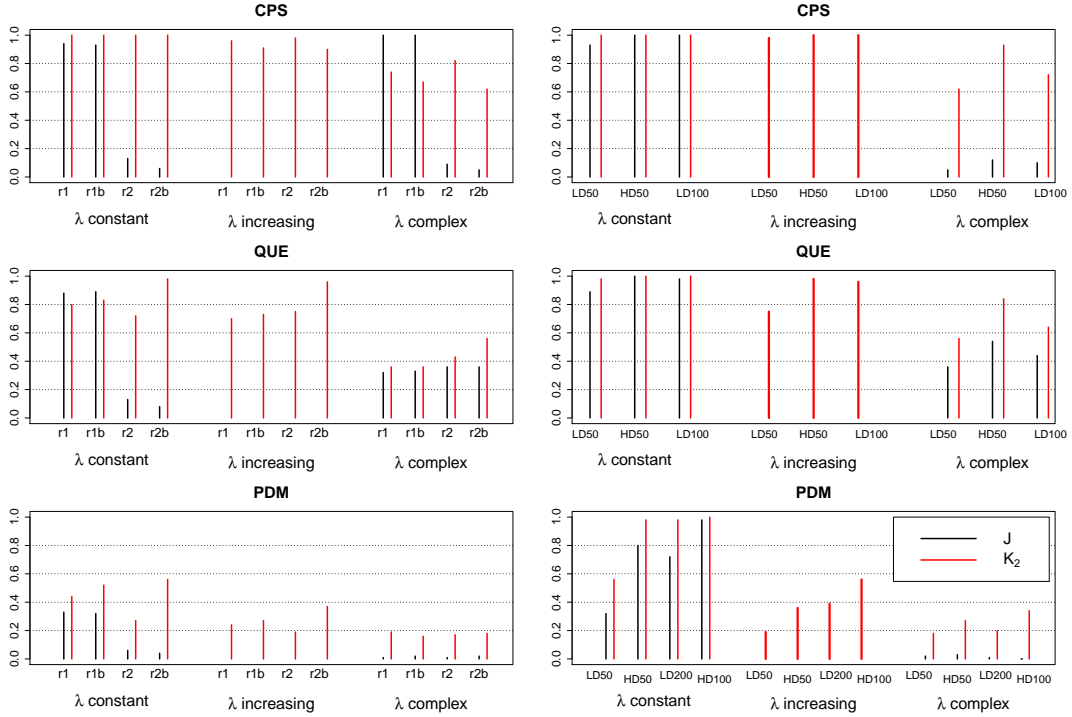


Figure 5. Power of the CROSS tests for $\alpha_1 = 0.05$. Left: Effect of the r -grid in setting LD50. Right: comparison of different settings with the selected r -grid.

for $\alpha_1 = 0.05$ with the different intensities under CPS, QUE and PDM, in HD50 and LD100. The power of the Normal test under CPS and QUE is good, specially in HD50, but slightly lower than the other tests. The power of the PaB test in NH processes is the highest except under QUE, where it is low with $\lambda^F(t)$. The power of K_2 is the highest in homogeneous PPs and satisfactory otherwise, except under PDM. A good property of the PaB test is that its power increases with the sample size much faster than in K_2 . For example, under QUE with $\lambda^F(t)$, the power of PaB increases to 0.70 in HD100 and 0.65 in LD200, approaching the K_2 counterparts, 0.87 and 0.72. On the other hand, the power of K_2 increases slowly under PDM with $\lambda^F(t)$, it is 0.34 in HD100 and 0.22 in LD200, while the PaB counterparts are 0.98 and 0.81.

To sum up, none of the tests shows the best performance in all settings. The estimated power of PaB and K_2 tests depends on the type of dependence but, if it is unknown, we suggest the use of the K_2 test in homogeneous PPs, and the PaB test if N_y follows a NH parametric model. Table 5 summarizes the tests which can be applied in each situation, and the recommended option if the dependence is unknown.

4.4. Size and power with estimated intensities

Most of the considered tests assume that $\lambda_y(t)$ and, in some cases, $\lambda_x(t)$, are known. This is not usual in real problems, but in many cases, they can be estimated from the observed process [26–28]. This section compares, in the settings in Section 4.3, the tests implemented with the true intensities $\lambda_x(t)$ and $\lambda_y(t)$, and their estimated counterparts. The homogeneous intensity is estimated as the mean number of points per time unit, n/T . In NH processes two estimators are used: a nonparametric approach

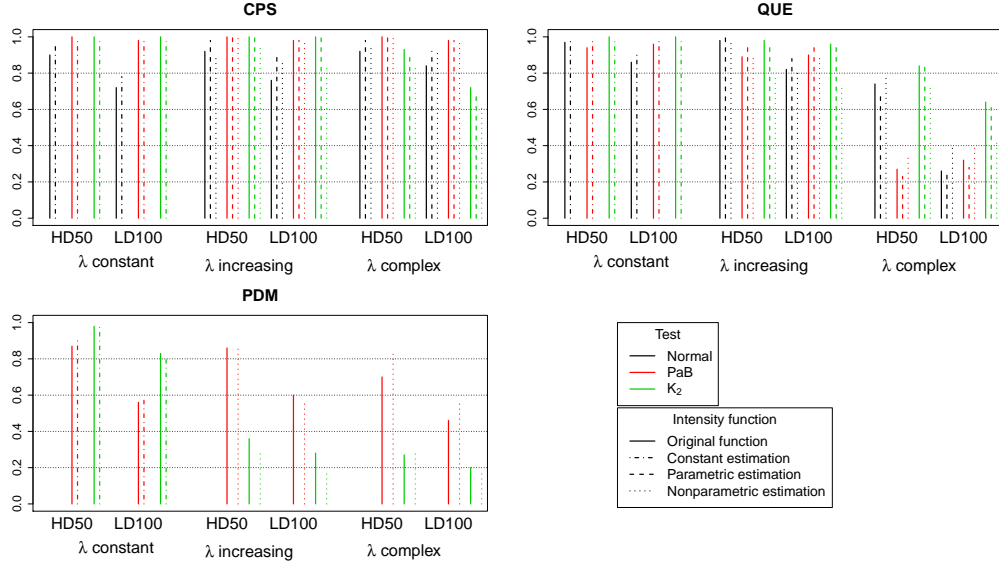


Figure 6. Power of the best tests for $\alpha_1 = 0.05$ with true and estimated intensities.

Table 5. Tests and conditions where they can be applied. The best option, under an unknown dependence model, is in bold, and tests with a low power in grey.

Assumptions - Intensity	Homogeneous	(log) Linear Intensity	General Intensity
$N_y \sim \text{Poisson}$ with $\lambda_y(t)$	Normal PaB K₂	Normal PaB <i>K₂</i>	Normal PaB <i>K₂</i> <i>J</i>
$N_y \sim \text{Parametric model}$ with $\lambda_y(t)$	PaB K₂	LoS <i>J</i> PaB <i>K₂</i>	PaB <i>K₂</i> <i>J</i>
Known $\lambda_x(t), \lambda_y(t)$	K₂	LoS <i>J</i> K₂	K₂ <i>J</i>
None		LoS	

that can be applied to any PP but requires m independent trajectories of PPs, and a parametric approach for Poisson processes that only requires one trajectory.

Parametric estimation (PE). In Poisson processes where the intensity is a function of known time-dependent covariates $\lambda(t) = \exp(\mathbf{x}^T(t)\beta)$, it can be estimated by conditional maximum likelihood [27]. This approach can be straightly applied to CPS and QUE models, where the marginal processes are Poisson, but not to the PDM model.

Nonparametric estimation (NPE). The NPE requires an i.i.d. sample of m trajectories $N_j = (t_{j1}, \dots, t_{jn_j})$. This is common in neurosciences, where the spike train issued from a neuron can be recorded m times, repeating the same stimulus. The mean of the empirical estimators $\hat{\lambda}_j(t)$ obtained from each N_j using moving averages or kernels, is a quite robust estimation.

A simulation study with estimated intensities, see Table 6, shows that the size of the Normal and PaB tests in homogeneous PPs and in NH with PE is even lower than the nominal value. The size of Normal and PaB tests with NPE and K_2 test is similar to the size with true intensities.

The previous Figure 6 compares the power with the true and the estimated intensities and shows that the results are equivalent, although with some minor differences. The power of all tests in homogeneous PPs and the Normal and PaB tests in NH processes is similar or slightly higher using the estimated intensities. The power of

Table 6. Size of the selected tests using estimated intensities. NPE are obtained with $m = 50$ repetitions.

α	λ^C			$\lambda^I(t)$						$\lambda^F(t)$					
	Normal	PaB	K_2	Normal		PaB		K_2		Normal		PaB		K_2	
				PE	NPE	PE	NPE	PE	NPE	PE	NPE	PE	NPE	PE	NPE
.05	.01	.02	.06	.01	.07	.01	.04	.06	.07	.00	.07	.02	.05	.07	.06
.10	.03	.07	.10	.02	.12	.04	.10	.12	.12	.01	.12	.04	.09	.12	.12

K_2 with the true intensity and PE is equivalent, but it decreases slightly with NPE, probably because this estimation is poorer and it is a nonparametric test. To sum up, the tests perform similarly with the true and the estimated intensities, provided that good estimators are used. The estimated size is lower or similar than its nominal value, and the power is similar or slightly higher. The PaB test seems to be more robust than K_2 to the estimation approach.

5. An application: occurrence of extreme heat events

The previous tests are used here to analyse the pairwise dependence between the occurrence of extreme heat events (EHEs) in three Spanish locations, Barcelona (B), Zaragoza (Z) and Huesca (H). The aim is to assess if the occurrence of the EHEs in those locations is independent or to identify the factors which explain the dependence. The locations are sited in a triangle with Barcelona on the East, around 250km away from the others, and Huesca located 67km to the North of Zaragoza. The daily maximum temperatures in the warm season (May to September) are available from 1951 to 2016. The days which are not observed in the three series are considered as missing observations, so that three series with 8262 complete observations are available. The extreme heat events may affect or not the three locations depending on if they are caused by large scale or local atmospheric situations.

An EHE is defined as a run of consecutive days where the temperature is over an extreme threshold, and its occurrence point is the day of maximum temperature in the run. The threshold is the 95th percentile of each series in the reference period (months of June, July and August in 1981-2010): 31.3, 37.8 and 36.4° C in Barcelona, Zaragoza and Huesca, where 121, 104 and 106 EHEs are observed respectively.

To quantify the dependence between the extremes, we calculate the correlation between the number of EHEs in intervals of length $l = 10$ days for a short-term dependence and $l = 100$ for medium-term dependence: $\varrho_{BZ}^{10} = 0.33$, $\varrho_{BH}^{10} = 0.31$ and $\varrho_{ZH}^{10} = 0.70$, and $\varrho_{BZ}^{100} = 0.55$, $\varrho_{BH}^{100} = 0.52$ and $\varrho_{ZH}^{100} = 0.68$. All of them are significantly different from zero, and show that there exist a pairwise dependence between the locations, and that it is stronger between Zaragoza and Huesca.

Our aim is to assess if this dependence can be explained by the seasonal behaviour of the phenomena (represented by an harmonic term) and a covariate representing the local temperature situation [29]. To that aim, a NH Poisson process whose intensity is a function of those covariates is fitted to each occurrence series, and all of them are satisfactorily validated. The second step is to study if the Poisson processes are independent, given the covariates which define their marginal intensities. If the tests do not reject the null hypothesis, it is concluded that the dependence between the EHE processes is explained by the considered covariates.

Since the validation analysis allows us to assume that the marginal processes are Poisson with the fitted intensities, the Normal, the PaB and the K_2 tests are used to assess the pairwise independence, although according to the rules in Section 4.3,

Table 7. P-values of the pairwise independence tests between the occurrence of EHEs in Barcelona, Zaragoza and Huesca.

	B-Z			B-H			Z-H		
	Normal	PaB	K_2	Normal	PaB	K_2	Normal	PaB	K_2
pv	.42	.20	.56	.30	.12	.40	.002	0	.002

the PaB test is the most adequate in this case. Table 7 shows the p-values of the three pairwise comparisons. The Normal test is applied with length $l = 20$, which guarantees the Normal approximation. The PaB and K_2 tests are implemented with 1000 repetitions and the last one with an r -grid of 50 values from 1 to around 500. It is also carried out with a grid focusing on short dependence, but a similar p-value is obtained. The three tests lead to the same conclusion the non rejection of independence in B-Z, and B-H (where and influential point is removed) and the rejection in Z-H. Since all the tests are built by keeping fixed N_x , the p-values resulting from changing the order of the locations are also calculated, and the decision is the same in all cases. Then, it is concluded that the occurrence of EHEs in Barcelona and Zaragoza and in Barcelona and Huesca is independent given the covariates, but those covariates do not explain all the dependence between Zaragoza and Huesca, the closest locations.

6. Conclusions

The assessment of independence between nonhomogeneous PPs is a frequent need in modelling problems. This work proposes three families of tests to assess the independence between two PPs, N_x and N_y . All the tests are built by keeping fixed their marginal structure but they are based on different assumptions, so that each one can be applied to different PPs (Poisson processes, PPs with a parametric model, NH processes, etc.) and dependence structures. All together they cover a wide range of situations appearing in real problems.

The first family assumes that N_y is a Poisson process with intensity $\lambda_y(t)$, and it is based on the fact that, under independence, the conditional distribution of $N_y|N_x$ is the distribution of N_y . It includes two tests based on a Poisson and an asymptotically Normal statistic. The second family includes two tests with the same statistic based on the close point distance, and p-values obtained from a parametric bootstrap and a Lotwick-Silverman approach. The first can be applied if N_y follows a parametric model, while the second does not need any assumption but is only valid for homogeneous PPs. The third family includes two tests whose statistics are based on a version of the spatial cross K and J functions for PPs in time. They are nonparametric and only need to know $\lambda_x(t)$ and $\lambda_y(t)$.

An extensive simulation study of the size and power of the tests is carried out. This study of the power covers four common dependence structures appearing in real modelling problems, so that it allows us to determine some practical rules to select the most adequate test in a particular situation. Each dependence model is analysed in different settings defined by two levels of dependence (low and high), three intensities (a constant, a log-linear time function and a complex function $\lambda^F(t)$), and three sample sizes (50, 100 and 200). The bulk of the comparisons is carried out using the true intensities, but a study comparing the results with the true and the estimated intensities shows that the size and power do not noticeably change when estimated intensities are used. The main conclusions are:

- In the considered settings, the Normal test, the parametric bootstrap test, and

the K_2 test show better results, in terms of size and power, than the other tests in their respective families.

- According to the simulation results, these three tests are well-sized and their power increases with the sample size and dependence level.
- The lowest power is obtained with $\lambda^F(t)$. The short-term dependence generated by CPS model is the most easily detected: the estimated power of PaB and K_2 tests is over 0.9 with $\bar{n}_x = 50$ and high dependence.
- None of the tests shows the highest power in all the settings, but some general rules about the recommended test in the situations where the dependence structure between the processes is unknown, can be stated:
 - The PaB test shows the most stable power in general, and it increases fast with the sample size. Its performance with estimated parameters is quite stable under different estimators. It is recommended for NH processes, if N_y follows a parametric model.
 - The K_2 test shows the best estimated power in homogeneous PPs. Its performance does not highly depend on the considered r -grid. It is recommended in homogeneous PPs, or if N_y does not follow a parametric model.
 - The Normal test is the fastest to compute and a good option with huge sample sizes, but it can only be applied to Poisson processes.

If a particular dependence model can be assumed, it would be better to base the election of the test on the results of the power for that model.

Notes

The data that support the findings of this study are available from the corresponding author upon request.

References

- [1] Trambly Y, Neppel L, Carreau J, et al. Non-stationary frequency analysis of heavy rainfall events in southern France. *Hydrological Sciences Journal*. 2013;58(2):280–294.
- [2] Abaurrea J, Asín J, Cebrián AC. Modeling and projecting the occurrence of bivariate extreme heat events using a non-homogeneous common Poisson shock process. *Stoch Environ Res Risk Assess*. 2015;29:309–322.
- [3] Møller J, Waagepetersen R. *Statistical inference and simulation for spatial point processes*. CRC Press; 2004.
- [4] Cronie O, van Lieshout MNM. Summary statistics for inhomogeneous marked point processes. *Ann Inst Statist Math*. 2016;68(4):905–928.
- [5] Schlather M, Ribeiro PJ, Diggle PJ. Detecting dependence between marks and locations of marked point processes. *J R Statist Soc B*. 2004;66(1):79–93.
- [6] Guan Y, Afshartous D. Test for independence between marks and points of marked point processes: a subsampling approach. *Environ Ecol Stat*. 2007;14:101–11.
- [7] Allard D, Brix A, Chadoeuf J. Testing local independence between two point processes. *Biometrics*. 2001;57:508–17.
- [8] Doguwa SI. On correlations of neuronal spike discharges. *Journal of Statistical Computation and Simulation*. 1993;47(1-2):11–24.
- [9] Tuleau-Malot C, Rouis A, Grammont F, et al. Multiple tests based on a gaussian approximation of the unitary events method with delayed coincidence count. *Neural computation*. 2014;26(7):1408–54.

- [10] Albert M, Bouret Y, Fromont M, et al. Bootstrap and permutation tests of independence for point processes. *Ann Statist.* 2015;43(6):2537–2564.
- [11] Dutilleul P. *Spatio-temporal heterogeneity: Concepts and analyses.* CUP; 2011.
- [12] Abaurrea J, Asín J, Cebrián AC. A bootstrap test of independence between three temporal nonhomogeneous Poisson processes and its application to heat wave modeling. *Environ Ecol Stat.* 2015;22(1):127–144.
- [13] Hennig C, Hausdorf B. Distance-based parametric bootstrap tests for clustering of species ranges. *Comput Statist Data Anal.* 2004;45(4):875 – 895.
- [14] Hennig C, Lin CJ. Flexible parametric bootstrap for testing homogeneity against clustering and assessing the number of clusters. *Stat Comput.* 2015;25(4):821–833.
- [15] Myllymäki M, Mrkvicka T, Grabarnik P, et al. Global envelope tests for spatial processes. *JR Statist Soc B.* 2017;79(2):381–404.
- [16] Li CS, Lee SM, Yeh MS. A test for lack-of-fit of zero-inflated negative binomial models. *Journal of Statistical Computation and Simulation.* 2019;89(7):1301–1321.
- [17] Caron PO. A comparison of the type i error rates of three assessment methods for indirect effects. *Journal of Statistical Computation and Simulation.* 2019;89(8):1343–1356.
- [18] Lotwick HW, Silverman BW. Methods for analysing spatial processes of several types of points. *JR Statist Soc B.* 1982;44:406–13.
- [19] Dixon PM. Ripley’s K function. In: El-Shaarawi AH, Piegorsch WW, editors. *Encyclopedia of environmetrics.* volume 3. Wiley Online Library; 2002. p. 1796–803.
- [20] Baddeley AJ, Møller J, Waagepetersen R. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica.* 2000;54:329–50.
- [21] Daley D, Vere-Jones D. *An introduction to the theory of point processes.* vol ii: General theory and structure. Springer-Verlag; 2008.
- [22] Lindskog P, McNeil A. Common Poisson shock models: applications to insurance and risk modelling. *ASTIN Bulletin.* 2003;33:209–38.
- [23] Neyman J, Scott EL. Statistical approach to problems of cosmology. *JR Statist Soc B.* 1958;20:1–43.
- [24] Keilson J, Servi L. Networks of nonhomogeneous $M/G/\infty$ systems. *J Appl Probab.* 1994; 31:157–68.
- [25] Isham V. Dependent thinning of point processes. *J Appl Probab.* 1980;17(4):987–95.
- [26] Carlos Simon de Miranda J, Morettin P. Estimation of the intensity of non-homogeneous point processes via wavelets. *Annals of the Institute of Statistical Mathematics.* 2011 02; 63:1221–1246.
- [27] Cebrián AC, Abaurrea J, Asín J. NH-Poisson: An R package for fitting and validating nonhomogeneous Poisson processes. *J Stat Softw.* 2015;64(6):1–25.
- [28] Dumanjug CF, Barrios EB, Lansangan JR. Bootstrap procedures in a spatial-temporal model. *Journal of Statistical Computation and Simulation.* 2010;80(7):809–822.
- [29] Abaurrea J, Asín J, Cebrián A. Modelling the occurrence of heat waves in maximum and minimum temperatures over Spain and projections for the period 2031-60. *Global and Planetary Change.* 2018;161:244 – 260.

Appendix A. Normal approximation of the Normal test

In Section 3.1 it was proved that the statistic O_{n_x} has an asymptotic Normal distribution under the null hypothesis of the test. Given the asymptotic character of the result, the following simulation study is performed to characterise the conditions where the Normal approximation can be reasonably used. The approximation depends on n_x , the number of points in N_x , and on the μ_i values, since the higher they are, the more Normal is the distribution of the variables Y_i , and the lower n_x is needed. The problem is that n_x and μ_i values cannot be completely fixed in a setting, since

Table A1. Minimum l and median of $\bar{\mu}$ leading to a valid Normal approximation of O_{n_x} .

Intensity	n_x	l	Median of $\bar{\mu}$
$\lambda_x = \lambda_y = \lambda^C$	50	26	.23
	100	16	.15
	200	8	.08
$\lambda_x(t) = \lambda_y(t) = \lambda^I(t)$	50	22	.24
	100	12	.15
	200	8	.10
$\lambda_x(t) = \lambda_y(t) = \lambda^F(t)$	50	14	.58
	100	4	.29
	200	2	.17

they are random (μ_i values are the integral of $\lambda_y(t)$ over the l -length intervals around the points t_{x_i} , which are random). However, we can fix the two main factors which determine μ_i ($\lambda_y(t)$ and l) and the expected value of n_x ,

$$\bar{n}_x = E(n_x) = E[N_x(T)] = \int_0^T \lambda_x(t) dt$$

and assuming that $\lambda(t)$ is constant in each time unit, $\bar{n}_x = \sum_{t=1}^T \lambda_x(t)$.

The simulation study aims to find, for different values \bar{n}_x , the minimum value $\bar{\mu} = \sum_{i=1}^{n_x} \mu_i / n_x$ that, in average, leads to a valid Normal approximation. That means that for a given $\lambda_y(t)$, we have to find the minimum l leading to a valid approximation.

Each setting is defined by $\lambda_x(t)$, $\lambda_y(t)$ and the value T which gives the considered \bar{n}_x . The three intensities described in Section 4.1, all with mean $\bar{\lambda}(t) = 0.01$, are considered. For simplicity, the same intensity is considered in the two processes. The steps of the validation are,

1. A trajectory N_{x1} of a Poisson process with intensity $\lambda_x(t)$ is generated in $(0, T]$.
2. An independent trajectory N_{yi} of a Poisson process N_y with intensity $\lambda_y(t)$ is generated in the same period. Since $\lambda_y(t)$ is usually unknown in real problems, the estimator $\hat{\lambda}_{yi}(t)$ is calculated using the approach by Cebrián et al. [27]. The value of the statistic O_{n_x} for N_{x1} and N_{yi} , using $\hat{\lambda}_{yi}(t)$, is obtained.
3. Step 2 is repeated 100 times, so that a sample of 100 values of the statistic under the null hypothesis is obtained.
4. Using the sample from Step 3, the Anderson-Darling test is applied to assess if the statistic under the null follows a $N(0, 1)$ distribution.

This process is repeated 1000 times. A setting leads to a valid Normal approximation if the number of rejections at level $\alpha = .01$ in the 1000 runs is lower or equal than 10.

Table A1 shows for the three intensities and for $\bar{n}_x = 50, 100, 200$, the minimum interval length which leads to a valid Normal approximation and the median of the $\bar{\mu}$ values in the 1000 runs. In homogeneous processes, the Normal approximation works for $\bar{\mu}$ values around 0.23 with $\bar{n}_x = 50$, but the value decreases to $\bar{\mu} = 0.08$ with $\bar{n}_x = 200$. In nonhomogeneous processes with $\lambda^F(t)$, the $\bar{\mu}$ value increases to 0.58 with $\bar{n}_x = 50$, but it decreases to less than 0.20 with $\bar{n}_x = 200$. The results confirm that to get a better Normal approximation, we can increase the mean sample size \bar{n}_x or increase the $\bar{\mu}$ values. Since the intensity of a process cannot be changed, to obtain high enough $\bar{\mu}$ values, longer l values have to be used.

Table B1. Correlation measures in LD and HD levels under the dependence models

Dep.	LD		HD	
	I_{50}	I_{200}	I_{50}	I_{200}
Int.				
CPS	.31	.32	.65	.65
MNS	.36	.71	.78	.88
QUE	.27	.65	.61	.88
PDM	.14	.30	.27	.61

The Normal approximation is quite robust to the considered estimator, since the same analysis was carried out using a simple empirical estimator of the intensity function, [27], and analogous results were obtained.

Appendix B. Correlation generated under dependence models

In order to compare the high and low dependence levels defined in each of the four dependence models, and to evaluate the difficulty to detect dependence in the different settings, it would be useful to quantify the dependence generated in each case. The considered independence tests are built by fixing the marginal structures, and since the same dependence models are used with λ^C , $\lambda^I(t)$ and $\lambda^F(t)$, the dependence level is only characterized in the homogeneous case, which is the easiest one. To somehow quantify the dependence between the PPs, we calculate the correlation coefficient between the number of points in N_x and N_y in intervals I_l of length l , $\rho_{xy, I_l} = Cor(X_{I_l}, Y_{I_l})$. Given the discrete character of X_{I_l} and Y_{I_l} , and that we aim to quantify not only the linear correlation but any type of dependence, the Spearman coefficient is used.

Table B1 shows $\hat{\rho}_{xy, I_l}$ in each dependence model, for their corresponding LD and HD settings, with intervals of length 50 and 200 (half and double of the mean interpoint distance) which measure the short and long term dependence, respectively. The sample size is always higher than 500 since periods of length $T = 100000$ are used. The estimator $\hat{\rho}_{xy, I_l}$ is calculated as the median of the sample correlations in 1000 runs. The results show that the dependence under each model is quite different. The CPS model only generates short term dependence so that the correlation in I_{50} and I_{200} is very similar. QUE and MNS models generate short and/or long term dependence depending on their parameters so that, under LD level, the correlation in interval I_{200} doubles the correlation in I_{50} , suggesting that mainly long term dependence is generated in that case. PDM model generates long term dependence, which is more difficult to be detected, even in the HD setting.

Appendix C. Results of the analysis of the power

C.1. Power of the tests

The values of the power for $\alpha_1 = 0.05$ plotted in all the figures in Section 4.2 are summarized in tables. The power for $\alpha_2 = 0.10$ is also shown in these tables.

Table C1 shows the numerical results plotted in Figure 3. The top table shows the power of the POISSON family (Poisson and Normal tests) with the automatic selection of l , under the dependence models CPS and QUE for $\alpha_1 = 0.05$ and $\alpha_2 = 0.10$ in the three settings LD50, HD50 and LD100 with the three intensities. The bottom table

Table C1. Power of the POISSON tests. Top: with an automatic selection of l . Bottom: Effect of l on the power of the Normal test, $\alpha_1 = 0.05$

Int. Test	λ^C				$\lambda^I(t)$				$\lambda^F(t)$				
	Poisson		Normal		Poisson		Normal		Poisson		Normal		
	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2	
CPS	LD50	.38	.49	.48	.57	.36	.48	.54	.63	.29	.39	.64	.71
	HD50	.79	.86	.90	.92	.79	.86	.92	.95	.58	.68	.92	.94
	LD100	.57	.67	.72	.78	.58	.68	.76	.84	.46	.55	.84	.88
QUE	LD50	.42	.56	.56	.68	.29	.44	.52	.62	.13	.22	.24	.35
	HD50	.92	.95	.97	.98	.91	.95	.98	.99	.39	.52	.74	.82
	LD100	.72	.83	.86	.92	.59	.71	.82	.88	.14	.22	.26	.36

Int.	setting	CPS				QUE			
		l=14	26	50	100	l=	26	50	100
λ^C	LD50	-	.93	.67	.34	.49	.58	.52	
	HD50	-	1	1	.82	1	1	.96	
	LD100	-	1	.91	.52	.77	.86	.82	
$\lambda^I(t)$	LD50	-	.89	.56	.30	.45	.52	.50	
	HD50	-	1	.98	.72	1	1	.95	
	LD100	1	.99	.87	.50	.71	.80	.78	
$\lambda^F(t)$	LD50	.66	.53	.42	.34	.25	.29	.31	
	HD50	.98	.93	.86	.77	.74	.77	.71	
	LD100	.84	.68	.55	.45	.27	.30	.30	

summarizes the power of the Normal test for $\alpha_1 = 0.05$ in the previous settings, for different values of l .

Table C2 shows the numerical results plotted in Figure 4, that is the power of the CLOSE family (PaB and LoS tests) under the four dependence models, for $\alpha_1 = 0.05$ and $\alpha_2 = 0.10$ in the usual settings and with the three intensities.

Table C3 shows the numerical results plotted in Figure 5, that is the power of the CROSS family (K_2 and J tests) under the dependence models CPS, QUE, and PDM, with the three intensities, and for $\alpha_1 = 0.05$ and $\alpha_2 = 0.10$. The top table shows the power of the CROSS tests in setting LD50 with the four considered r -grids, $r1$, $r1b$, $r2$ and $r2b$. The bottom table shows the power of the tests with the r grid suggested for the previous analysis under different settings: LD50, HD50, LD100, and LD200 and HD100 when necessary.

Table C2. Power of the CLOSE family

Int. Test	setting	λ^C				$\lambda^I(t)$		$\lambda^F(t)$	
		PaB		LoS		PaB		PaB	
		α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
CPS	LD50	.79	.87	.91	.96	.78	.87	.78	.89
	HD50	1	1	1	1	1	1	1	1
	LD100	.98	.99	1	1	.98	1	.98	1
QUE	LD50	.64	.79	.90	.97	.55	.71	.18	.29
	HD50	.94	.98	1	1	.89	.95	.27	.44
	LD100	.96	.99	1	1	.90	.97	.32	.48
PDM	LD50	.29	.43	.30	.45	.28	.44	.26	.36
	HD50	.87	.93	.88	.93	.86	.94	.70	.83
	LD100	.56	.70	.57	.70	.60	.73	.46	.62
MNS	LD50	.41	.62	.70	.85	.24	.43	.26	.48
	HD50	.95	.98	.97	.99	.90	.96	.65	.82
	LD100	.84	.93	.97	1	.70	.85	.46	.65

Table C3. Power of the CROSS family

Int.	Dep.	r-grid	$r1$		$r1b$		$r2$		$r2b$	
			α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
K_2	λ_C	CPS	1	1	1	1	1	1	1	1
		QUE	.80	.91	.83	.92	.72	.87	.98	1
		PDM	.44	.59	.52	.64	.27	.39	.56	.70
	$\lambda^I(t)$	CPS	.96	.98	.91	.95	.98	.99	.90	.94
		QUE	.70	.82	.73	.83	.75	.95	.96	.99
		PDM	.24	.35	.27	.39	.19	.34	.37	.48
	$\lambda^F(t)$	CPS	.74	.84	.67	.78	.82	.88	.62	.70
		QUE	.36	.50	.36	.51	.43	.68	.56	.71
		PDM	.19	.31	.16	.28	.17	.27	.18	.28
J	λ_C	CPS	.94 (.92)	.96 (.96)	.93	.96	.13 (.11)	.26 (.26)	.06	.18
		QUE	.88 (.87)	.94 (.93)	.89	.95	.13 (.14)	.40 (.42)	.08	.36
		PDM	.33 (.32)	.43 (.42)	.32	.42	.06 (.06)	.13 (.12)	.04	.09
	$\lambda^F(t)$	CPS	1 (1)	1 (1)	1	1	.09 (.07)	.34 (.33)	.05	.29
		QUE	.32 (.32)	.44 (.44)	.33	.46	.36 (.38)	.48 (.48)	.36	.49
		PDM	.01 (.01)	.02 (.02)	.02	.02	.01 (.01)	.03 (.02)	.02	.05

Int. Test	Dep. setting	λ^C				$\lambda^I(t)$		$\lambda^F(t)$			
		$J (r1b)$		$K_2 (r2b)$		$K_2 (r2)$		$J (r2b)$		$K_2 (r2b)$	
		α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
CPS	LD50	.93	.96	1	1	.98	.99	.05	.29	.62	.70
	HD50	1	1	1	1	1	1	.12	.50	.93	.96
	LD100	1	1	1	1	1	1	.10	.43	.72	.78
QUE	LD50	.89	.95	.98	1	.75	.95	.36	.49	.56	.71
	HD50	1	1	1	1	.98	1	.54	.65	.84	.92
	LD100	.98	.99	1	1	.96	1	.44	.56	.64	.78
PDM	LD50	.32	.42	.56	.70	.19	.34	.02	.05	.18	.28
	HD50	.80	.86	.98	.99	.36	.56	.03	.06	.27	.42
	LD200	.72	.81	.98	.99	.39	.54	.01	.01	.20	.33
	HD100	.98	.99	1	1	.56	.76	.00	.01	.34	.48

C.2. Power comparison with true and estimated intensities

This section shows the results of the simulation study in Section 4.4 to compare the power of the tests implemented with true and estimated intensities. Table C4 summarizes the values of the power plotted in Figure 6, that is the power of the best test in each family, *Normal*, *PaB* and K_2 , with the three intensities, in settings HD50 and LD100, for $\alpha_1 = 0.05$ and with true and estimated intensities. In the NH processes, the power using both the parametric and nonparametric estimators is compared.

Table C4. Power of selected tests using the true and the estimated intensities for $\alpha_1 = 0.05$

	Test Dep.	Normal		PaB			K_2		
		CPS	QUE	CPS	QUE	PDM	CPS	QUE	PDM
HD50	λ^C	.90	.97	1	.94	.87	1	1	.98
	$\hat{\lambda}^C$.96	1	1	.99	.91	1	1	.98
	$\lambda^I(t)$.92	.98	1	.89	.86	1	.98	.36
	$\hat{\lambda}^I(t)$ (PE)	.98	1	1	.96	-	1	.97	-
	$\hat{\lambda}^I(t)$ (NPE)	.90	.99	1	.89	.87	.96	.78	.28
	$\lambda^F(t)$.92	.74	1	.27	.70	.93	.84	.27
	$\hat{\lambda}^F(t)$ (PE)	.98	.70	1	.23	-	.91	.84	-
	$\hat{\lambda}^F(t)$ (NPE)	.94	.79	1	.34	.84	.84	.73	.29
LD100	λ^C	.72	.86	.98	.96	.56	1	1	.83
	$\hat{\lambda}^C$.78	.92	.98	.99	.58	1	1	.82
	$\lambda^I(t)$.76	.82	.98	.90	.60	1	.96	.28
	$\hat{\lambda}^I(t)$ (PE)	.89	.88	.98	.94	-	1	.96	-
	$\hat{\lambda}^I(t)$ (NPE)	.86	.81	.98	.90	.57	.83	.72	.18
	$\lambda^F(t)$.84	.26	.98	.32	.46	.72	.64	.20
	$\hat{\lambda}^F(t)$ (PE)	.92	.25	.98	.28	-	.70	.61	-
	$\hat{\lambda}^F(t)$ (NPE)	.92	.41	.99	.39	.56	.63	.43	.19