A KPSS Test for Stationarity for Spatial Point Processes

Yongtao Guan
Division of Biostatistics, Yale University,
New Haven, CT 06520-8034
email: yongtao.guan@yale.edu

August 7, 2007

SUMMARY. We propose a formal method to test stationarity for spatial point processes. The proposed test statistic is based on the integrated squared deviations of observed counts of events from their means estimated under stationarity. We show that the resulting test statistic converges in distribution to a functional of a two-dimensional Brownian motion. To conduct the test, we compare the calculated statistic with the upper tail critical values of this functional. Our method requires only a weak dependence condition on the process but does not assume any parametric model for it. As a result, it can be applied to a wide class of spatial point process models. We study the efficacy of the test through both simulations and applications to two real data examples that were previously suspected to be nonstationary based on graphical evidences. Our test formally confirmed the suspected nonstationarity for both data.

KEY WORDS: Spatial Point Process, KPSS Test for Stationarity
1. Introduction

Spatial point pattern data consist of locations of events that are often of interest in biological and ecological studies. Examples include locations of the nuclei of cells in a tissue (Diggle 2003), of trees in a tropical forest (Waagepetersen 2007), and of wildfires in a region (Schoengberg 2004). Such data are commonly viewed as a realization from a stochastic process called the spatial point process, which is often assumed to be stationary in practice. Under the assumption of stationarity, the main interests of the analysis are to test complete spatial randomness (CSR; i.e. whether the process is Poisson) and to model the interaction between events (e.g. clustered or inhibitive interactions) if CSR is rejected (e.g. Diggle 2003). When stationarity is not true, however, the analysis often becomes more complicated. In particular, the form of nonstationarity may have to be modeled properly first before any similar analysis used in the stationary case can be conducted.

Although a useful assumption, stationarity may be questionable in real applications. Many authors have in fact expressed concerns over the validity of stationarity for their data. For example, for the longleaf pine data to be detailed in Section 5, Cressie (1993, p.600) reported that there was “a clear trend of increasing intensity from the eastern to the western half of the study region”. For the Ambrosia Dumosa data also to be detailed in Section 5, Perry et al. (2002) observed “two or three bands of relatively low or zero density” that appeared to “run roughly north-west to south-east across the area”. Thus both data examples appeared to be nonstationary rather than stationary. We note that the form of nonstationarity being suspected in these two examples (and that being considered in the article subsequently)
in fact refers to the inhomogeneity of the intensity of the process, which is only a special but nevertheless important case of the more general nonstationarity. With a slight abuse of notation, we will use nonstationarity for inhomogeneity.

A major research interest in recent spatial point process literature is to model nonstationary spatial point patterns (e.g. Møller and Waagepetersen 2007). However, before doing so, a fundamental question yet to be answered is whether it is really necessary to treat a process as being nonstationary. As explained in the first paragraphs of this section, different modeling strategies need to be taken for stationary and nonstationary processes. To assess stationarity, graphical methods are often used in practice. For example, Cressie (1993) used a contour plot of nonparametric kernel estimates of the intensity for the longleaf pine data, whereas Perry et al. (2002) used a scatter plot of the observed plant locations for the Ambrosia Dumosa data. Often graphical diagnostics are difficult to assess and may be open to interpretation. This is especially the case for spatial point processes since the eye can be “a poor judge” (Diggle 2003, p.91). Therefore, it would be of interest to develop a formal method to test stationarity.

We note that testing stationarity has been studied extensively in the time series analysis literature (e.g., Kwiatkowski et al. 1992; Hobijn et al. 2004; Giraitis et al. 2006). However, little has been done for spatial point processes. In this paper, we develop a spatial point process version of the popular KPSS test developed by Kwiatkowski et al. (1992) in time series analysis. Our test statistic is constructed in terms of the integrated squared deviations of the observed counts in some sets from the (estimated) means
under stationarity. A nice property of the proposed method is that it requires only a weak dependence assumption on the process but is not restricted to any specific class of models. This is desirable considering the fact that spatial point processes often possess rather complicated dependence structures (e.g. clustered and inhibitive structures). To our best knowledge, this is the first test of its kind for spatial point processes.

The remainder of the article is organized as follows. In Section 2, we give necessary background on the KPSS test in time series. In Section 3, we introduce the proposed approach and investigate its asymptotic properties. We apply the test to some simulated data in Section 4 and the two motivating examples in Section 5. Theoretical justifications of the test are given in the web appendices.

2. KPSS Test for Stationarity for Time Series

Let \( Y_i, i = 1, \cdots, n \), be the observed time series for which we wish to test stationarity, and \( \bar{Y}_n \) be the sample mean of \( Y_i \). Consider the following partial sum process of derivations from the (estimated) mean of the process:

\[
S_n(u) = \sum_{i=1}^{u} (Y_i - \bar{Y}_n).
\]  

Under the null hypothesis of stationarity, \( E[S_n(u)] = 0 \) for all \( u = 1, \cdots, n \). If the process is nonstationary, however, we may expect \( E[S_n(u)] \) to be different from zero for some \( u \). In terms of the partial sum process \( S_n(u) \), Kwiatkowski et al. (1992) proposed the following KPSS test statistic for stationarity:

\[
T_n = \frac{1}{n^2 S_n^2} \sum_{u=1}^{n} [S_n(u)]^2,
\]  

4
where
\[ s_n^2 = m_n^{-1} \sum_{i,j=1}^{m_n} \tilde{\gamma}_{i-j}, \]
(3)

\[ \tilde{\gamma}_j = (n-j)^{-1} \sum_{i=1}^{n-j} (Y_i - \bar{Y}_n)(Y_{i+j} - \bar{Y}_n) \text{ for } 0 \leq j < n, \]
and \( m_n \) is a bandwidth satisfying \( m_n \to \infty \) and \( m_n/n \to 0 \). Here \( s_n^2 \) is a nonparametric estimator for the limit of \( n\text{Var}(\bar{Y}_n) \), provided that it exists. Under some suitable conditions, Kwiatkowski et al. (1992) showed that \( T_n \) converged in distribution to \( \int_0^1 V(t)^2 \, dt \), where \( V(t) \) is a standard Brownian bridge. The test can then be conducted by comparing \( T_n \) with the upper tail critical values of \( \int_0^1 V(t)^2 \, dt \), which can be approximated by simulating the Brownian bridge \( V(t) \). Specifically, if \( T_n \) is larger than the desired upper tail critical value, then we should reject stationarity.

3. KPSS Test for Stationarity for Spatial Point Processes

3.1 The proposed method

Let \( N \) be a spatial point process that is observed over \([0, n_1] \times [0, n_2]\), where \( n_i = c_i n \), \( c_i > 0 \). If the point process is stationary, then the intensity of the process is equal to a constant, say \( \lambda \). Let \( N(u_1, u_2) \) denote the number of events of \( N \) in the set \([0, u_1] \times [0, u_2]\), where \( 0 < u_i \leq n_i \). Note that under stationarity, \( E[N(u_1, u_2)] = u_1 u_2 \lambda \) for all \( 0 < u_i \leq n_i \). Thus an unbiased estimator for \( \lambda \) is given by \( \hat{\lambda}_n = N(n_1, n_2)/(n_1 n_2) \). For any fixed \( u_1 \) and \( u_2 \), we define the spatial point process version of \( S_n \) in (1) as:

\[ S_n(u_1, u_2) = N(u_1, u_2) - u_1 u_2 \hat{\lambda}_n. \]
(4)

Note that \( S_n(u_1, u_2) \) essentially compares the actually observed number of events in the set \([0, u_1] \times [0, u_2]\) with the expected number of events estimated
under stationarity in the same set. If the point process is indeed stationary, then clearly $E[S_n(u_1, u_2)] = 0$. Otherwise we expect $E[S_n(u_1, u_2)]$ to be different from zero for some $0 < u_i \leq n_i$. In light of this observation, it’s natural to consider a test statistic taking the following general form:

$$
\tilde{T}_n = \frac{1}{n_1^n n_2^n} \int_0^{n_1} \int_0^{n_2} [S_n(u_1, u_2)]^2 du_1 du_2. 
$$

(5)

In the above, $\sigma^2$ is a standardizing term used to ensure that the statistic $\tilde{T}_n$ for any (weakly dependent) stationary process, either clustered or inhibitive or both, converges to a common limiting distribution. Specifically, let

$$
\sigma^2 = \lambda^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [g(l_1, l_2) - 1] dl_1 dl_2 + \lambda, 
$$

(6)

where $g(l_1, l_2)$ is the pair correlation function (PCF; Stoyan and Stoyan 1994) at lag $(l_1, l_2)$. Here the PCF is a function that describes the second-order structure of the point process. If two point events separated by a lag $(l_1, l_2)$ are completely independent, then $g(l_1, l_2) = 1$. We show in the Web Appendix B that under some suitable conditions requiring the dependence of the process is weak enough,

$$
\tilde{T}_n \xrightarrow{d} \int_0^1 \int_0^1 [V(t_1, t_2)]^2 dt_1 dt_2,
$$

(7)

where $V(t_1, t_2) = W(t_1, t_2) - t_1 t_2 W(1, 1)$ and $W(t_1, t_2)$ is a two-dimensional Brownian motion, i.e. a process that has continuous sample paths and independent increments in disjoint sets, and satisfies $E[W(t_1, t_2)] = 0$ and $\text{VAR}[W(t_1, t_2)] = t_1 t_2$. To conduct the test, we can then compute $\tilde{T}_n$ defined in (5) directly, if $\sigma^2$ is known, and compare it with the upper tail critical values obtained from its limiting distribution given in (7). Note that the desired critical values can be easily obtained by first simulating $V(t_1, t_2)$ and
then evaluating the integral in (7). Specifically, the critical values at the 1%, 5% and 10% levels are approximately equal to 0.5013, 0.3244 and 0.2542, respectively.

A complication in practice is that $\sigma^2$ used to define $\tilde{T}_n$ is typically unknown and thus has to be estimated. From its expression in (6), we can fit a parametric PCF model to the data and then plug the estimated PCF into (6) in order to estimate $\sigma^2$. The accuracy of the estimator, however, will inevitably depend on how close the used parametric PCF model is to the true PCF. To avoid the need to specify a parametric PCF model, we will instead use the following nonparametric estimator for $\sigma^2$:

$$\hat{\sigma}^2_n = \frac{1}{(n_1 - |x - x'|)(n_2 - |y - y'|)} \sum_{(x-x')^2+(y-y')^2 \leq m_n^2} 1 - \hat{\lambda}^2_n \pi m_n^2 + \hat{\lambda}_n, \quad (8)$$

where $\hat{\lambda}^2_n$ is an estimator for $\lambda^2$, e.g. $\hat{\lambda}^2_n$, $m_n \to \infty$, $m_n/n \to 0$ and the $\neq$ sign over the double summation signifies summation over all distinct pairs of events $(x, y)$ and $(x', y')$. We show in the Web Appendix C that $\hat{\sigma}^2_n$ is a consistent estimator for $\sigma^2$ under some weak conditions. This in turn leads to the following test statistic:

$$T_n = \frac{1}{n_1^2 n_2^2 \hat{\sigma}^2_n} \int_0^{n_1} \int_0^{n_2} [S_n(u_1, u_2)]^2 du_1 du_2, \quad (9)$$

Clearly $T_n$ here has the same limiting distribution as being given in (7) due to the consistency of $\hat{\sigma}^2_n$ for $\sigma^2$. Furthermore, note that $T_n$ is essentially the spatial point process version of the KPSS test statistic in (2) in the time series case. We thus call the proposed test also as a KPSS test. In the next section, we will study the statistical properties of $T_n$. 
3.2 Consistency of the test

A question of interest regarding the proposed test is whether it is consistent. We say a test is consistent if the probability for it to conclude a given correct alternative hypothesis converges to one as \( n \) increases to infinity. In our setting, this is equivalent to that the test statistic \( T_n \), where \( T_n \) is defined in (9), increases to infinity with \( n \).

To establish the consistency of the test, let \( C \) denote a constant that may take different values, and let \( \lambda(x, y) \) and \( \lambda_2[(x, y), (x', y')] \) denote the first- and second-order intensity functions of the process (see the Web Appendix A for their definitions), respectively. For \( \lambda(x, y) \), we assume

\[
n^{-\gamma} \lambda(n_1 t_1, n_2 t_2) \to \lambda_0(t_1, t_2) \text{ for some } \gamma > -2,\tag{10}
\]

where \( \lambda_0(t_1, t_2) \) is a bounded function that takes nonconstant values on a set with positive measure in \([0,1] \times [0,1]\). For \( \lambda_2[(x, y), (x', y')] \), we assume

\[
\lambda_2[(x, y), (x', y')] \leq C \lambda(x, y) \lambda(x', y'),\tag{11}
\]

\[
\int_0^\infty \int_0^\infty |\lambda_2[(x, y), (x', y')] - \lambda(x, y)\lambda(x', y')| dx' dy' \leq C n^\gamma \lambda(x, y).\tag{12}
\]

Under conditions (10), (11) and (12), we show in the Web Appendix D that \( T_n \to \infty \) in probability as \( n \to \infty \), i.e. the proposed test is consistent with detecting the alternative hypothesis that satisfies (10). Condition (10) requires that the intensity function \( \lambda(\cdot) \) has a discernible difference at a rate \( n^\gamma \). This is a rather weak condition given that \( \gamma \) only needs to be larger than \(-2\). Conditions (11) and (12) are also fairly weak conditions. In the special case when \( N \) is second-order reweighted stationary (Baddeley et al. 2000), i.e. \( \lambda[(x, y), (x', y')] = \lambda(x, y)\lambda(x', y')g(x - x', y - y') \), provided the PCF
$g(\cdot)$ exists, then (11) and (12) hold if $g(\cdot)$ is bounded and \( \int_0^\infty \int_0^\infty |g(l_1,l_2) - 1| dl_1 dl_2 < C \). Many commonly used point process models such as the Poisson cluster process and the simple inhibition process models considered in the simulation study satisfy these conditions.

3.3 Practical guidelines

To calculate the test statistic $T_n$, it is important to decide the bandwidth parameter $m_n$ in (8). Generally a small $m_n$ leads to an estimator with a large bias, whereas a large $m_n$ leads to one with a large variance. A balance has to be taken between controlling the bias and variance. To select $m_n$, we adopt an ad hoc procedure by examining the empirical PCF plot. Specifically, we define $m_n$ as the smallest lag distance beyond which the empirical PCF plot becomes flat and close to one, where the level of flatness and closeness to one has to be judged by eye. Our main goal here is to control the bias. By using the “smallest” possible $m_n$, however, we also control the variance in an indirect way. It would be desirable if $m_n$ can be selected by some automatic data-driven method. This should be considered in future research.

Additionally, we can fit a parametric PCF model to the data and calculate the fitted PCF at lag $m_n$, where $m_n$ is the suggested bandwidth by the empirical PCF plot. If the fitted PCF is indeed very close to one, then this provides further evidence in support of the choice of $m_n$. Otherwise a larger value of $m_n$ may be needed. The parametric approach can help reduce the arbitrariness involved in using the graphical approach alone.

Another practical issue with applying the proposed test is which location of a study region should be defined as the origin, i.e. $(0,0)$. For a rectangular region as being considered in this article, it appears to be equally reasonable
to define any of its four corners as the origin. A sensible choice is to first calculate the test statistic $T_n$ by using each of the four corners as the origin and then combine these statistics in some way. To do so, let $S_i^n(u_1, u_2), i = 1, \cdots , 4$, be the random processes defined in (4), where the four corners of the rectangular region, starting from the lower left and going clockwise toward the bottom right, are in turn defined as the four origins. We then define the following test statistic:

$$T^*_n = \frac{1}{n_1n_2\hat{\sigma}^2_n} \sum_{i=1}^{4} \int_0^{n_1} \int_0^{n_2} [S_i^n(u_1, u_2)]^2 du_1 du_2.$$ (13)

We show in the Web Appendix E that the test statistic $T^*_n$ converges in distribution to

$$\int_0^1 \int_0^1 [W(t_1, t_2) - t_1t_2W(1, 1)]^2 dt_1 dt_2$$

$$+ \int_0^1 \int_0^1 [W(t_1, 1) - W(t_1, 1 - t_2) - t_1t_2W(1, 1)]^2 dt_1 dt_2$$

$$+ \int_0^1 \int_0^1 [W(1, t_2) - W(1 - t_1, t_2) - t_1t_2W(1, 1)]^2 dt_1 dt_2$$

$$+ \int_0^1 \int_0^1 [W(1, 1) - W(1 - t_1, 1) - W(1, 1 - t_2) - W(1 - t_1, 1 - t_2) - t_1t_2W(1, 1)]^2 dt_1 dt_2, $$ (14)

where $W$ is a two-dimensional Brownian motion as before. In view of this result, a test for stationarity is to reject the null hypothesis if $T^*_n$ is larger than its corresponding upper tail critical value from (14). Note that the desired critical values can be easily obtained by first simulating the two-dimensional Brownian motion $W(t_1, t_2)$ and then evaluating the integrals in (14). Specifically, the critical values at the 1%, 5% and 10% levels are approximately equal to 1.5618, 1.1196 and 0.9322, respectively.
4. A Simulation Study

We simulated realizations from both a Poisson cluster process and a simple inhibition process of Matérn’s first type (e.g. Diggle 2003) on a $n \times n$ square region, where $n = 10, 20$. For both processes, the first-order intensity function was defined as $\lambda(x, y) = \alpha \exp(\beta x/n)$ for $\beta = 0, 1, 2$. Note that $\beta = 0$ led to a stationary model whereas $\beta = 1, 2$ both led to nonstationary models. Furthermore, the model for $\beta = 2$ is more nonstationary than $\beta = 1$. For each $\beta$, we selected $\alpha$ such that the expected number of events per realization was roughly $n^2$.

For the Poisson cluster process, we used a homogeneous Poisson process for the parent process. The intensity of the process was equal to 0.25, which led to 25 and 100 expected number of parents for $n = 10, 20$, respectively. Each parent generated a Poisson number of offspring. The position of each offspring relative to its parent was defined as a radially symmetric Gaussian random variable (e.g. Diggle 2003) with a standard deviation, $\sigma = 0.2, 0.4$. Each offspring was then thinned as in Waagepetersen (2007) where the probability to retain an offspring was equal to the intensity at the offspring location divided by the maximum intensity in the region, i.e. $\alpha \exp(\beta)$ in our case.

For the simple inhibition process, we first simulated a stationary simple inhibition process of Matérn’s first type. The inhibition parameter was set to be 0.2, i.e. no pair could have an inter-pair distance less than 0.2. Each process was then thinned as in the Poisson cluster process case.

We simulated one thousand realizations using each set of parameters for the two types of processes. For each realization, we applied the proposed method to test stationarity. Specifically, we calculated the test statistic de-
fined in (9) by setting each of the four corners of the square as the origin as well as the test statistic defined in (13). The nonparametric variance estimator given in (10) was used to calculate \( \hat{\sigma}^2_n \). For the Poisson cluster processes, note that \( 4\sigma \) can be roughly treated as the dependence range, as suggested by Diggle (2003). We thus set the bandwidth parameter \( m_n = 3\sigma, 4\sigma, 5\sigma \).

For the simple inhibition processes, note that the dependence completely dies out if the lag distance is larger than 0.4. We thus set the bandwidth parameter \( m_n = 0.2, 0.4, 0.6 \). The different \( m_n \) values for both types of processes allowed us to evaluate the sensitiveness of the proposed test to the choice of \( m_n \) around the dependence range. In practice, the dependence range can be roughly estimated by examining an empirical PCF plot. We rejected the null hypothesis if the calculated test statistic was larger than its corresponding critical value at the 5% level, where the critical values were approximated by using (8) and (15) for the test statistics defined in (9) and (13), respectively.

Table 1 lists the empirical sizes and powers in the Poisson cluster process case. For sizes, we see that all empirical sizes were reasonably close to the nominal size. In general, the sizes were closer to the nominal size for (9) than (13). This was expected since (13) took a more complicated form than (9) and thus might converge more slowly to its limiting distribution. For powers, we see that the empirical powers increased when \( \beta \) increased and when \( n \) increased. This provided evidence for our theoretical results on the consistency of the test. The powers typically decreased when \( m_n \) increased. Thus a larger \( m_n \) tended to yield a more conservative test. However, the changes were rather small in general. In all cases, (13) was more powerful than (9). The difference was often quite substantial, especially in the small
sample size case.

Table 2 lists the empirical sizes and powers in the simple inhibition process case. Compared to the Poisson cluster case, we see that the empirical sizes were even closer to the nominal size and the powers were also much higher. In other words, the test became better both in terms of its ability to achieve the nominal size and to obtain a high power. This improvement was likely due to the fact that the dependence was much weaker in this case, which in turn led to much improved variance estimates $\hat{\sigma}^2_n$ in the null case and more evident signals for nonstationarity in the alternative case. The former led to a better size whereas the latter increased the power. As in the Poisson cluster process case, (13) was more powerful than (9) in general. The choice of $m_n$ did not affect the results significantly.

Table 2 about here.

5. Applications

5.1 Longleaf pine data

The longleaf pine data consist of locations of 584 longleaf pine trees in a $200m \times 200m$ square region (see Figure 1) in an old-growth forest in Thomas County, Georgia. Cressie (1993) estimated the first-order intensity function of the process by kernel smoothing, based on which he concluded nonstationarity for the data. Stoyan and Stoyan (1996) fit a generalized Thomas process model, which was stationary, to these locations and obtained a good fit. Thus Stoyan and Stoyan’s result appeared to contradict with Cressie’s conclusion. Several other analysis, e.g. Politis and Sherman (2001), Guan et al. (2007), assumed stationarity for the data but did not formally test it.
We applied our proposed method to test stationarity for the *longleaf pine* data. A formal test was needed due to the contradictive conclusions in literature regarding whether the process was stationary. Based on the simulation evidence in Section 4, we used the test statistic defined in (13) because of its better performance. To estimate $\sigma^2$, we used the nonparametric estimator given in (8). From the empirical PCF plot (see Web Figure 1), the dependence appeared to be ignorable after $15 - 20m$. This was further supported by fitting the stationary parametric Poisson cluster process model used in the simulation to the data. Using Cressies (1993)’s estimates, the PCF was equal to 1.052 for 15m and 1.003 for 20m. We also considered the generalized Thomas process models fitted by Stoyan and Stoyan (2007). The PCF for 20m was equal to 1.009 and 1.016 for the two models fitted by least squares therein and equal to 1.043 for the model fitted by trial-and-error. In view of these results, we set $m_n = 20m$. The resulting test statistic was equal to 1.11, which was only slightly smaller than the critical value at the 5% level ($= 1.1196$) but larger than that at the 10% level ($= .9322$). Thus there was moderately strong evidence that the underlying process was nonstationary.

5.2 *Ambrosia Dumosa* data

The *Ambrosia Dumosa* data consist of locations of 4358 *Ambrosia Dumosa* plants, recorded in the 1984 census within a $100m \times 100m$ square region in the Colorado Desert (see Figure 2). *Ambrosia Dumosa* is an extremely abundant, long-lived and drought deciduous shrub which is often found on well drained soils below 1061 m elevation (Miriti *et al.* 1998). In the study site, it accounts for approximately 62% of all the encountered perennial plant species.
In a previous study of the data, Perry et al. (2002) noted several low-intensity strips running roughly in the north-west to south-east direction. Guan et al. (2006) attributed the cause of these strips to anisotropy. By using a formal testing method, they found strong evidence for anisotropy. They further suggested that the effects of directional shading on seed germination and the survival of young seedlings might have caused the anisotropic pattern in the data.

To conduct their test, Guan et al. (2006) assumed stationarity for the underlying process. This assumption should be formally tested due to two reasons. Firstly, it was an important assumption for Guan et al. (2006)’s test and thus should be verified. Secondly, and most importantly, the result of the test would have a strong implication on how to interpret and model the data. Specifically, a rejection of stationarity would warrant the necessity to consider the effect of possibly nonstationary environmental variables on the distribution of Ambrosia Dumosa. It might well be the case that the reported strips were due to nonstationarity in the underlying environment but not due to anisotropy.

We applied our proposed method to test stationarity for the Ambrosia Dumosa data. As in the longleaf pine data example, we used the test statistic defined in (13). We used $m_n = 2m$ as the bandwidth for the nonparametric variance estimator given in (8). $2m$ was used based on the finding in Guan et al. (2006) that the dependence appeared to vanish after a distance of $1.5−2m$. We further verified the validity of this choice by fitting the stationary Poisson cluster model used in the simulation to the data. Based on the fitted model,
the PCF for \(2m\) was almost equal to one, which supported our choice. The resulting test statistic was equal to 1.561, which was only slightly smaller than the critical value at the .01 level (= 1.562). Thus there was a strong evidence that the underlying process was nonstationary.

REFERENCES


the Null Hypothesis of Stationarity against the Alternative of a Unit Root”, *Journal of Econometrics*, 54, 159–178.


Figure 1. The longleaf pine data.
Figure 2. The *Ambrosia Dumosa* data.
Empirical sizes ($\beta = 0$) and powers ($\beta = 1, 2$) for the proposed tests in the Poisson cluster process case. The nominal level was 5%. The test statistic defined in (9) used the lower-left corner of the region as the origin.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$n$</th>
<th>$\sigma$</th>
<th>(9) for different $m_n$</th>
<th>(13) for different $m_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$3\sigma$</td>
<td>$4\sigma$</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>0.2</td>
<td>0.052</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.037</td>
<td>0.034</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.2</td>
<td>0.065</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.063</td>
<td>0.049</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>0.2</td>
<td>0.074</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.048</td>
<td>0.043</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.2</td>
<td>0.476</td>
<td>0.420</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.406</td>
<td>0.339</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.2</td>
<td>0.253</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.182</td>
<td>0.123</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.2</td>
<td>0.966</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.935</td>
<td>0.899</td>
</tr>
</tbody>
</table>

21
Table 2

Empirical sizes ($\beta = 0$) and powers ($\beta = 1, 2$) for the proposed tests in the simple inhibition process case. The nominal level was 5%. The test statistic defined in (9) used the lower-left corner of the region as the origin.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$n$</th>
<th>(9) for different $m_n$</th>
<th>(13) for different $m_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>0.044</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>0.606</td>
<td>0.580</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.996</td>
<td>0.996</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>