Bootstrap testing for first-order stationarity on irregular windows in spatial point patterns

Christine Kraamwinkel

University of Pretoria, South Africa

Inger Fabris-Rotelli*

University of Pretoria, South Africa

Alfred Stein

University of Pretoria, South Africa and University of Twente, Netherlands

Abstract

Kernel smoothing is commonly used in spatial point patterns to construct intensity plots. Kernels allow for visually and subjectively inferring on first-order stationarity. Formal objective tests exist for testing first-order stationarity that assume independence of spatial regions. We propose to extend inference for first-order stationary by using bootstrapping in existing hypothesis tests to deal with the violation of independence. More specifically we compare Poisson intensities from bootstrapped spatial quadrat samples, providing a test for first-order stationarity without violating the assumption of independence of the tests. Five hypothesis testing methods are investigated. The choice of grid mesh size and window shape used in these tests is discussed and guidance is provided through testing the power of the tests. The application considers the household locations in rural villages in Northern Tanzania as an unmarked point pattern. A clear effect of the village sizes on the relation between grid mesh size and confidence intervals of bootstrap sampling is shown. We conclude that bootstrapping provides a novel contribution to inference of first-order stationarity for spatial point patterns.

Keywords: first-order stationarity, hypothesis tests, bootstrap sampling, independence, stationarity, village household locations, Tanzania

1. Introduction

Understanding the intensity of a spatial point process is often an important aim of a spatial scientific study. For instance, spatial variation of house prices could reflect poverty patterns in an area. To establish clustering of a pattern we first need to eliminate alternative explanations such as spatially

*Inger Fabris-Rotelli

Email address: inger.fabris-rotelli@up.ac.za (Alfred Stein)

varying intensity. When it is known that intensity varies across a pattern, adjustments for this effect can be made to the tools used to investigate clustering. [1, 2]

Intensity is defined as follows. Let $\mathbf{X}$ be a point process in a two-dimensional space D. In the stationary case, for any subregion R of D, the intensity λ of a point process is taken as the expected number of points of $\mathbf{X}$ falling in R and is proportional to the area of R:

$$E\left[n\left(\mathbf{X}\cap R\right)\right] = \lambda\left|R\right| \tag{1}$$

where $n\left(\mathbf{X}\cap R\right)$ is the number of points in $\mathbf{X}$ within R and $|R|$ is the area of R. The intensity can therefore be considered as the average density of points per unit area where λ is a constant [1, 2, 3]. In the non-stationary case, the expected number of points falling within R, with spatial varying intensity function of the process, $\lambda\left(u\right)$, is defined [3] as

$$E\left[n\left(\mathbf{X}\cap R\right)\right] = \int_R \lambda\left(u\right) du.$$

According to [1], analysis often starts with the tentative assumption that a process is stationary and the validity of this assumption is assessed. In point processes, the empirical density of points,

$$\hat{\lambda} = \frac{n\left(\mathbf{x}\right)}{|W|}, \tag{2}$$

is an unbiased estimate of λ, assuming that $\mathbf{X}$ has stationary intensity. In this equation, $\mathbf{x}$ is the point pattern observed in a window W, $n\left(\mathbf{x}\right)$ is the number of points in W and $|W|$ is the area of W.

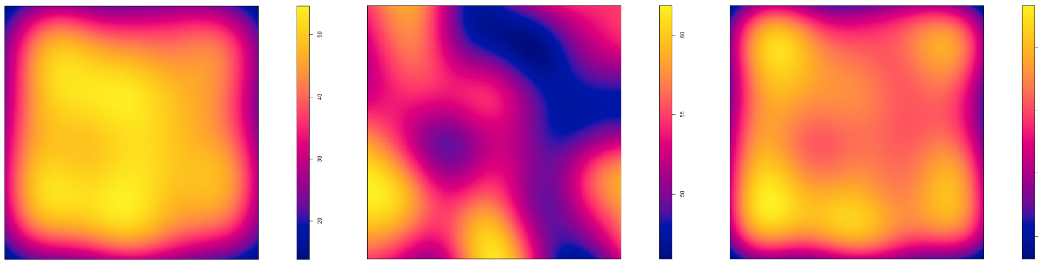
Before we are able to choose and fit an appropriate model to a point pattern, it is important to understand the underlying data structure by considering first- and second-order stationarity [4, 5, 6]. A first-order stationary process has the same intensity at different locations on a map. This implies that the expected number of observations is proportional to the size of the observation area. Stationarity is related to the expected number of pairs of points in those sections. To study second-order non-stationarity, we consider interactions between the points of the process. Complete spatial randomness (CSR) describes a stationary point pattern where the points are randomly distributed [2, 7]. Three possibilities can occur: points are spread randomly, i.e. without interactions, they are clustered, i.e. with positive interactions or they are regular, i.e. with negative interactions [1, 3, 7]. Regularity implies that points tend to avoid each other, leading to a pattern within the data with points distributed in a regular fashion. A clustered pattern, also referred to as an 'aggregated' point pattern, implies that points are grouped or clustered together. A random pattern point implies that the there is no pattern to the points, neither regular nor clustered. In this paper we focus on first-order stationarity.

In order to obtain an accurate impression of the degree of second-order stationarity in a point pattern, i.e. whether the pattern is clustered, regular or random, we require a sound knowledge of the process intensity. Several tests have been developed to test for second-order stationarity in spatial point processes. However, to our knowledge, tests for first-order stationarity in the spatial context are not yet well-developed, with researchers mostly relying on the visual inspection of density plots or assumptions about the data to decide whether the pattern is stationary or not [1]. These intensity plots are obtained through kernel estimation and are affected by the choice of whether or not to apply edge corrections. An example is shown in Figure 1 where kernel estimation was applied to a simulated first-order stationary point pattern with and without various edge corrections. If a judgement call had to be made regarding the first-order stationarity of this pattern using Figure 1b, a researcher might mistakenly classify the pattern as first-order non-stationary. The need for more formal testing is thus apparent. Tests for stationarity in point patterns compare intensities at different locations on the map, leading us to apply tests developed in the context of time intervals to this setting. These tests, originally applied to one-dimensional Poisson processes for statistically testing stationarity of $\lambda(t)$, where t is time, are now decribed for the spatial context. So we consider $\lambda(u)$ where u is a spatial location. These tests, however, are to our knowledge not widely applied in practice and their application has software limitations that are discussed later on.

Beside the five tests discussed in this paper, other tests for first-order stationarity have been developed in the past. Guan [8] extended the KPSS test[1] developed for time series first-order stationarity testing [9] to the spatial domain. This method does not have the strict requirement of independence of subregions within the domain which the tests discussed in this paper do but is restricted to rectangular domains. In addition, neither the exact nor approximate distribution for the asymptotic null test statistics is known and thus requires simulation to compute critical values. Zhang and Zhou [10] extended this method to arbitrary regions $S \subseteq \mathbb{R}^d$ and also derived a closed form expression for the asymptotic null distribution of the test statistic. The simulations and applications therein, however, only consider rectangular and circular regions. The use of the technique in practice for other regions e.g. a convex hull and other irregular shapes, while theoretically possible, are not trivial. In this paper we extend the current tests to better deal with any window shape in a trivial way.

The original five tests are based on choosing subregions within a domain and testing equal Poisson intensities across the subregions i.e. determining whether the intensities of K underlying spatial processes are the same [1, 12]. In this context, the K underlying spatial processes refer to subdividing the observation window into K quadrats and comparing the intensities of each quadrat. These tests however assume that the subregions are independent which is in spatial statistics not true by Tobler's law [13]. If first-order stationarity is present, the result could be misleading [1, 2]. While it is true that the tests such as Moran's I [14] and Geary's C [15], as well as variograms, their application does not

[1]Kwiatkowski-Phillips-Schmidt-Shin test



(a) Uncorrected (b) Uniform correction (c) Diggle's correction [11]

Figure 1: Kernel density estimates of the intensity of a simulated point pattern using various edge corrections and a Gaussian kernel.

confirm first-order stationarity. Such methods usually requires marked data as well. Here we consider unmarked data, that is, points characterised by their spatial locations only. Assuming dependence in spatial data is thus natural according to Toblers first law of geography when investigating testing for first-order stationarity. In the absence of spatial interaction i.e. independence, the technique in this paper is still valid.

The objective of this paper is to extend the methods of testing for first-order stationarity without violating the assumption of independence. Five hypothesis tests are discussed for this purpose, only two of which are currently available in the `spatstat` package in `R`, extending their use to deal with independence appropriately. These tests are discussed in Section 2 with an application to simulated point patterns in Section 3 and real-world point patterns of household locations in rural villages in Northern Tanzania in Section 4. The simulations are extended to irregular shaped windows and the independence assumption dealt with through bootstrap sampling of subregions of the windows.

2. Methodology

For data modeled by the Poisson distribution, counts are used to test for the equality of K intensities on K spatial regions. These tests are closely linked to the method of quadrat counts as the window W is divided into quadrats R_1, ..., R_K. The number of points within each quadrat, x_1, ..., x_K is then counted and the size of each quadrat, A_1, ..., A_K, is noted. Under the null hypothesis, if the areas of each quadrat are equal, $A = A_1 = ... = A_K$, and the unknown intensity is λ, then the counts x_i are independent Poisson random variables with equal mean $A\lambda$ [1]. The fixed observation areas need not be of equal size, but are generally assumed to be independent. Several studies have been performed to compare the size and power of numerous test statistics if $K = 2$ for both equal sampling frames (see [16, 17]) and unequal sampling frames (see [18, 19, 20]).

Consider the case where independence is assumed. We observe K independent random variables $X_1, X_2, ..., X_K$ and let x_i be the observed number of occurrences in a spatial Poisson process with an unknown rate λ_i in a sampling frame with known area A_i. Note that this notation makes allowances for areas A_i of different sizes. The density function of X_i is given by

$$f(x_i) = \frac{e^{-\lambda_i A_i} (\lambda_i A_i)^{x_i}}{x_i!}, \qquad \text{for} \quad i = 1,\ 2,\ \ldots,\ K.$$

The maximum likelihood estimate of λ_i is given by the empirical rate, $\hat{\lambda}_i = \frac{x_i}{A_i}$. The stationarity test for K independent Poisson variates is given by $H_0:\ \lambda_1 = \lambda_2 = \ldots = \lambda_K$ against $H_A:\ \lambda_i \neq \lambda_j$ for at least one combination $i,\ j \in \{1,\ \ldots,\ K\}$.

Under the assumption of independence, the likelihood function is given by

$$L(\lambda_1,\ \ldots,\ \lambda_K \mid x_1,\ \ldots,\ x_K) = e^{-\sum_{i=1}^{K} \lambda_i A_i} \frac{\prod_{i=1}^{K} (\lambda_i A_i)^{x_i}}{\prod_{i=1}^{K} x_i!}.$$

Under the null hypothesis, each X_i follows a Poisson distribution with parameter $\lambda_i = \lambda$, implying that the point pattern is stationary. The maximum likelihood estimator of λ is then given by

$$\hat{\lambda} = \frac{\sum_{i=1}^{K} x_i}{\sum_{i=1}^{K} A_i}. \tag{3}$$

Several statistics have been proposed to test for the equality of the rates of Poisson processes for time, i.e. in one dimension. Here we extend these tests, in terms of notation, to be applicable in the spatial domain. Chiu and Wang [12] provide the details of these tests but they do not apply the tests on spatial data and do not deal with the violation of independence. Here we do not state these tests as new but contribute to their correct application to spatial data. We state the tests here for completeness.

1. Pearson's χ^2-test [21] is given in the current context of a point pattern by

$$\chi^2_{\text{Pearson}} = \sum_{i=1}^{K} \frac{\left(x_i - \hat{\lambda} A_i\right)^2}{\hat{\lambda} A_i} = \frac{\sum_{i=1}^{K} A_i}{\sum_{i=1}^{K} x_i} \sum_{i=1}^{K} \frac{x_i^2}{A_i} - \sum_{i=1}^{K} x_i.$$

   This test can be applied in two different ways. Firstly, it can be used to test goodness-of-fit to the Poisson distribution assuming stationary intensity [22]. Secondly, it can be used to test for stationarity by assuming independence [23]. Since we are interested in testing for first-order stationarity we will be applying the second, so called the χ^2 test of stationarity. Under the null hypothesis, this test statistic has an asymptotic χ^2-distribution with $K - 1$ degrees of freedom. Traditionally, this approximation is acceptable if the expected number of points in each of the quadrats is at least 5 [21]. The power of the test depends on the size of the quadrats and is optimal when these are neither very large nor very small [1, 24, 25].

2. The likelihood ratio test statistic was originally proposed by Neyman and Pearson in 1928 [26] with the asymptotic distribution of the test derived by Wilks in 1938 [27]. The likelihood ratio is obtained as

$$\Lambda = \frac{L\left(\hat{\lambda},\ \ldots,\ \hat{\lambda} \mid x_1,\ \ldots,\ x_K\right)}{L\left(\hat{\lambda}_1,\ \ldots,\ \hat{\lambda}_K \mid x_1,\ \ldots,\ x_K\right)}.$$

It is convention to take that $0/0 = 0$ in case the sum of the x_i's equals 0. Based on this test statistic, Chiu and Wang [12] developed a likelihood ratio test statistic to test for the equivalence of Poisson intensities, namely

$$LB = -2\Lambda = 2\left(\sum_{i=1}^{K} x_i \ln \frac{x_i}{A_i} - \sum_{i=1}^{K} x_i \ln \frac{\sum_{i=1}^{K} x_i}{\sum_{i=1}^{K} A_i}\right),$$

which under the convention that $0 \ln 0 = 0$, has an asymptotic χ^2-distribution with $K-1$ degrees of freedom. Rao and Chakravarti [28] considered the particular case where the time periods (in our case area) are of length 1, however Chiu and Wang [12] were the first to discuss the more general case.

3. In 1948, the score test was introduced by Rao as an alternative to the likelihood ratio test [29]. This test is simpler than the likelihood ratio test as it only requires estimation under the null hypothesis. This is particularly applicable when the success of an experiment depends on disproving the null hypothesis. The test is computationally quite simple but was only put to serious use in the 1980s when its use was popularized in the field of Econometrics [30]. In the context of point processes, this test statistic is given by

$$SC = u\left(\hat{\lambda}, \ldots, \hat{\lambda}\right)' I\left(\hat{\lambda}, \ldots, \hat{\lambda}\right)^{-1} u\left(\hat{\lambda}, \ldots, \hat{\lambda}\right) = \left(\frac{\sum_{i=1}^{K} x_i}{\sum_{i=1}^{K} A_i}\right)^2 \sum_{i=1}^{K} \frac{A_i^2}{x_i} - \sum_{i=1}^{K} x_i$$

where

$$u\left(\lambda_1, \ldots, \lambda_K\right) = \left(\frac{\partial \ln L\left(\lambda_1, \ldots, \lambda \mid x_1, \ldots, x_K\right)}{\partial \lambda_1}, \ldots, \frac{\partial \ln L\left(\lambda_1, \ldots, \lambda \mid x_1, \ldots, x_K\right)}{\partial \lambda_K}\right)',$$

$$I\left(\lambda_1, \ldots, \lambda_K\right) = \left[-\frac{\partial^2 \ln L\left(\lambda_1, \ldots, \lambda \mid x_1, \ldots, x_K\right)}{\partial \lambda_i \partial \lambda_j}\right]_{i,j=1,\ldots,K}.$$

To avoid division by zero it is common to add 0.5 to every x_i whenever one of the observed x_i is zero. Under the stationarity hypothesis, SC has an asymptotic χ^2-distribution with $K-1$ degrees of freedom. Ng and Cook [31] investigated score tests of stationarity for Poisson processes with equal observation periods. Chiu et al. [12] performed a comparative study to consider the stationarity of $K \geq 3$ Poisson count data with different observation periods.

4. In 1966, Potthoff et al. [21] proposed two test statistics that deal with the stationarity of Poisson count data with different observation periods. The test statistic was developed in the context of

time intervals, t_i, but we will consider the case of areas, A_i. The first of these is the VT test statistic based on the statistic

$$V = \left(\sum_{i=1}^{K} A_i\right) \sum_{i=1}^{K} \frac{x_i (x_i - 1)}{A_i}$$

where

$$\frac{V - \left(\sum_{i=1}^{K} x_i\right)\left(\sum_{i=1}^{K} x_i - 1\right)}{\sqrt{2(K-1)\left(\sum_{i=1}^{K} x_i\right)\left(\sum_{i=1}^{K} x_i - 1\right)}}$$

follows the standard normal distribution asymptotically under the null hypothesis. A more mathematically refined approximation can be found as

$$VT = eV + f \sim \chi^2_{\nu_1} \tag{4}$$

with degrees of freedom

$$\nu_1 = e^2 (K-1)\left(\sum_{i=1}^{K} x_i\right)\left(\sum_{i=1}^{K} x_i - 1\right).$$

The constants in Equation 4 are given by

$$e = \frac{2(K-1)}{\left(\sum_{i=1}^{K} A_i\right)\left(\sum_{i=1}^{K} \frac{1}{A_i}\right) - 3K + 2 + 2(K-1)\left(\sum_{i=1}^{K} x_i - 2\right)}$$

and

$$f = e\left[(K-1)e - 1\right]\left(\sum_{i=1}^{K} x_i\right)\left(\sum_{i=1}^{K} x_i - 1\right)$$

and are determined in such a way that VT and $\chi^2_{\nu_1}$ have the same first three moments.

5. The second test statistic proposed by Potthoff et al. [21] is also an asymptotic χ^2-test which is the locally most powerful test against a selected alternative if λ is known [12, 21]. As with the VT test, we will consider areas, not time intervals. The UT test statistic is based on the statistic

$$U = \sum_{i=1}^{K} x_i^2 - \sum_{i=1}^{K} x_i - 2\lambda \sum_{i=1}^{K} A_i x_i$$

and

$$\frac{U + \lambda^2 \sum_{i=1}^{K} A_i^2}{\sqrt{2\lambda^2 \sum_{i=1}^{K} A_i^2}} \tag{5}$$

follows the standard normal distribution asymptotically under the null hypothesis.

A more mathematically refined approximation can be found as

$$UT = gU + h \sim \chi^2_{\nu_2} \tag{6}$$

with degrees of freedom

$$\nu_2 = g^2 \lambda^2 \sum_{i=1}^{K} A_i^2.$$

The constants in Equation 6 are given by

$$g = \frac{\sum_{i=1}^{K} A_i^2}{\sum_{i=1}^{K} \frac{A_i^2}{2} + \lambda \sum_{i=1}^{K} A_i^3}$$

and

$$h = g\left(g+1\right) \lambda^2 \sum_{i=1}^{K} A_i^2$$

and are determined in such a way that UT and $\chi^2_{\nu_2}$ have the same first three moments. Potthoff and Whittinghill [21] proposed an alternative estimator, λ^*, for the case where λ is unknown. This estimator is the value of λ that minimizes Equation 5 and is given by

$$\lambda^* = \sqrt{\frac{\sum_{i=1}^{K} x_i^2 - \sum_{i=1}^{K} x_i}{\sum_{i=1}^{K} A_i^2}}.$$

Chiu and Wang [12] confirmed through a simulation study that the UT test is sometimes too conservative in small experiments and shows a marked loss of power in such a situation. This handicap is not shared by the Pearson's χ^2 and VT tests.

These tests for first-order stationarity all have χ^2-distributions that are valid only asymptotically. Tests using p-values calculated by these asymptotic distributions may encounter size distortions which can be addressed through the use of parametric bootstrapping. This approximates the p-value through Monte Carlo resamples generated according to the stationarity hypothesis [12, 32]. The procedure for approximating p-values is as follows. The test statistic, denoted by τ, is calculated based on the observed counts. Next we draw a sample of K independent Poisson variates with mean $A_i \hat{\lambda}$, $i = 1, \ldots K$, where $\hat{\lambda}$ is the maximum likelihood estimate of the intensity, that is under H_0, see Equation 3. The test statistic is calculated for this sample and is denoted by τ_1^*. This is repeated B times so that we have $\tau_1^*, \ldots, \tau_B^*$. The bootstrap p-value as proposed by Davison and Hinkley [32] is

$$\text{bootstrap } p\text{-value} = \frac{\#\left\{i : \tau_i^* \geq \tau\right\} + 1}{B+1}. \tag{7}$$

We reject the null hypothesis of first-order stationarity if the bootstrap p-value is less than or equal to the nominal significance level α. An alternative bootstrap p-value can be calculated without adding 1 in the numerator and denominator in Equation 7 (see [33, 34]), however this is less accurate. Here we will follow the original testing procedure.

Hope [35] illustrated that, in the case that there is no nuisance parameter, the power loss resulting from using parametric bootstrap tests is slight compared to the corresponding uniformly most powerful test. This implies that B need not necessarily be large. Besag and Diggle [36] suggested that at $\alpha = 0.05$, $B = 99$ would be adequate. A simulation study conducted by Davison and Hinkley [32] showed that for $\alpha \geq 0.05$, the loss of power with $B = 99$ is not serious but that $B = 999$ should generally be a safe choice. Here we obtain the approximate p-values from parametric bootstrap samples of size 999.

2.1. Accounting for dependence in spatial data

The hypothesis tests discussed above assume that counts used in the calculation of the test statistics are obtained from areas that are independent of each other. Points in a spatial pattern that are close together are however unlikely to be independent as already stated, violating this assumption. In some cases, points that are far apart may also be considered dependent. An example of this occurs when analysing anisotropic weather patterns along a coastline or property damage along a fault line, when in one direction short distances will exhibit independence, while in a different direction we observe dependence. If however points that are further apart could validly be considered independent, the tests give valuable insight into the first-order stationarity. Some level of intuition or judgement should be applied by the researcher when deciding whether the assumption of independence is justifiable, as well as deciding on the number of areas to compare. Here we consider alternative methods of applying these tests such that the independence assumption is no longer violated, and dependence can be assumed.

We propose here an alternative for dealing with the independence violation by randomly sampling a subset of the quadrats and using these to calculate a test statistic. This is repeated 99 times, thereby calculating 99 bootstrap test statistics, each time drawing a different sample of 50% of the quadrats and calculating a test statistic for each of these. If a quadrat has 0 counts in it, its count is recorded as 0.5 to avoid division by 0. To balance this all other counts are also increased by 0.5. A sample of $B = 999$ here would more likely violate independence since certain samples will have dependence e.g. if the 50% are selected as the top half of the quadrats.

3. Simulation study

The power of the proposed technique is tested on a stationary as well as non-stationary patterns on a rectangular window and convex window.

3.1. A stationary point pattern

Consider the point pattern shown in Figure 2. This pattern was simulated from a stationary Poisson point process with intensity $\lambda = 50$ generated in a square window with a length of 4 units. It can be seen from this pattern that the average number of points in areas of equal size are roughly the same.

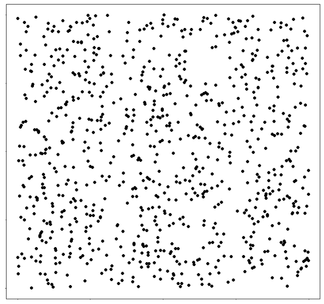


Figure 2: Simulated point pattern from a stationary Poisson point process with intensity $\lambda = 50$, containing 844 points in a window of 16 square units.

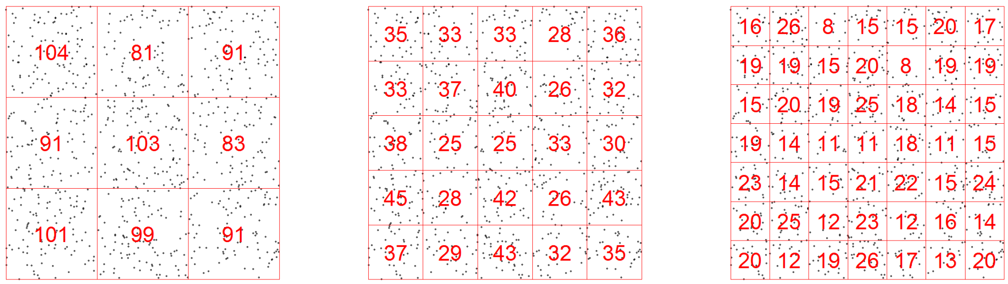
	Expected count per quadrat	χ^2	LR	SC	VT	UT
3×3	93.78	0.631	0.626	0.617	0.631	0.631
5×5	33.76	0.414	0.427	0.469	0.414	0.414
7×7	17.22	0.185	0.170	0.178	0.184	0.185

Table 1: The p-values for the hypothesis tests applied to the simulated stationary point pattern on a square window without adjusting for spatial dependence

The simulated stationary point pattern is divided into quadrats of different sizes to assess the first-order stationarity. In order to illustrate the application of the five hypothesis tests to a point pattern, we divide the simulated pattern into $n \times n$ size quadrats where $n = 3, 5, 7$ as illustrated in Figure 3. The five hypothesis tests, while not correcting for violation of independence, were applied to the simulated point pattern with results shown in Table 1. For each of the grid sizes, the expected number of points in each quadrat is greater than 5. It is clear that for all grid sizes, each of the tests indicate that the pattern was generated from a stationary point process.

The asympototic power [37] of the five hypothesis tests was investigated for a stationary pattern such as the one in Figure 2. and is illustrated in Table 2. It can be seen that as the grid size increases (more quadrats) the power increases. The asymptotic power of the tests was also tested on the convex hull window as shown in Figure 4. The results are presented in Table 3. Table 3 shows that as the number of quadrats increases the power decreases but it is not consistent. The use of these tests on irregular type windows is thus not advised without adjusting for dependency.

The choice of the size of the quadrats is a difficult one. There is a tradeoff between bias and variability. While a large number of quadrats reduce the relative error of the counts n_j, it also smooths out the variation in intensity within each of the quadrats. For non-square windows the choice warrants more thought as well as the quadrats on the boundaries will be incomplete. Here a finer grid



(a) 3×3 (b) 5×5 (c) 7×7

Figure 3: Number of points observed within quadrats for the simulated stationary point pattern in Figure 2 divided into different size grids.

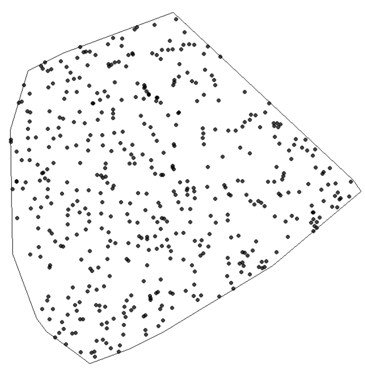


Figure 4: The simulated stationary pattern with $\lambda = 50$ on a convex hull window

n	χ^2	LR	SC	VT	UT
3	0.3626	0.3657	0.3745	0.3612	0.341
5	0.8487	0.8472	0.8582	0.8419	0.8323
7	0.9922	0.9936	0.9986	0.9902	0.9892

Table 2: Asympototic powers obtained for the stationary simulated pattern with intensity $\lambda = 50$ on the square window without adjusting for spatial dependence.

n	K_{orig}	K	χ^2	LR	SC	VT	UT
3	9	9	0.867867	0.65949	0.741527	0.853934	0.431437
5	23	19	0.198725	0.202075	0.218456	0.191812	0.174772
7	42	34	0.269062	0.27485	0.303173	0.258531	0.245134

Table 3: Asympototic power values obtained for the stationary simulated pattern with intensity $\lambda = 50$ on the convex hull window without adjusting for spatial dependence. K_{orig} is the number of quadrats and K the number of quadrats that have expected counts larger than 5.

	Expected count per quadrat	χ^2	LR	SC	VT	UT
3×3	92.44	0.001	0.001	0.001	0.001	0.001
5×5	34.67	0.001	0.001	0.001	0.001	0.001
7×7	17.33	0.001	0.001	0.001	0.001	0.001

Table 4: The p-values for the hypothesis tests applied to the simulated non-stationary point pattern in Figure 5(a) without adjusting for spatial dependence.

will always be better, however going too fine results in losing the structure.

Next we look at the power of the test for first-order non-stationary point patterns.

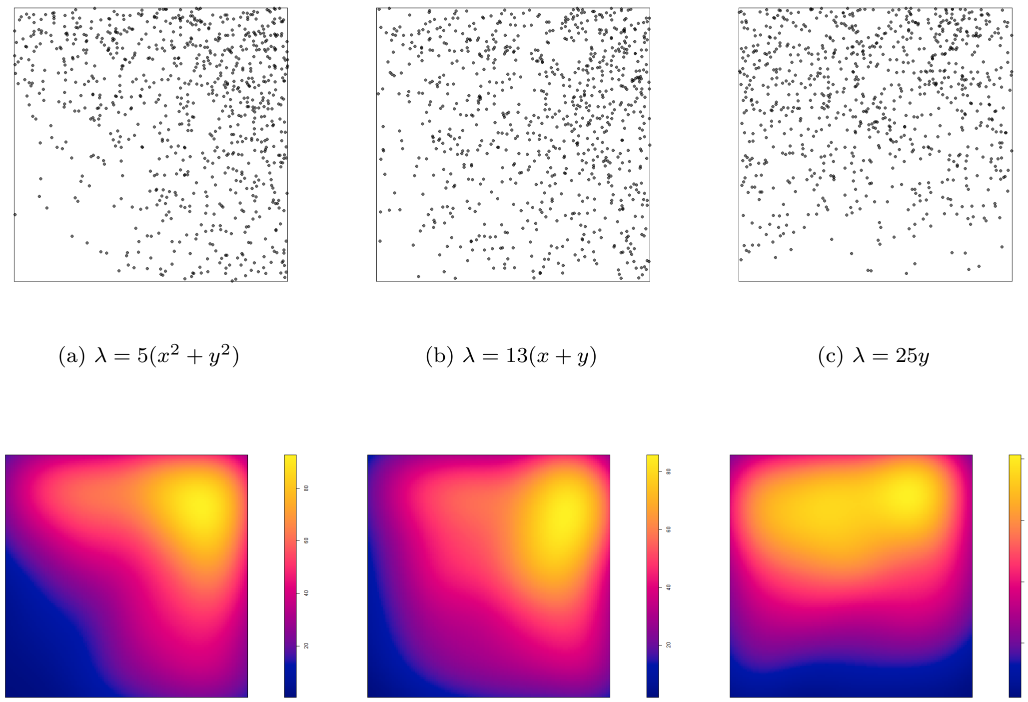
3.2. First-order non-stationary point patterns

Figure 5 shows examples of point patterns with non-stationary intensity functions simulated over a square of 16 square units with various non-stationary structures, namely quadratic, radial and gradient. The kernel density estimates for these patterns are shown in teh second row of Figure 5. The parametric bootstrap p-values for the five hypothesis tests are shown in Tables 4, 5 and 6. Here once again the violation of indepedence is not considered. For each of the grid sizes, the expected number of points in each quadrat is greater than 5. All p-values are found to be significant, indicating that the point patterns exhibit non-stationary intensities. The power of the tests for the three non-stationary cases were also obtained and were all obtained as 1.

The power for these non-stationary cases on the convex hull window are shown in the lefthand sides of Tables 8 and 9. In Table 8 the powers are good except for the coarse grid case ($n = 3$). In Table 9 the intensity of the simulated points was halved. It is seen that the powers then drop and are less consistent on this convex hull window compared to the square window.

3.3. Dealing with dependence: bootstrap sampling

To deal with the inherent dependence in spatial data the quadrats are sampled. Doing this with bootstraping provides a mechanism to avoid dependence of quadrats. The power of the five hypothesis



(a) $\lambda = 5(x^2 + y^2)$ (b) $\lambda = 13(x + y)$ (c) $\lambda = 25y$

(d) $\lambda = 5(x^2 + y^2)$ (e) $\lambda = 13(x + y)$ (f) $\lambda = 25y$

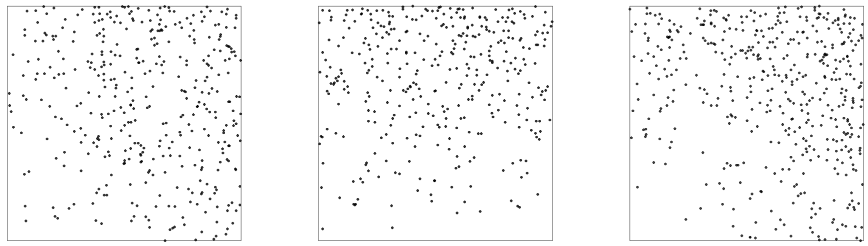
Figure 5: Simulated non-stationary point patterns from a Poisson process with different intensity functions, containing 832, 827 and 830 points respectively in windows of 16 square units, with their fitted intesities shown on the second row.

	Expected count per quadrat	χ^2	LR	SC	VT	UT
3×3	91.89	0.001	0.001	0.001	0.001	0.001
5×5	33.08	0.001	0.001	0.001	0.001	0.001
7×7	16.88	0.001	0.001	0.001	0.001	0.001

Table 5: The p-values for the hypothesis tests applied to the simulated non-stationary point pattern in Figure 5(b) without adjusting for spatial dependence.

	Expected count per quadrat	χ^2	LR	SC	VT	UT
3×3	92.22	0.001	0.001	0.001	0.001	0.001
5×5	33.20	0.001	0.001	0.001	0.001	0.001
7×7	16.94	0.001	0.001	0.001	0.001	0.001

Table 6: The p-values for the hypothesis tests applied to the simulated non-stationary point patternin Figure 5(c) without adjusting for spatial dependence.



(a) $\lambda = 5(x^2 + y^2)$ (b) $\lambda = 13(x + y)$ (c) $\lambda = 25y$

Figure 6: Simulated non-stationary point patterns from a Poisson process with different intensity functions with a weaker average intensity $\lambda = 25$

tests was investigated as follows [38]. A stationary pattern was simulated 100 000 times yielding 100 000 bootstrap test statistics for each grid size 3×3, 5×5 and 7×7. Then 1000 simulations of each of the three non-stationary patterns in Figure 5 were obtained. The test statistics for each case are compared to the p-values obtained using the 100 000 simulations in the first step.

The non-stationary patterns in Figure 5 are dense patterns. The average intensity was thus halved and the power study redone to investigate the effect of less points. The simulations were thus done for the sample patterns in Figure 5 for an average intensity of $\lambda = 50$ as well as a $\lambda = 25$. Examples of the weaker simulated patterns are shown in Figure 6. The power of the tests are shown in Table 7 for comparison. For the first case the power values in Table 7 show high power. The second case retains high power. The case $n = 7$ for radial non-stationarity shows weaker power however.

Under the assumption of stationarity, Equation 1 implies that the expected count in each quadrat is proportional to the area of the quadrat. This in turn implies that the average intensity given by Equation 2 is an unbiased estimator for the stationary intensity λ. We can therefore also compare the estimated intensities in different sized and shaped quadrats, should we wish, in order to make an inference about the stationarity of a point pattern, although the a regularly structured grid is important. A finer grid will result in more robustness at the boundary.

In order to assess the capability of bootstrap sampling in non-rectangular windows, the power tests were repeated on the convex window in Figure 4. The results are presented in Tables 8 and 9. The powers for the larger grid sizes (larger number of quadrats) are good compared to testing without bootstrapping. The powers are slightly less, but the importance of dealing with the inherent dependence has been dealt with.

	Using all quadrats					Bootstrap sampling				
n	χ^2	LR	SC	VT	UT	χ^2	LR	SC	VT	UT
$\lambda = 50$										
Radial Non-stationarity										
3	1	1	1	1	1	0.975	1	1	0.975	0.975
5	1	1	1	1	1	0.999	1	1	0.999	0.999
7	1	1	1	1	1	1	1	1	1	1
Gradient Non-stationarity										
3	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1
Quadratic Non-stationarity										
3	1	1	1	1	1	0.998	1	1	0.998	0.998
5	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1
$\lambda = 25$										
Radial Non-stationarity										
3	1	1	1	1	1	0.89	1	1	0.89	0.89
5	1	1	1	1	1	0.945	1	1	0.944	0.944
7	0.995	0.973	0.951	0.996	0.995	0.892	0.986	0.965	0.901	0.891
Gradient Non-stationarity										
3	1	1	1	1	1	0.981	1	1	0.981	0.981
5	1	1	1	1	1	0.999	1	1	0.999	0.999
7	1	0.999	1	1	1	0.99	1	1	0.997	0.991
Quadratic Non-stationarity										
3	1	1	1	1	1	0.985	1	1	0.985	0.985
5	1	1	1	1	1	1	1	1	1	1
7	1	0.998	1	1	1	0.994	1	1	0.998	0.994

Table 7: Power values obtained for the three non-stationary patterns with a strong and weak average intensity on the rectangular window.

	Using all quadrats					Bootstrap sampling				
n	χ^2	LR	SC	VT	UT	χ^2	LR	SC	VT	UT
Radial Non-stationarity										
3	0	0	0.001	0	0	0	0	0.001	0	0.004
5	1	1	1	1	1	0.967	1	1	0.971	0.946
7	1	0.998	1	1	1	0.947	0.997	0.998	0.95	0.943
Gradient Non-stationarity										
3	0	0	0.004	0	0.003	0	0	0.003	0	0.003
5	1	1	1	1	1	0.995	1	1	0.994	0.989
7	1	1	1	1	1	0.993	1	1	0.996	0.988
Quadratic Non-stationarity										
3	0	0	0	0	0.211	0	0	0	0	0.029
5	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1

Table 8: Power values obtained for the three non-stationary patterns with a strong average intensity on the convex hull window in Figure 4

	Using all quadrats					Bootstrap Sampling Approach				
n	χ^2	LR	SC	VT	UT	χ^2	LR	SC	VT	UT
Radial Non-stationarity										
3	0	0.001	0.005	0	0.05	0.118	0	0.004	0.116	0.122
5	0.969	0.96	0.956	0.964	0.959	0.741	0.956	0.955	0.755	0.717
7	0.877	0.739	0.803	0.886	0.857	0.585	0.73	0.794	0.462	0.552
Gradient Non-stationarity										
3	0	0.003	0.012	0	0.212	0.152	0	0.012	0.15	0.177
5	0.998	0.986	0.999	0.999	0.994	0.896	0.99	0.996	0.917	0.883
7	0.956	0.779	0.937	0.973	0.947	0.737	0.786	0.94	0.776	0.572
Quadratic Non-stationarity										
3	0	0.004	0.003	0	0.674	0.279	0.007	0.003	0.275	0.311
5	1	0.999	1	1	1	0.986	0.999	1	0.989	0.984
7	1	0.978	0.999	1	0.999	0.941	0.976	0.998	0.967	0.937

Table 9: Power values obtained for the three non-stationary patterns with a weak average intensity on the convex hull window in Figure 4

4. Application of tests to household locations in rural African villages

The spatial analysis of houses in more rural locations is an interesting one since the building of such houses does not usually follow the strict guidelines of government and the road network. The village is of interest as intensive rabies eradication campaigns have started where the efficiency of the interventions is affected by the accessibility of the houses and hence of stationarity in their pattern [39]. Yang and Jia [40] study such a spatial pattern in China and further interest is shown in India[2] and Turkey[3]. We thus look at the spatial patterns of such households locations in four rural villages in the Mara province, Northern Tanzania[4]. The locations of these households are shown in Figures 7 and 8 along with the convex hull windows used for analysis. The choice of window in such cases is complicated as the natural process responsible for the position of houses is one of need, access, culture and simplicity, not one based on ultimately planning for a bigger city with service facilities. The convex hull, the smallest convex set containing all the points of X, is however a common technique for irregular areas. Chosing a window for geographical data is a complicated matter. The convex hull is a natural alternative to rectangular windows if the point pattern does not naturally follow a rectangular window, as the convex hull ensures all distances (Euclidean) between points are within the window. Ripley and Rasson discuss the convex hull and its mathematics in detail [41, 42]. The mathematics of the Euclidean distance measure is thus not violated.

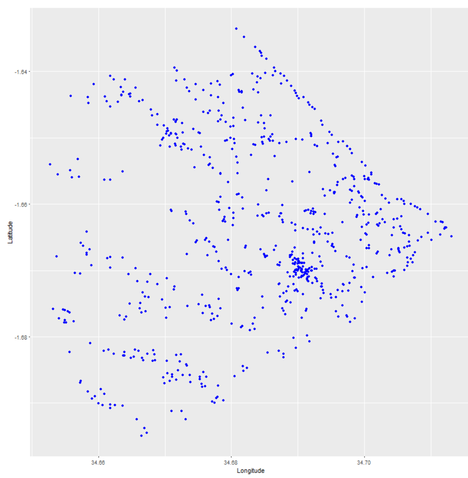
In order to apply the tests described in the previous section, the windows were divided into $n \times n$ grids with $n = 3, 5, 7$. These divisions with quadrat counts are shown in Figure 9.

This concept is illustrated in Figure 10. Since the quadrats that are used in the calculations are not directly adjacent to each other, it is less likely that the number of points in one quadrat is strongly dependent on the number of points in a non-adjacent quadrat; and the dependence will decrease with distance. Of course, this cannot be stated in an equal way for every spatial data set and intuition does need to be involved. The level of strength of the dependence will differ for different data. For these tests the parametric bootstrap p-values are obtained as follows. First a stationary Poisson point pattern is generated on the same window as the original pattern using the maximum likelihood estimate obtained for the intensity of the original pattern. The quadrats in the simulated pattern corresponding to those used to calculate the test statistic for the original pattern, are then used to calculated the first bootstrap test statistic. The quadrats included in the calculation of the test statistic can lead to different conclusions as is illustrated in Table 10, in which odd and even quadrats are illustrated.

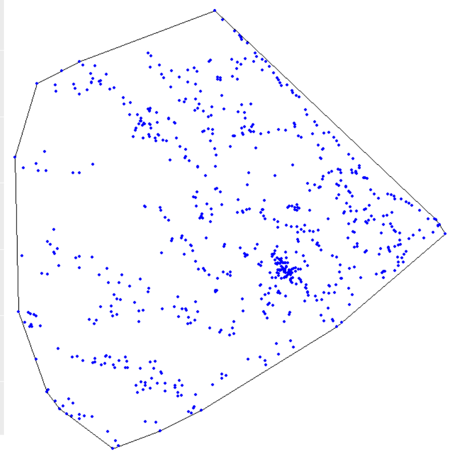
[2] http://www.preservearticles.com/2012013022140/short-essay-on-rural-settlement-patterns-in-india.html (Accessed on 9 April 2018)

[3] https://archnet.org/publications/3964 (Accessed 9 April 2018)

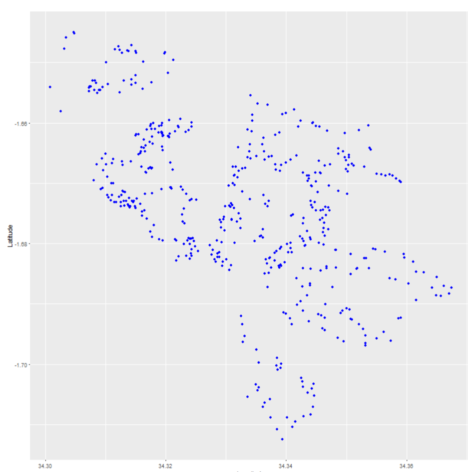
[4] http://www.gla.ac.uk/researchinstitutes/bahcm/staff/katiehampson/, http://www.katiehampson.com/



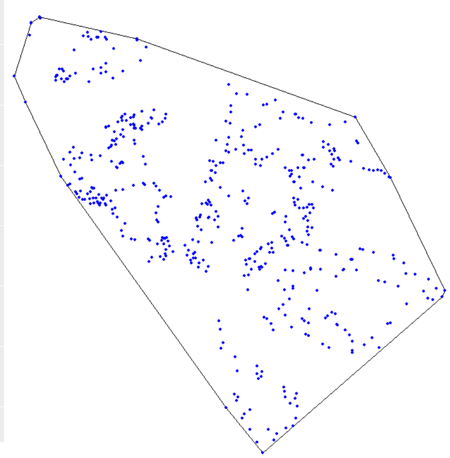
(a) Machochwe. Locations of households.



(b) Machochwe. Convex hull window.

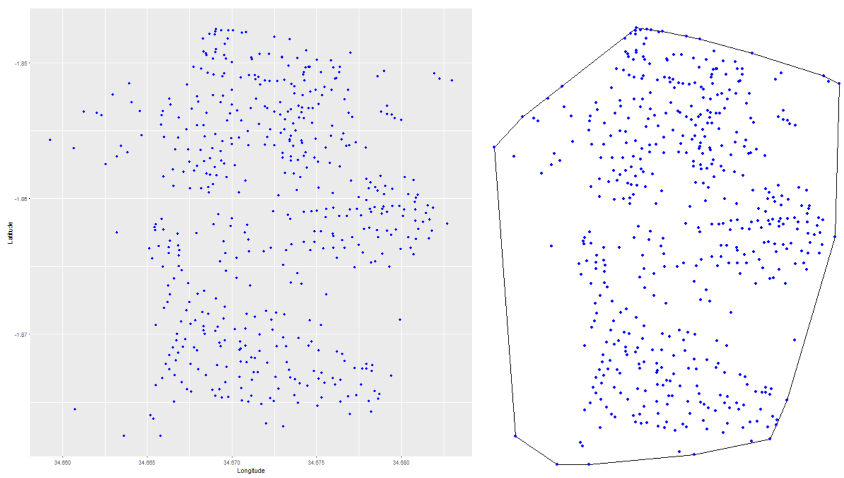


(c) Mesaga. Locations of households.

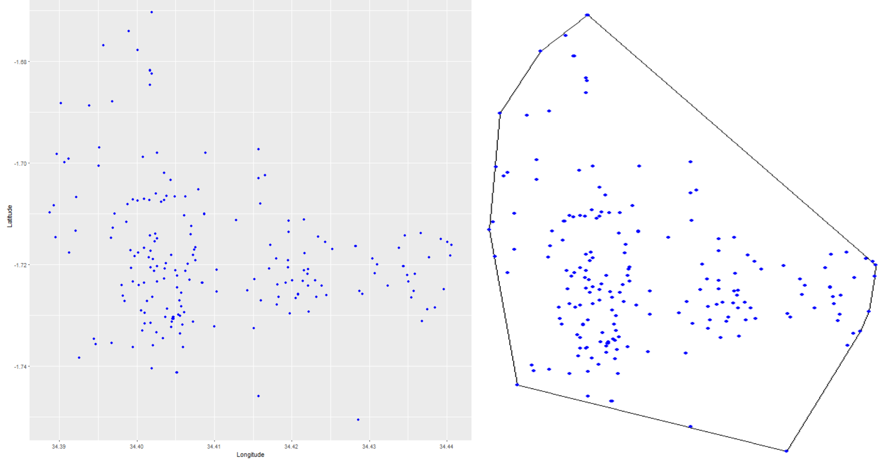


(d) Mesaga. Convex hull window.

Figure 7: Point patterns of households in Machochwe (top) and Mesaga (bottom), Mara province, Northern Tanzania.

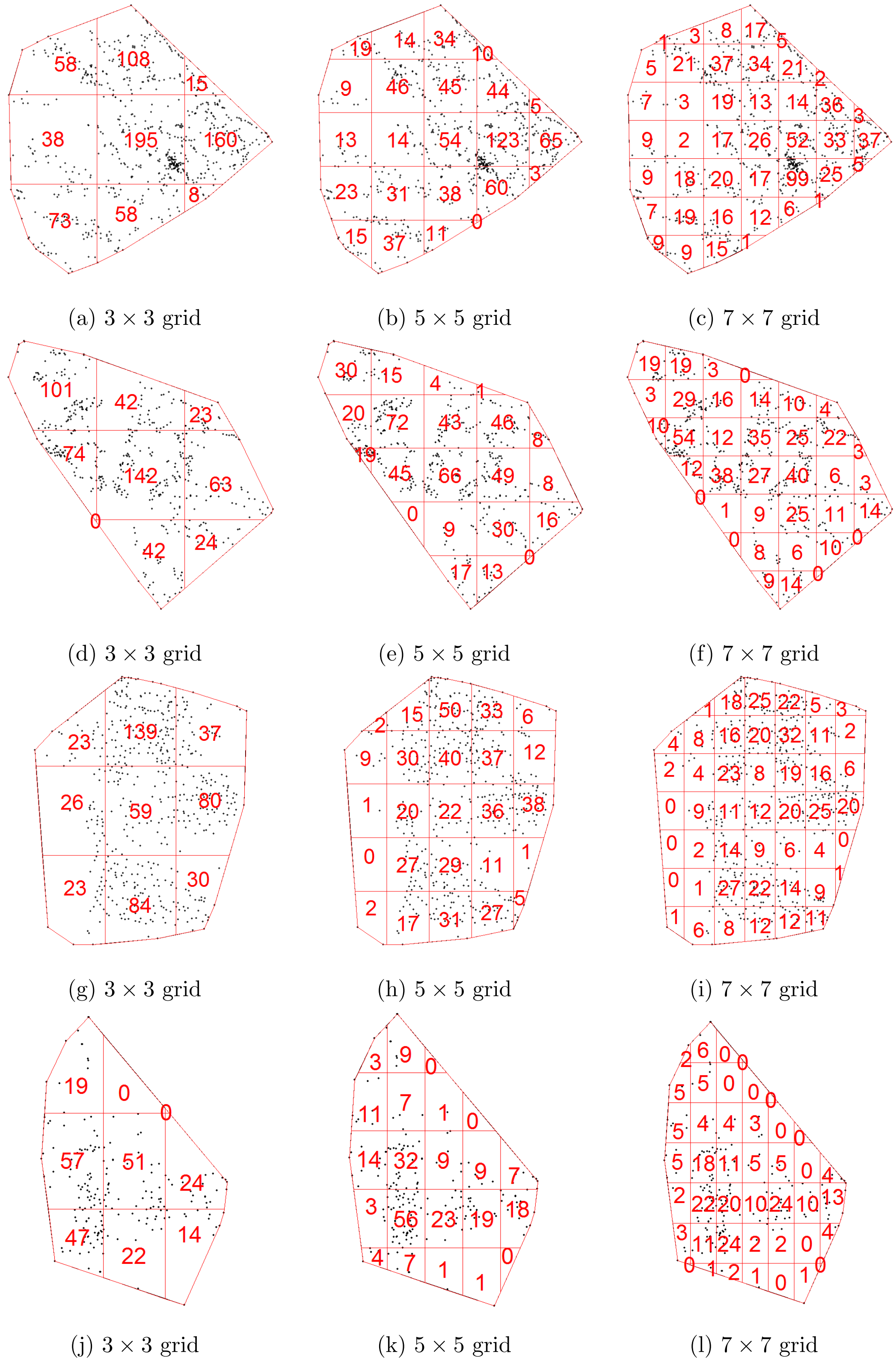


(a) Morotonga. Locations of households. (b) Morotonga. Convex hull window.



(c) Ring'wani. Locations of households. (d) Ring'wani. Convex hull window.

Figure 8: Point patterns of households in Morotonga (top) and Ring'wani (bottom), Mara province, Northern Tanzania.



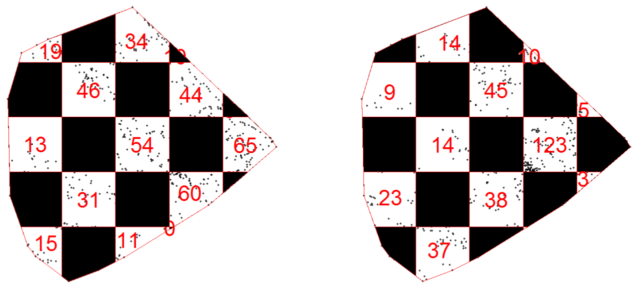
(a) 3×3 grid (b) 5×5 grid (c) 7×7 grid

(d) 3×3 grid (e) 5×5 grid (f) 7×7 grid

(g) 3×3 grid (h) 5×5 grid (i) 7×7 grid

(j) 3×3 grid (k) 5×5 grid (l) 7×7 grid

Figure 9: Point patterns divided into different grid sizes. Machochwe is shown in Figures (a)-(c), Mesaga is shown in Figures (d)-(f), Morotonga is shown in Figures (g)-(i) and Ring'wani is shown in Figures (j)-(l)



(a) Odd-numbered quadrats (b) Even-numbered quadrats

Figure 10: Point patterns divided into a 5×5 grid with either odd or even-numbered quadrats included in the calculation of the test statistic.

Village	Odd	Even
Machochwe	0.994	0.001
Mesaga	0.240	0.946
Morotonga	0.087	0.001
Ring'wani	0.001	0.364

Table 10: Parametric bootstrap p-values obtained for real-life point patterns using regular-spaced non-adjacent quadrats for the calculation of test statistics.

Applying bootstraping instead is more promising.

Thus we sampled 50% of the quadrats that are then used to calculate the proposed bootstrap test statistic, by sampling 50% of the quadrats 99 times. Only non-empty quadrats are used as the nature of this geograhical data of houses leads one to understand that empty regions may indicate areas not habitable e.g. mountains, ponds or rivers. The justification for sampling 99 times instead of 999 is again relevant here - sampling too many times could result in including more samples having dependencies. The p-values for each of the 99 test statistics are summarized in Tables 11 - 14 for the villages under consideration to show the minimum, maximum, average and median p-value. The standard deviation along with an empirical 95% confidence interval for the p-values are also shown.

4.1. Ring'wani - A small village

The village of Ring'wani is the smallest village considered, consisting of 243 households. When the pattern is divided into a 3×3 grid, we obtain 7 non-empty quadrats of which 4 quadrats were sampled at random. The five test statistics along with p-values were calculated for each of the samples. For 5×5 and 7×7 grids, 10 out of 19 and 16 out of 31 non-empty quadrats were sampled respectively.

Table 11 shows that for a 3×3 grid, the tests do not give consistent results with p-values ranging from 0.001 to 0.988. The empirical confidence intervals are wide and also do not lead to conclusive results regarding the stationarity of the pattern. The average p-value is insignificant for all tests as is the median p-value, except in the case of the UT test. When the grid size is increased to 5×5, the empirical confidence intervals become much narrower. All p-values are significant, indicating that the point pattern is non-stationary. When considering the average or median p-values, the same conclusion is reached. For a 7×7 grid, the p-values show slightly more variation between different samples. For the χ^2, likelihood ratio and UT tests, the empirical confidence intervals indicate only significant p-values. The empirical confidence intervals for the p-values of the score and VT tests include both significant and insignificant p-values, however the average and median p-values for these tests are shown to be significant.

4.2. Mesaga and Morotonga - Medium sized villages

The village of Mesaga consists of 512 households. When the pattern is divided into a 3×3 grid, we obtain 8 non-empty quadrats of which 4 quadrats were sampled at random. The five test statistics along with p-values were calculated for each of the samples. For 5×5 and 7×7 grids, 10 out of 19 and 16 out of 32 non-empty quadrats were sampled respectively. From Table 12 it is clear that for a 3×3 grid, again the tests do not give consistent results with p-values ranging from 0.008 to 0.989. The empirical confidence intervals are wide and also do not lead to conclusive results regarding the stationarity of the pattern. The average and median p-values are all insignificant. When the grid size is increased to 5×5 or 7×7, the empirical confidence intervals become much narrower. All p-values are significant, indicating that the point pattern is non-stationary. When considering the average or median p-values, the same conclusion is reached.

The village of Morotonga is similar in size to Mesaga, consisting of 501 households. However, due to the shape of the village, in other words the shape of the convex hull window, when the pattern is divided into a 3×3 grid, all quadrats are non-empty. Again, 4 of these non-empty quadrats were sampled at random and used for the calculation of the five test statistics and their associated p-values. For 5×5 and 7×7 grids, 12 out of 24 and 22 out of 43 non-empty quadrats were sampled respectively. From Table 13 it is clear that for a 3×3 grid, again the tests do not give consistent results with p-values ranging from 0.001 to 0.181. The empirical confidence intervals are wide and include both significant and insignificant p-values. However, the average and median p-values are all significant. When the grid size is increased to 5×5 or 7×7, the empirical confidence intervals become much narrower. All p-values are significant, indicating that the point pattern is non-stationary. When considering the average or median p-values, the same conclusion is reached.

4.3. Machochwe - the largest village

Machochwe is the largest village, consisting of 713 households. When the pattern is divided into a 3×3 grid, we obtain 9 non-empty quadrats of which 4 quadrats were sampled at random. For 5×5 and 7×7 grids, 11 out of 22 and 21 out of 42 non-empty quadrats were sampled respectively. The five test statistics along with p-values were calculated for each of the samples.

From Table 14 it is clear that for a 3×3 grid, the tests give inconsistent results for different samples of quadrats with p-values ranging from 0.001 to 0.831. The empirical confidence intervals for all tests except the UT test are wide and contain only insignificant p-values, leading to the pattern being incorrectly classified as stationary. The empirical confidence interval for the UT test includes both significant and insignificant p-values. For all tests, the average and median p-values are insignificant, again leading to the conclusion that the pattern is stationary. When the grid size is increased to 5×5 or 7×7, the empirical confidence intervals become narrow and include only significant p-values, now indicating that the point pattern is non-stationary. When considering the average or median p-values, the same conclusion is reached.

In this section the two alternative techniques for testing first-order stationarity were applied to rural village point patterns. Alternating quadrats are not consistent, whereas bootstrap sampling proves useful provided the grid size is appropriately chosen.

5. Discussion

In this paper we investigated five hypothesis tests for first-order stationarity in the context of spatial point patterns. The tests themselves are not new but their use on spatial point patterns has only had limited use. We investigated the power of these tests on rectangular windows as well as on irregular windows i.e. the convex hull type window. The powers for bootstrap sampling are good for a large enough grid, namely a large number of quadrats. When the number of quadrats is small none of the tests perform well. It is thus important to use a grid fine enough to cover the point pattern effectively in terms of the window as well as to pick up the nature of the point pattern. However, too fine a grid should be cautioned against.

It was shown that, when using regular-spaced non-adjacent quadrats, the choice of which quadrats to include can have a significant effect on the outcome of the tests with contradictory conclusions obtained for the two scenarios. This shows that the choice of which quadrats to include in the calculation of the test statistic can have a major impact on the outcomes of the tests. To use bootstrap random sampling, it is better to select a subset of quadrats to include in the calculation of the test statistics, after which p-values were obtained as before. Bootstraping was repeated 99 times and the resulting p-values were used to set up empirical 95% confidence intervals. For small grid sizes, these intervals tend to be wide, including both significant and insignificant p-values and no conclusion can be reached. When the grid size was increased, p-values stabilised and the empirical confidence intervals became more accurate. The only exception occurred for Ring'wani, a pattern that contains significantly fewer points than the other patterns. This improves testing first-order stationarity within a pattern and bypassing the violation of independence of quadrats.

Two of the hypothesis tests discussed here, namely Pearson's χ^2 and the likelihood ratio test, are already available for use in spatial statistical packages like `spatstat` in `R` [1, 43]. The χ^2 and likelihood ratio tests can be performed using the `quadrat.test` function available in the `spatstat` package in `R`. The first two are currently available in `spatstat` but the assumption of independence is not dealt with.

Baddeley [1] pointed out several limitations of the χ^2 test. Firstly, the alternative hypothesis that the point pattern has non-stationarity intensity merely negates the null hypothesis of a stationary intensity. There are two possible reasons why a point process may fail to be a stationary point process, namely that the process does in fact not have stationary intensity, or that it violates the property of independence between points. In addition, the built-in function should be used with caution. During the application phase of this article we discovered several irregularities in the test statistics obtained. Although the built-in functions correctly calculated the test statistic when using a rectangular window with no empty quadrats, discrepancies occurred if there were empty quadrats or when using a convex hull window. In the presence of empty quadrats, their areas are still included in the calculation of the statistic and no option is available to exclude these. We were unable to determine the exact source of the discrepancy in the test statistic calculations when using a convex hull window. In addition, small areas where the expected number of points is less than 5 are also not excluded from the calculation of the test statistic. Secondly, by default the function performs a two-sided hypothesis test. In this context we consider only one-sided tests, specifically upper-tailed tests, as a large absolute difference between observed and expected counts is an indication of a deviation from the null. Thirdly, when specifying in the function that a one-sided test must be performed, the p-value given in the output is for a lower-tailed test and needs to be adjusted as we are performing an upper-tailed test.

The choice of window is somewhat callenging. Even the convex hull may not capture the true situation, e.g. a hill in the village means that houses are unlikely to occur there and that the distance between houses on opposite sides of the hill creates artificial Euclidean distances. Possibly a type of network distance measure based on the communication paths between houses is more appropriate. Incorporating such covariate data to find non-convex type windows can prove useful here. The resulting window should then be used in testing for first-order stationarity and mean more precise conclusions.

In addition how to define a quadrat warrants further thought. Should they be regularly spaced and equal size? Could quadrats be chosen based on covariate information similar to the mechanism for window choice? It then follows that the definition of neighbour quadrats could also be extended, perhaps using a weight matrix or other methods.

6. Conclusion

This paper presents an extended methodology for first-order non-stationarity in spatial point patterns. Bootstrap sampling is able to deal with the assumption of independence in existing tests. The tests were applied to simulated as well as real-world point patterns. It showed good results in the case study. On regular and irregular windows, power tests were also performed. Tests that are based on using bootstrap random samples of quadrats, although more computationally intensive, provides us with an indication of how p-values differ when different quadrats are included in the calculation of the test statistics. When p-values are highly variable, these empirical confidence intervals also give us an indication of how appropriate the choice of grid size is. In our application, we showed how a small grid size (coarse grid) can negatively influence the results of the first-order stationarity tests and that larger grid sizes (finer grids) could potentially be more informative. The number of points within the pattern, however, should also be taken into account as a too fine a grid could also influence the tests negatively, as shown for the village Ring'wani. Future work could focus on the selection of an automatic optimal choice of quadrat size and determination of an irregular window using appropriate covariate information.

Acknowledgements

The authors acknowledge funding recieved from the South African National Research Foundation (NRF) under CSUR grant 90315, Center of Excellence in Mathematics and Statistical Science (COE-MaSS) statistics subtheme grant and the South African Statistical Association's NRF-SASA grant.

References

[1] A. Baddeley, E. Rubak, R. Turner, Spatial Point Patterns: Methodology and Applications with R, Chapman and Hall/CRC Press, London, 2015.

[2] A. C. Gatrell, T. C. Bailey, P. J. Diggle, B. S. Rowlingson, Spatial point pattern analysis and its application in geographical epidemiology, Transactions of the Institute of British Geographers 21 (1) (1996) 256–274.

[3] A. Baddeley, Analysing spatial point patterns in R, Tech. Rep. Version 4.1, CSIRO and University of Western Australia (December 2010).

[4] R. Haining, Spatial Data Analysis: Theory and Practice, Cambridge University Press, 2003.

[5] B. Matérn, Spatial variation, Lecture Notes in Statistics, Springer-Verlag, 1986.

[6] J. Wang, R. Haining, Z. Cao, Sample surveying to estimate the mean of a heterogeneous surface: reducing the error variance through zoning, International Journal of Geographical Information Science 24 (4) (2010) 523–543.

[7] P. J. Diggle, Statistical Analysis of Spatial and Spatio-Temporal Point Patterns, 3rd Edition, Vol. 128 of Monographs on Statistics and Applied Probability, CRC Press, Taylor & Francis Group, 2014.

[8] Y. Guan, A KPSS test for stationarity for spatial point processes, Biometrics 64 (2008) 800–806.

[9] D. Kwiatkowski, P. Phillips, P. Schmidt, Y. Shin, Testing the null hypotheisis of stationarity against the alternative of a unit root, Journal of Econometrics 54 (1992) 159 – 178.

[10] T. Zhang, B. Zhou, Test for the first-order stationarity for spatial point processes in arbitray regions, Journal of Agricultural, Biological and Environmental Statistics 19 (4) (2014) 387–404.

[11] M. Berman, P. Diggle, Estimating weighted integrals of the second-order intensity of a spatial point process, Journal of the Royal Statistical Society. Series B (Methodological) (1989) 81–92.

[12] S. N. Chiu, L. Wang, Homogeneity tests for several Poisson populations, Computational Statistics & Data Analysis 53 (12) (2009) 4266 – 4278.

[13] W. R. Tobler, A computer movie simulating urban growth in the Detroit region, Economic Geography 46 (1970) 234–240.
URL `http://0-www.jstor.org.innopac.up.ac.za/stable/143141`

[14] P. A. P. Moran, Notes on continuous stochastic phenomena, Biometrika 37 (1/2) (1950) 17–23.
URL `http://www.jstor.org/stable/2332142`

[15] R. C. Geary, The contiguity ratio and statistical mapping, The Incorporated Statistician 5 (3) (1954) 115–146.
URL `http://www.jstor.org/stable/2986645`

[16] D. J. Best, The difference between two Poisson expectations, Australian Journal of Statistics 17 (1) (1975) 29–33.

[17] K. Detre, C. White, The comparison of two Poisson-distributed observations, Biometrics 26 (4) (1970) 851–854.

[18] S. N. Chiu, Parametric bootstrap and approximate tests for two Poisson variates, Journal of Statistical Computation and Simulation 80 (3) (2010) 263–271.

[19] H. Ng, K. Gu, M. Tang, A comparative study of tests for the difference of two Poisson means, Computational Statistics & Data Analysis 51 (6) (2007) 3085 – 3099.

[20] H. K. T. Ng, M.-L. Tang, Testing the equality of two Poisson means using the rate ratio, Statistics in Medicine 24 (6) (2005) 955–965.

[21] R. F. Potthoff, M. Whittinghill, Testing for homogeneity: II. the Poisson distribution, Biometrika 53 (1/2) (1966) 183–190.

[22] G. E. Blackman, A study by statistical methods of the distribution of species in grassland associations, Annals of Botany 49 (196) (1935) 749–777.

[23] P. Greig-Smith, The use of random and contiguous quadrats in the study of the structure of plant communities, Annals of Botany 16 (2) (1952) 293–316.

[24] K. J. Koehler, K. Larntz, An empirical investigation of goodness-of-fit statistics for sparse multinomials, Journal of the American Statistical Association 75 (370) (1980) 336–344.

[25] T. R. C. Read, Small-sample comparisons for the power divergence goodness-of-fit statistics, Journal of the American Statistical Association 79 (388) (1984) 929–935.

[26] J. Neyman, E. S. Pearson, On the use and interpretation of certain test criteria for purposes of statistical inference: Part I, Biometrika 20A (1/2) (1928) 175–240.

[27] S. S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, The Annals of Mathematical Statistics 9 (1) (1938) 60–62.

[28] C. R. Rao, I. M. Chakravarti, Some small sample tests of significance for a Poisson distribution, Biometrics 12 (3) (1956) 264–282.

[29] C. R. Rao, Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation, Mathematical Proceedings of the Cambridge Philosophical Society 44 (1) (1948) 50–57.

[30] A. K. Bera, Y. Bilias, Rao's score, Neyman's C(α) and Silvey's LM tests: an essay on historical developments and some new results, Journal of Statistical Planning and Inference 97 (1) (2001) 9–44.

[31] E. T. M. Ng, R. J. Cook, Adjusted score tests of homogeneity for Poisson processes, Journal of the American Statistical Association 94 (445) (1999) 308–319.

[32] A. Davison, D. Hinkley, Bootstrap Methods and Their Application, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1997.

[33] R. Davidson, J. G. MacKinnon, Graphical methods for investigating the size and power of hypothesis tests, The Manchester School 66 (1) (1998) 1–26.

[34] R. Davidson, J. G. MacKinnon, Improving the reliability of bootstrap tests with the fast double bootstrap, Computational Statistics & Data Analysis 51 (7) (2007) 3259–3281.

[35] A. C. A. Hope, A simplified Monte Carlo significance test procedure, Journal of the Royal Statistical Society. Series B (Methodological) 30 (3) (1968) 582–598.

[36] J. Besag, P. J. Diggle, Simple Monte Carlo tests for spatial pattern, Journal of the Royal Statistical Society. Series C (Applied Statistics) 26 (3) (1977) 327–333.

[37] G. Oyeyemi, A. Adewara, F. Adebola, S. Salau, On the estimation of power and sample size in test of independence, Asian Journal of Mathematics & Statistics 3 (2010) 139–146.

[38] Z. Zhang, K.-H. Yuan, Practical Atatistical Power Analysis Using WebPower and R, ISDSA Press, India, 2017.

[39] C. Kraamwinkel, I. Fabris-Rotelli, G. Fosgate, D. L. Knobel, K. Hampson, A study on the apparent randomness of an animal sample, in: South African Statistical Association 2014 Conference Proceedings, November 2014.

[40] Y. Yang, B. Jia, A comparative study on the spatial patterns of Chinese row houses in colonial cities in the early 20th century, Open and Sustainable Building .

[41] B. D. Ripley, Modelling spatial patterns, Journal of the Royal Statistical Society. Series B (Methodological) 39 (2) (1977) 172–212.

[42] J. Rasson, M. Rémon, F. Henry, Finding the edge of a poisson forest with inside and outside observations: the discriminant analysis point of view, From Data to Knowledge: Theoretical and Practical Aspects of Classification, Data Analysis, and Knowledge Organization (2013) 94.

[43] A. Baddeley, R. Turner, Spatstat: An R package for analyzing spatial point patterns, Journal of Statistical Software 12 (6) (2005) 1–42.

		χ^2	LR	SC	VT	UT
3×3	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.988	0.988	0.988	0.988	0.988
	Average	0.357081	0.356879	0.247838	0.356606	0.152758
	Median	0.554	0.554	0.166	0.554	0.005
	Standard deviation	0.348608	0.348588	0.301118	0.348074	0.304328
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.8902	0.891	0.892	0.887	0.856
5×5	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.035	0.02	0.004	0.03	0.032
	Average	0.001657	0.001313	0.001101	0.001727	0.001646
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.003634	0.002013	0.000463	0.003434	0.003382
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.0072	0.003	0.003	0.0102	0.007
7×7	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.047	0.068	0.128	0.116	0.054
	Average	0.002636	0.002909	0.007556	0.004566	0.002727
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.007577	0.008986	0.01823	0.017567	0.008405
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.0317	0.0267	0.0614	0.06025	0.03075

Table 11: Summary of p-values for hypothesis tests applied to household location point pattern for Ring'wani. For a 3×3 grid, test statistics were calculated using random samples of 4 non-empty quadrats. For 5×5 and 7×7 grids, random samples of respectively 10 and 16 non-empty quadrats were used for the calculations.

		χ^2	LR	SC	VT	UT
3×3	Min	0.016	0.042	0.008	0.066	0.02
	Max	0.989	0.989	0.988	0.988	0.986
	Average	0.433	0.43	0.431	0.434	0.509
	Median	0.166	0.161	0.407	0.119	0.683
	Standard deviationv	0.369686	0.35847	0.322995	0.364816	0.353046
	$P_{2.5}$	0.023	0.046	0.012	0.071	0.024
	$P_{97.5}$	0.9698	0.971	0.970	0.967	0.940
5×5	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.003	0.003	0.01	0.003	0.004
	Average	0.001061	0.001051	0.001253	0.001061	0.00104
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.000345	0.000299	0.001181	0.000345	0.000317
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.0021	0.00155	0.0041	0.0021	0.001
7×7	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.002	0.001	0.01	0.002	0.001
	Average	0.00101	0.001	0.001182	0.00101	0.001
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.000101	0	0.000973	0.000101	0
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.001	0.001	0.002	0.001	0.001

Table 12: Summary of p-values for hypothesis tests applied to household location point pattern for Mesaga. For a 3×3 grid, test statistics were calculated using random samples of 4 non-empty quadrats. For 5×5 and 7×7 grids, random samples of respectively 10 and 16 non-empty quadrats were used for the calculations.

		χ^2	LR	SC	VT	UT
3×3	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.175	0.173	0.176	0.181	0.152
	Average	0.007879	0.008071	0.008071	0.008121	0.008576
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.026754	0.027258	0.026745	0.027749	0.027657
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.1088	0.111	0.0981	0.112	0.113
5×5	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.01	0.009	0.027	0.011	0.006
	Average	0.001343	0.001121	0.001556	0.001323	0.001101
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.001334	0.000836	0.002775	0.001384	0.000562
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.0046	0.002	0.005	0.00455	0.002
7×7	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.002	0.002	0.004	0.002	0.001
	Average	0.00101	0.00101	0.001091	0.00101	0.001
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0.000101	0.000101	0.00038	0.000101	0
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.001	0.001	0.002	0.001	0.001

Table 13: Summary of p-values for hypothesis tests applied to household location point pattern for Morotonga. For a 3×3 grid, test statistics were calculated using random samples of 4 non-empty quadrats. For 5×5 and 7×7 grids, random samples of respectively 12 and 22 non-empty quadrats were used for the calculations.

		χ^2	LR	SC	VT	UT
3×3	Min	0.091	0.038	0.057	0.094	0.001
	Max	0.822	0.814	0.804	0.831	0.818
	Average	0.446202	0.417909	0.431667	0.446273	0.288838
	Median	0.441	0.392	0.435	0.441	0.209
	Standard deviation	0.158406	0.200491	0.169582	0.158038	0.238708
	$P_{2.5}$	0.1218	0.1118	0.0888	0.12465	0.001
	$P_{97.5}$	0.7263	0.724	0.724	0.726	0.726
5×5	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.001	0.001	0.003	0.001	0.016
	Average	0.001	0.001	0.00103	0.001	0.001152
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0	0	0.000224	0	0.001508
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.001	0.001	0.001	0.001	0.001
7×7	Min	0.001	0.001	0.001	0.001	0.001
	Max	0.001	0.001	0.003	0.001	0.001
	Average	0.001	0.001	0.001061	0.001	0.001
	Median	0.001	0.001	0.001	0.001	0.001
	Standard deviation	0	0	0.000314	0	0
	$P_{2.5}$	0.001	0.001	0.001	0.001	0.001
	$P_{97.5}$	0.001	0.001	0.002	0.001	0.001

Table 14: Summary of p-values for hypothesis tests applied to household location point pattern for Machochwe. For a 3×3 grid, test statistics were calculated using random samples of 4 non-empty quadrats. For 5×5 and 7×7 grids, random samples of respectively 11 and 21 non-empty quadrats were used for the calculations.