Statistical Inference for Stochastic Processes (2006) 9:31–46
DOI 10.1007/s11203-004-2894-2


# Spatial Point Process Models of Defensive Strategies: Detecting Changes

JOHN KORNAK, MARK E. IRWIN and NOEL CRESSIE*
*Department of Statistics, The Ohio State University, Columbus, OH 43210, USA*

(Received 27 July 2003; Accepted 18 April 2004)

**Abstract.** The study of stochastic processes can take many forms. Theoretical properties are important to ensure consistent model definition. Statistical inference on unknown parameters is equally important but can be difficult. This is principally because many of the standard assumptions for proving consistency and asymptotic normality of estimators involve independence and homogeneity. In the case where inference is concerned with detecting change in a spatial process from one time point to another, a statistical-computing approach can be rewarding. Regardless of the complexity of the stochastic process, if simulating from it is relatively easy, then detecting change is possible using a Monte Carlo approach. The methodology is applied in a military scenario, where a country's defensive posture changes as a function of its perceived threat. For tactical-decision purposes, it is extremely important to know whether the country's perceived threat level has changed.

**AMS Subject Classification:** Primary: 62M30, Secondary: 60G55.

**Key words:** Poisson point process, kernel smoothing, intensity function, Monte Carlo testing, conditional power, geopolitical tendencies.

## 1. Introduction

This article is concerned with inference on spatial stochastic processes. The two general problems considered are: (1) detect whether an observed, possibly incomplete realization of the process comes from a hypothesized probability measure; and (2) detect whether the underlying probability measure has changed from one realization to another. Crucial to our approach is the ability to simulate from the relevant classes of spatial stochastic processes.

In the research presented below, we solve an applied-probability problem of relevance to the intelligence community. Understanding and quantifying the *behavioral psychology* of enemy nations under conflict situations (or perceived conflict situations) can lead to improved tactical counter-measures. A variety of important strategic questions can be answered that can provide extremely useful intelligence; for example, when threatened, does the enemy

*Author for correspondence: Tel.: 614-292-5194; Fax: 614-292-2096; e-mail: ncressie@stat.ohio-state.edu.

tend to retreat and defend, or aggressively counter-attack? Does the enemy attack in isolated pockets, or in a more uniform manner?

In this article, we look to answer these questions in specific situations where intelligence data give the position and readiness-state of hostile mobile launcher systems. The affiliation and potential threat of mobile launcher systems can vary significantly under different readiness states. By smoothing the spatial point process of mobile launchers (at a given snapshot in time), we obtain intensity maps that quantify the potential threat the launchers imply.

Our interpretation and application of this has led to two approaches. The first is a *global* approach that models spatial locations of all detectable mobile launchers in a potentially hostile country at successive snapshots in time. The objective is to determine whether, and if so to what extent, the country is deploying positional changes in readiness to attack.

The second is a *local* approach. Here we characterize the country's readiness-state by the behavior of the mobile-launcher intensity at the same successive snapshots in time, but over *selected regions* of the country. The objective is the same as for the global problem.

Intelligence data is not generally good enough to track individual launcher locations through time, and hence the applied-probability problem we pose involves spatial point processes observed at different snapshots in time. As can be seen in Figure 1, it may be difficult to detect differences between intensity functions generating the launcher locations by just viewing the observed launcher fields. When going from left to right in the figure, the true intensity functions for the fields involve launchers being shifted towards the northern and eastern regions of the country, representing low-, intermediate-, and high-threat situations. While it is clear from the figure that the high-threat situation in (c) is different than the other two, differentiating between (a) and (b) by eye is difficult. However, the difference between the first two fields can be described and detected through the use of summaries and hypothesis tests.

In Section 2 of this article, we describe our approach, which is based on Monte Carlo inference. We employ Monte Carlo procedures on the global and local problems based on data like the mobile-launcher intensity maps referred to earlier. Section 3 gives the smoothing method used to obtain the mobile-launcher intensity maps and suggests summary functionals that can be used for decision making, in both quantitative and descriptive ways. Section 4 contains the results of our change-detection procedure applied to synthetic data, as well as results on the power of the global and local approaches. Section 5 gives some discussion and conclusions.

## 2. Stochastic Spatial Model of Mobile-launcher Sites

As there are no unclassified data of this spatial type available, we generated synthetic data according to the following model. Assume that the country

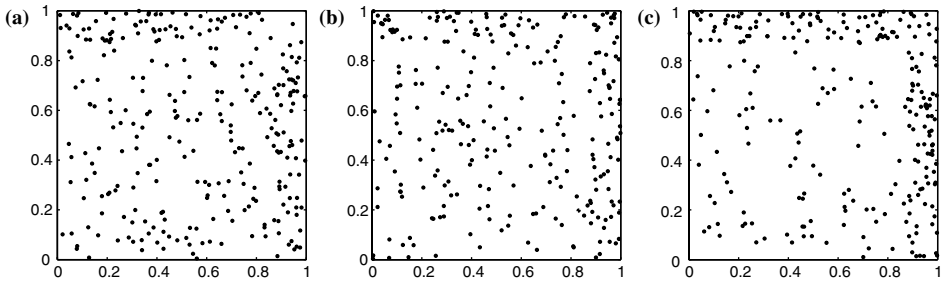


*Figure 1.* (a) low-threat (b) intermediate-threat and (c) high-threat launcher locations.

$D \subset \mathbb{R}^2$ has been divided into $I$ disjoint regions $A_1, A_2, \ldots, A_I$. At fixed time $t$, assume that for a region $A_i$ $(i = 1, \ldots, I)$, the intensity of launchers (i.e. the average number of launchers per unit area) is constant (locally homogeneous), equal to $\lambda_i(t)$. This assumption is made for convenience of presentation and it is a simple matter to generalize the development to more smoothly varying spatial intensities. Then, for each region $A_i$, launchers are assumed distributed according to a Poisson process with intensity $\lambda_i(t)$ at time $t$. For the Poisson process, the expected number of launchers in region $A_i$ is $\lambda_i(t)|A_i|$, where $|A_i|$ is the area of region $A_i$. Furthermore, the locations of the launchers in $A_i$, conditional on their number, is uniformly distributed in $A_i$ (e.g. Cressie, 1993, Chapter 8).

### 2.1. LOW-THREAT AND HIGH-THREAT

The condition of low-threat could be considered as an equilibrium (or resting) defensive state. This would be a typical level of readiness for the country when its neighbor is not preparing to attack. We consider a high-threat/high-readiness state to occur when the country is moving more launchers into strategic areas, here the northern and eastern border regions. However, another potentially interesting change would occur if the country pulls its launchers back inland, perhaps to defend major cities or weapons caches from possible air attacks. In what is to follow, we use the mobile-launcher intensity map to summarize the spatial distribution of mobile launchers throughout the country. Then we determine tactical behavioral changes by summarizing the intensity map and using the summary functionals to test for increases or decreases over time.

A body of work exists in the point-process literature for detecting regions of elevated intensity in a homogeneous background; see, for example, Allard and Fraley (1997), Dasgupta and Raftery (1998), and Byers and Raftery (1998). However, the methods therein differ from the approach described here in that we are looking to detect relative changes in intensity, at *different* snapshots in time and between pre-specified local regions. Naus (1965)

proposed the scan statistic to test for the homogeneity (in terms of intensity) of a Poisson point process in both space and time. Recent literature exists on developing tests based on a derived null distribution of this statistic under the null model of a constant intensity function; see, for example, Chen and Glaz (1996), Alm (1997), Priebe et al. (1997), Kulldorff (1999), and Glaz et al. (2001). Our contribution is to test for *changes* in intensity patterns with test statistics that are measurable functionals of the point pattern, to admit specifically the global/local aspects of the problem, and to solve the associated inferences with Monte Carlo methodology like that used by Turnbull et al. (1990) and Kulldorff (1997). Our aim is to demonstrate how certain inferences on stochastic processes are possible when simulation is easy.

### 2.2. EXAMPLE

To illustrate the inference procedures, we adopt a simple scenario. Let the 'country' of interest be represented by the unit square ($D = [0,1] \times [0,1]$). Assume that there are two important regions, $A_1 = [0, 0.875] \times [0, 0.875]$, the mostly interior part of the country, and $A_2 = D/A_1$, the northern and eastern border region where the launcher intensity will likely increase at a time of increased threat from these directions.

In what is to follow, we generate synthetic data that represent launcher locations at three snapshots $t_1, t_2, t_3$. We assume that $\lambda_1(t_1)|A_1| + \lambda_2(t_1)|A_2| = \lambda_1(t_2)|A_1| + \lambda_2(t_2)|A_2| = \lambda_1(t_3)|A_1| + \lambda_2(t_3)|A_2|$, which keeps the number of launchers at the three snapshots approximately the same. (That is, the expected number of launchers at $t_1, t_2$, and $t_3$ are equal. This assumption can be relaxed in a straightforward manner.) We choose intensities given by $\lambda_1(t_1) = 199.77$, $\lambda_2(t_1) = 499.42$; $\lambda_1(t_2) = 191.47$, $\lambda_2(t_2) = 526.54$; and $\lambda_1(t_3) = 139.35$, $\lambda_2(t_3) = 696.77$. These choices were based on defining threat level via the ratio $\delta_{jk}(t) = \frac{\lambda_k(t)}{\lambda_j(t)}$. The initial time $t_1$ corresponds to a period of low threat ($\delta_{12}(t_1) = 2.5$), the middle time $t_2$ corresponds to a period of intermediate threat ($\delta_{12}(t_2) = 2.75$), and the final time $t_3$ corresponds to a period of high threat ($\delta_{12}(t_3) = 5$) from the north and the east. The resulting data are the simulated point patterns shown in Figure 1.

## 3. Detecting Changes

Assume that at a particular time point there are $N_D$ launchers in the country $D$, occurring at locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{N_D}$. We wish to smooth this spatial point process to obtain the intensity of launchers at any location $\mathbf{s} \in D$. The approach proposed is based on kernel smoothing, which has the advantage of not making any assumptions about the true intensity function and the methodology is easily implemented. That is, we avoid inference based on the likelihood, which requires parametric assumptions about the intensity function. In practice, we choose a discrete, regular grid upon which to carry out

kernel smoothing. The choice of resolution for the grid requires a trade-off between spatial resolution and computational time, and it is not critical for the inference methodology described below. Since the grid is used for visual mapping, we shall refer to each point on the grid as a pixel. (Though strictly speaking, it is the pixel location).

### 3.1. KERNEL INTENSITY ESTIMATION

To obtain an estimate of the launcher intensity at any point $\mathbf{s} \in D \subset \mathbb{R}^2$, a kernel smoother can be used (e.g. Cressie, 1993, p. 653). The smoother is of the form,

$$\hat{\lambda}_h(\mathbf{s}) = \frac{1}{p_h(\mathbf{s})} \sum_{i=1}^{N_D} h^{-2} w\Big(\frac{\mathbf{s} - \mathbf{s}_i}{h}\Big), \tag{1}$$

where

$$p_h(\mathbf{s}) = \int_D h^{-2} w\Big(\frac{\mathbf{u} - \mathbf{s}}{h}\Big) \mathrm{d}\mathbf{u}. \tag{2}$$

The function $w$ can be any symmetric probability density function around the origin, and $h > 0$ is known as the *bandwidth*. Two popular choices for the kernel function $w$ are the bivariate Gaussian density function and the Epanechnikov kernel (Cressie, 1993, p. 654). We propose the use of the Gaussian density function,

$$w(\mathbf{u}) \equiv \frac{1}{2\pi} \exp(-\mathbf{u}'\mathbf{u}/2), \tag{3}$$

for its ease of implementation. The choice of the bandwidth $h$ determines the smoothness of the estimated intensity surface. Factors involving its choice will be discussed in the next section. The quantity $p_h(\mathbf{s})$ is an edge-correction term accounting for the fact that launchers do not occur outside the region $D$. If $D$ is the whole plane, then $p_h(\mathbf{s}) = 1$, for all $\mathbf{s} \in D$. For bounded $D$, such as in our situation, $p_h(\mathbf{s})$ is typically less than 1 in regions close to the border and decreases as $\mathbf{s}$ approaches it. (Here 'close' is defined as within the kernel bandwidth.)

### 3.2. KERNEL-BANDWIDTH SELECTION

When the Gaussian kernel is adopted, the bandwidth $h$ is in fact the standard deviation of the Gaussian density, and it is centered on each launcher location $\mathbf{s}_i$. As mentioned above, the choice of $h$ determines the smoothness of the estimated intensity surface. A small value for $h$ will tend to lead to a rough, bumpy surface, whereas a larger value will tend to lead to a smooth, flat surface. An important factor in choosing $h$ is knowing how clustered launcher locations tend to be.

We provide a means to obtain a default value for $h$, based on the *Matched Filter Theorem* (MFT) (e.g. Rosenfeld and Kak, 1976, pp. 302–306). In words,

the MFT states that the optimal form of a linear smoothing kernel that is required to extract features of interest from a signal should match the shape of the features that are being searched for. In the simulation used to generate the data, we have a raised vertical wall of intensity of width 0.125 at the northern and eastern borders; therefore, in order to match the shape of the raised signal, the optimal filter would be a raised square of dimension $0.125 \times 0.125$. However, in real scenarios, we would expect to see intensity functions that change more smoothly. For example, we may expect very high intensity within a couple of miles of a border, but then the intensity will decrease gradually as one moves further away from the border. Therefore, we suggest (and consequently adopt) a 2-D Gaussian kernel, which has the ability to match features that vary more smoothly. With this choice of kernel in mind, we attempt to match the expected 'width' of features of interest even though we do not match the shape. In this way, it is possible to observe how robust the model is to either sub-optimal choices of kernel shape, or unusually extreme intensity maps. This width-matching is achieved by choosing a kernel that matches the expected width of map features (in this case, border regions) to the *full width at half maximum* (FWHM) of the marginal (1-D) Gaussian kernel. FWHM is defined as the width of the Gaussian kernel at half of its maximum height and has been adopted as a measure of bandwidth by the engineering, signal-processing, and imaging communities (e.g. Tubbs et al., 2002). Friston (2002) uses FWHM kernel-matching to smooth functional brain imaging maps in order to improve the signal-to-noise ratio that occurs in response to a stimulus. The choice of the Gaussian smoothing kernel width is made by matching the FWHM of the kernel with that of the smallest expected dimension of the region of activation.

The following formula is used to obtain the required bandwidth $h$ of a Gaussian kernel for a particular FWHM,

$$h = \frac{\text{FWHM}}{2\sqrt{2\log 2}}. \tag{4}$$

For example, when setting the FWHM equal to the border width of 0.125, as used in our simulated example, we find that $h = 0.0531$.

### 3.3. SUMMARIES AND COMPARISONS

A global approach to describing the region of interest $D$ is *via* the (estimated) spatial variance. This summary functional is used as a measure of the 'roughness' of the intensity surface over $D$. It has the form,

$$\hat{S}^2(t) \equiv \frac{1}{|D|}\int_D (\hat{\lambda}(\mathbf{u}) - \widehat{\bar{\lambda}})^2 d\mathbf{u}, \tag{5}$$

where

$$\widehat{\bar{\lambda}} \equiv \frac{1}{|D|}\int_D \hat{\lambda}(\mathbf{u})\mathrm{d}\mathbf{u} \tag{6}$$

is the mean estimated intensity of launchers over $D$ at time $t$, and $\hat{\lambda}(\cdot)$ is given by Equation (1). (For ease of notation, we have dropped dependence on the bandwidth $h$). Note that we cannot obtain $\hat{S}^2(t)$ directly from Equation (5) due to the gridding of $D$. However, integrals of this type can be approximated for any measurable function $v(\mathbf{u})$ using,

$$\frac{1}{|C|}\int v(\mathbf{u})\mathrm{d}\mathbf{u} \approx \frac{1}{N_C}\sum_{i=1}^{N_C} v(\mathbf{u}_i), \tag{7}$$

where $C \subset D$ is the region that the statistic is to be calculated over, and the summation is over the $N_C$ pixels in the region.

The larger the value of $\hat{S}^2(t)$, the more variable is the estimated intensity surface. Thus, it can be used to examine global changes in the threat level over time. Define the variance ratio,

$$\hat{R}(t_1, t_2) \equiv \frac{\hat{S}^2(t_2)}{\hat{S}^2(t_1)}; \tag{8}$$

no change from $t_1$ to $t_2$ corresponds to $\hat{R}(t_1, t_2)$ taking values close to 1. Note that these global measures do not depend on breaking up $D$ into regions (which is required for local measures). However, their global nature implies that locally different intensity surfaces can give the same global spatial variance and variance ratios; distinguishing between the surfaces requires local measures.

In order make more local inferences, we consider summary functionals based on breaking up the country into sub-regions of interest, $B_1, B_2, \ldots, B_J$. (These sub-regions may be different from the disjoint regions $A_1, \ldots, A_I$ of homogeneous Poisson intensity described in Section 2). The $J$ regions can be chosen based on military intelligence or geopolitical concerns (e.g. $B_1$ might be a sensitive border region). While the method would allow for the possibility of regions that vary over time, for the results presented in this paper we shall assume that they are fixed. For time $t$ and each region $B_j$, define

$$\hat{M}_j(t) \equiv \int_{B_j} \hat{\lambda}(\mathbf{u})\mathrm{d}\mathbf{u}, \tag{9}$$

where $\hat{\lambda}(\cdot)$ is given by Equation (1). Also, define

$$\hat{\Lambda}_j(t) \equiv \frac{\hat{M}_j(t)}{|B_j|}. \tag{10}$$

The summary functional $\hat{M}_j(t)$ estimates the mean number of launchers in region $B_j$ under the readiness conditions observed at time $t$, whereas $\hat{\Lambda}_j(t)$ estimates the average intensity in the region. To compare the launcher deployment in two regions, $B_j$ and $B_k$, at a particular time $t$, we propose the summary functional,

$$\hat{\delta}_{jk}(t) \equiv \frac{\hat{\Lambda}_j(t)}{\hat{\Lambda}_k(t)}, \tag{11}$$

the ratio of the average intensities for the two regions. When the ratio $\hat{\delta}_{jk}(t) > 1$, the launcher locations are more dense in region $B_j$, while $\hat{\delta}_{jk}(t) < 1$ implies the opposite.

To examine how launcher deployment changes from time $t_1$ to time $t_2$ in region $B_j$, two possible summaries are

$$\hat{\gamma}_j(t_1, t_2) \equiv \frac{\hat{\Lambda}_j(t_2)}{\hat{\Lambda}_j(t_1)} = \frac{\hat{M}_j(t_2)}{\hat{M}_j(t_1)}, \tag{12}$$

and

$$\hat{\eta}_j(t_1, t_2) \equiv \hat{M}_j(t_2) - \hat{M}_j(t_1). \tag{13}$$

A natural summary to compare how the launcher-deployment ratio for the two regions changes over time is

$$\hat{\Delta}_{jk}(t_1, t_2) \equiv \frac{\hat{\delta}_{jk}(t_2)}{\hat{\delta}_{jk}(t_1)}. \tag{14}$$

Note that $\hat{\delta}_{jk}(t)$ is similar to an odds measure, and hence $\hat{\Delta}_{jk}(t_1, t_2)$ is similar to an odds ratio.

### 3.4. MONTE CARLO-BASED INFERENCE

For many of the usual models for a spatial point process (e.g. Cressie, 1993, Section 8.5), the distributions of the summary functionals discussed in the previous section are analytically intractable. Since it is assumed that we can simulate from the spatial point process, we propose the use of Monte Carlo tests (Barnard, 1963; Hope, 1968; Birnbaum, 1975; Besag and Diggle, 1977; Cressie, 1993, p. 635) to examine whether the launcher intensities have changed.

Let $U$ be a test statistic of interest, such as $\hat{S}^2(t)$, $\hat{\Lambda}_j(t)$ or $\hat{\delta}_{jk}(t)$. Suppose that at time $t_1$, the launcher-allocation model is known from historical data (e.g. corresponding to low threat); that is, the intensity function at $t_1$ is known. Call this the null model. We wish to determine whether the launchers are allocated under the same model at time $t_2$. That is, we wish to know whether the statistic observed at time $t_2$ is consistent with the known low-threat intensity map that underlies the locations observed at time $t_1$.

We now describe the Monte Carlo test that is the basis of our inference methodology to detect changes. We simulate $(K-1)$ realizations of the launcher field under the null model at time $t_1$ and calculate the corresponding test statistics, denoted by $U_2, \ldots, U_K$. Let the observed value of $U$ at time $t_2$ be denoted by $U_1$. The observed value $U_1$ is checked to see if it is 'consistent with' the other simulated $U_2, \ldots, U_K$. To achieve this, we look at the ordered $\{U_i\}$, denoted by $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(K)}$, and we reject the null model if $U_1$ is one of the more extreme values. For example, suppose that we are interested in testing whether there is increased clustering of launchers at time $t_2$

using the spatial variance, $U_1 = \hat{S}^2(t_2)$. In this case, large values of $U_1$ indicate evidence of clustering of launchers.

Suppose that $U_1 = U_{(l)}$ for some $l \in \{1, \ldots, K\}$; then we reject the null model of no change in launcher intensity (in favor of increased heterogeneity) at a significance level $\alpha$ if $(K + 1 - l)/K \leq \alpha$. This yields an exact level-$\alpha$ test. So, for example, if $K = 100$, rejection at the $\alpha = 0.05$ level occurs if $U_1$ is one of $U_{(96)}, \ldots, U_{(100)}$. For a two-sided test, $H_0$ should be rejected if $l/K \leq \alpha/2$ or $(K + 1 - l)/K \leq \alpha/2$. The two-sided situation would occur, for example, if one wanted to see whether the mean intensity in region $B_j$, as measured by $\hat{\Lambda}_j(t)$, either increases or decreases.

The Monte Carlo tests described above examine whether the intensity at time $t_2$ is consistent with the null-intensity model at time $t_1$; this is done by comparing what we have observed at time $t_2$ with a set of simulated realizations generated under the null model at time $t_1$. Thus, the test statistics used in these tests only depend on $t_2$, the time of interest (e.g. $\hat{S}^2(t_2)$, $\hat{\Lambda}_j(t_2)$, or $\hat{\delta}_{jk}(t_2)$). Summary functionals that depend on two times, such as $\hat{\gamma}_j(t_1, t_2)$, can also be used for testing, and they are appropriate when we have little information as to what is a low-threat or a high-threat intensity map for the country; that is, we are unable to define specifically the null-intensity model, but we do have a realization at each time $t_1$ and $t_2$ from which to detect whether a change has occurred.

The summaries based on two times can be used to examine whether two observed launcher patterns could have been generated by the same intensity function, without having to pre-specify a null-intensity model. The behavioral question we are asking here is: 'Is there any *change* in the intensity map between times $t_1$ and $t_2$ when the only information that we have about launcher-allocation patterns, at either time point, comes from observed launcher locations?'. This question could be posed in terms of either global or local summaries (Section 3.3). The testing procedures that we present are analogous to the standard two-sample $t$-test for comparing two means. Suppose that the launcher patterns at times $t_1$ and $t_2$ are both generated by an intensity function $\lambda(\mathbf{s})$. Denote the intensity functions estimated from the data for the two times by $\hat{\lambda}(\mathbf{s}; t_1)$ and $\hat{\lambda}(\mathbf{s}; t_2)$. If both launcher patterns were generated under the same intensity function, then both $\hat{\lambda}(\mathbf{s}; t_1)$ and $\hat{\lambda}(\mathbf{s}; t_2)$ are estimates of the common function. Further, the combined estimate, $\hat{\lambda}^c(\mathbf{s}) = (\hat{\lambda}(\mathbf{s}; t_1) + \hat{\lambda}(\mathbf{s}; t_2))/2$ is a better estimator under the *null model of no change* than either of the previous two, since it pools the information from both data sources. The procedure for a Monte Carlo test based on the combined estimate is as follows. First, generate $(K - 1)$ *pairs* of independent point patterns from the estimate $\hat{\lambda}_h^c(\mathbf{s})$; for any pair, the first point pattern corresponds to time $t_1$ and the second corresponds to time $t_2$. Then obtain an estimate of the intensity for each of the pairs using the kernel smoother described in Section 3.1. For the $i$-th pair of estimates,

obtain a summary based on the two times $t_1$ and $t_2$; for example, suppose we choose the local summary $\hat{\Delta}_{jk}(t_1, t_2)$ and compute $U_i \equiv \hat{\Delta}^i_{jk}(t_1, t_2)$ for the $i$-th pair; $i = 2, \ldots, K$. This gives $(K-1)$ simulated statistics under the null hypothesis. Then we compare the observed value $U_1 \equiv \hat{\Delta}_{jk}(t_1, t_2)$ with the $(K-1)$ simulated values $\{U_2, \ldots, U_K\}$, as was done for the single-time-point test. This will yield an *approximate* level-$\alpha$ test; we lose exactness because the data are generated from an (combined) estimate of the intensity.

For all of the tests described herein, the choice of the number of simulated data sets to be acquired, $(K-1)$, depends principally on two factors. The first is $\alpha$, the type-I error rate for the test. The smaller $\alpha$ is, the larger $K$ needs to be. In this setting, $\alpha$ should not be very small, since we are interested in detecting threat changes early; $\alpha$ somewhere in the range 0.01–0.10 is reasonable. For this range of $\alpha$, $(K-1) = 999$ simulated data sets will be adequate. The second factor is the speed with which an answer to the hypothesis test is needed. Fewer simulated data sets yields a quicker result. However, when required, a large number of data sets may be generated ahead of time under known intensity functions (e.g. under low-threat, peace-time conditions), with the necessary summaries stored. In this case, a newly observed test statistic can be compared quickly with the stored values in order to perform the desired test. Past intelligence data concerning the spatial intensity of mobile launchers could be used to determine the null-intensity model of interest.

### 3.5. POWER OF THE MONTE CARLO TESTS

For the Monte Carlo tests that were used to examine whether a spatial intensity function differs from a pre-specified null intensity function, a number of factors influence the tests' operating characteristics. Two important ones are $U$, the summary of the spatial point process used as the test statistic, and $(K-1)$, the number of simulations from the null intensity function.

To study the effect of these two factors, a small simulation study was performed. Two test statistics were studied: the spatial variance $\hat{S}^2(t)$ (global test) and the mean intensity ratio $\hat{\delta}_{jk}(t)$ (local test). For the local test, the regions used for generating the data were used for the summaries (i.e. $B_i = A_i$, for all $i$). Two values for $(K-1)$ were chosen: 199 and 999. Further factors include the type-I-error rate (3 levels: $\alpha = 0.01, 0.05, 0.10$), and the mean-intensity ratio under the null hypothesis (2 levels: $\delta^0_{jk} = 1, 2.5$). All tests' powers were evaluated at the alternative-hypothesis, mean-intensity ratios, $\delta^A_{jk} = 1, 1.25, 1.5, \ldots, 5$.

Let $U_{2:K} \equiv \{U_2, U_3, \ldots, U_K\}$ denote the $(K-1)$ realizations of a test statistic simulated under the null hypothesis $H_0$. Then the power of the test is,

$$\pi(\delta^A_{jk}; U, K-1, \alpha, \delta^0_{jk}) = P_{H_A}[\text{Reject} H_0] \tag{15}$$

$$= E_{H_0}[P_{H_A}[\text{Reject} H_0 | U_{2:K}]] \tag{16}$$

$$\equiv E_{H_0}[\tilde{\pi}(\delta^A_{jk}; U, K-1, \alpha, U_{2:K})], \tag{17}$$

where $\tilde{\pi}(\delta^A_{jk}; U, K-1, \alpha, U_{2:K})$ is the conditional power, that is, the power given the sample $U_{2:K}$.

Now assume that the form of the Monte Carlo test is to reject $H_0$ if $U_1 > z^*(U_{2:K}, \alpha)$. The critical value $z^*(U_{2:K}, \alpha)$ will be the $K - \lfloor K\alpha \rfloor$ order statistic of $U_{2:K}$, where $\lfloor x \rfloor$ denotes the largest integer $\leqslant x$. Then

$$\pi(\delta^A_{jk}; U, K-1, \alpha, \delta^0_{jk}) = E_{H_0}[P_{H_A}[U > z^*(U_{2:K}, \alpha)]]. \tag{18}$$

Note that similar relationships hold for the the lower one-sided alternative and for the two-sided alternative. For the lower one-sided alternative, the critical value is the $\lfloor K\alpha \rfloor$ order statistic from $U_{2:K}$, and for the two-sided alternative, the two critical values are the $\lfloor K\frac{\alpha}{2} \rfloor$ and the $(K - \lfloor K\frac{\alpha}{2} \rfloor)$ order statistics from $U_{2:K}$.

Based on Equation (18), the power can be estimated easily by the following Monte Carlo scheme. For each $H_0$ of interest, simulate $L$ sets of $(K-1)$ simulated launcher patterns, and calculated the statistics, $U^{(1)}_{2:K}, \ldots, U^{(L)}_{2:K}$. For the $l$-th launcher pattern $(l = 1, \ldots, L)$, calculate the test's critical value $z^*(U^{(l)}_{2:K}, \alpha)$ for each significance level $\alpha$. Then, under each $H_A$, independently generate $M$ launcher patterns and calculate their summary functionals $U^{(1)}, \ldots, U^{(M)}$. Finally $\pi(\delta^A_{jk}; U, K-1, \alpha, \delta^0_{jk})$ can be unbiasedly estimated by,

$$\hat{\pi}(\delta^A_{jk}; U, K-1, \alpha, \delta^0_{jk}) \equiv \frac{1}{LM} \sum_{l=1}^{L} \sum_{m=1}^{M} I(U^{(m)} > z^*(U^{(l)}_{2:K}, \alpha)). \tag{19}$$

These two sets of independent simulations, one set under $H_0$ and the other set under $H_A$, are the basis of a Monte Carlo approach to estimating power for all combinations of $H_0, H_A$, and $\alpha$. This approach also has the property that for a fixed $H_A$, the estimated power is a nondecreasing function of $\alpha$, as is the case with the true power function. Several statistics are compared this way in Section 4.

## 4. Results

The true intensity maps at each of the times $t_1$ (low-threat) to $t_3$ (high-threat) are piecewise constant functions for which the intensity levels for the inland and border regions become more and more separated as the threat level increases. In Figure 1, we show individual point-process realizations of each of the intensity maps, which serve as our synthetic data for times $t_1$, $t_2$, and $t_3$. The increased concentration of launchers in the border region at time $t_3$, compared to that at times $t_1$ and $t_2$, is immediately obvious. However, between times $t_1$ and $t_2$ there is no obviously discernable difference.

Figure 2 shows the smoothed intensity maps at times $t_1, t_2$, and $t_3$. The increased activity at the border for the high-threat condition in panel (c) compared to panels (a) and (b), the lower-threat conditions, is obvious; notice that the intensity map shown in panel (c) also captures the reduced intensity inland.

Monte Carlo tests were performed using $K - 1 = 999$ and test statistics obtained from the three smoothed intensity maps. Both the global test based on $\hat{S}^2(t)$ and the local test based on $\hat{\delta}_{12}(t)$ are testing the (low-threat) null hypothesis $H_0 : \delta_{12}(t) = 2.5$ versus the alternative hypothesis, $H_A : \delta_{12}(t) > 2.5$, resulting in a one-sided test. Notice that in the case of Figure 2(a), $H_0$ is true, whereas for Figure 2(b) and (c), $H_A$ is true. The results of the inference are summarized in Table I. For the high-threat case, both test statistics give strong evidence that the launcher pattern was not generated under $H_0$. In fact, for both statistics, the observed test statistic was more extreme than those for any of the patterns simulated under $H_0$ ($\text{range}(\hat{S}^2(t)) = (10912, 44336)$, $\text{range}(\hat{\delta}_{12}(t)) = (1.503, 3.015)$). For the intermediate-threat case, there is moderate evidence that $H_A$ is true, with the local test rejecting and the global test almost rejecting the null hypothesis at the $\alpha = 0.05$ level. The test statistics for the low-threat case, where $H_0$ is in fact true, are consistent with the null hypothesis.

We also developed tests based on $\hat{R}(t_i, t_j)$ (global) and $\hat{\Delta}_{12}(t_i, t_j)$ (local) to see whether we could detect a change in mobile-launcher deployment between the two times $t_i \neq t_j$. Here the null hypothesis is: No change in the threat, versus the alternative hypothesis: Threat has changed. For the comparison of Low versus High and for Intermediate versus High, there is strong evidence that the pattern generated under the high-threat condition was indeed obtained from a different intensity model than that of the lower-threat conditions, as can be seen in Table II. However, when the patterns generated under the low- and intermediate-threat models are compared, there is little evidence to show that they were obtained from different models, as the $p$-values for the two tests are well above 0.05. Note that this is not in conflict

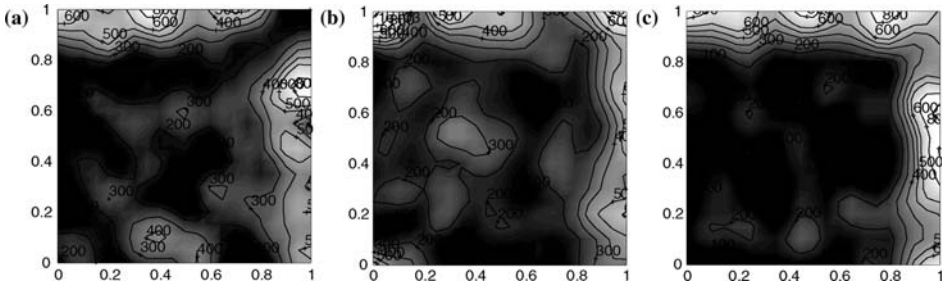


*Figure 2.* Contour plots of estimated intensity field under (a) low-threat (b) intermediate-threat and (c) high-threat conditions.

*Table I.* Test statistics and *p*-values (one-sided) for the three smoothed intensity maps given in Figure 2

| Threat level | $\hat{S}^2(t)$ | *p*-value ($\hat{S}^2(t)$) | $\hat{\delta}_{12}(t)$ | *p*-value ($\hat{\delta}_{12}(t)$) |
|---|---|---|---|---|
| Low | 25689 | 0.257 | 2.193 | 0.339 |
| Intermediate | 30738 | 0.073 | 2.540 | 0.050 |
| High | 54417 | 0.001 | 4.362 | 0.001 |

with the earlier results, since a less-specific null hypothesis is being examined. Also note that for the statistics based on two time points, the associated tests are two-sided, since we are examining if there is a change in the launcher intensity function in either direction (increased or decreased threat).

Because the results of all tests involving the high-threat case were so significant, we conducted a power study, like that described in Section 3.5, to gauge the sensitivity of the tests for threat conditions with much lower differences in intensity. We chose $L = 50$ and $M = 1000$ and compared both the global and local test statistics.

For all appropriate cases examined, the power of the local test based on $\hat{\delta}_{12}(t)$ was greater than the power of the global test based on $\hat{S}^2(t)$; under the null hypothesis, both power functions are approximately equal to $\alpha$. This is not surprising since the spatial variance does not involve any local partitioning and so provides less information about a possible change in intensity. Two examples of power curves are shown in Figure 3; clearly, the local test is consistently more powerful.

The results when $K - 1 = 199$ are similar to the case of $K - 1 = 999$. Only when $\alpha = 0.01$ do noticeable changes in power exist, with the power being lower when $K - 1 = 199$. Since for $\alpha = 0.01$ and $K - 1 = 199$, the critical values are based on the most extreme simulated statistic (two-sided test) or the second most extreme simulated statistic (one-sided test), this result is not particularly surprising. The small differences in power that occur are due to the variability of $\{z^*(U^{(l)}_{2:K}, \alpha) : l = 1, \ldots, 50\}$ in Equation (19). The larger $(K - 1)$ is, the less variable are the Monte Carlo test's critical values, and therefore the less variable is the conditional power. Figure 4 shows a comparison of the power curves for $K - 1 = 199$ versus $K - 1 = 999$, in the case

*Table II.* Test statistics and *p*-values (two-sided) for two-time-point comparisons of the smoothed intensity maps given in Figure 2

| Threat-level comparison | $\hat{R}^2(t_i, t_j)$ | *p*-value ($\hat{R}^2(t_i, t_j)$) | $\hat{\Delta}_{12}(t_i, t_j)$ | *p*-value ($\hat{\Delta}_{12}(t_i, t_j)$) |
|---|---|---|---|---|
| Low versus Intermediate | 1.197 | 0.550 | 1.158 | 0.296 |
| Low versus High | 2.118 | 0.008 | 1.989 | 0.002 |
| Intermediate versus High | 1.770 | 0.046 | 1.718 | 0.002 |

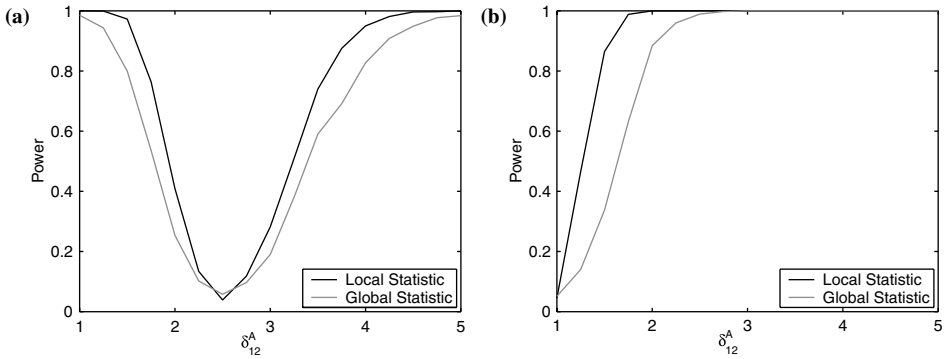


*Figure 3*. Power curves with $\alpha = 0.05$ and $K - 1 = 999$ for (a) $\delta_{12}^0 = 2.5$ and two-sided test, and (b) $\delta_{12}^0 = 1$ and one-sided test.

where $\alpha = 0.01$. Even here, the curves are reasonably close together, suggesting that it may not be necessary to use a very large Monte Carlo sample size $(K - 1)$. For the two choices of $(K - 1)$, Figure 5 shows histograms of the variability in the estimated conditional power, $\{\frac{1}{M}\sum_{m=1}^{M} I(U^{(m)} > z^*(U_{2:K}^{(l)}, \alpha)) : l = 1, \ldots, 50\}$, for $K - 1 = 199$ and 999. The extra variability in conditional power as $(K - 1)$ decreases is clearly seen *via* the increased dispersion in the histogram in panel (b) when compared to that in panel (a).

## 5. Discussion

We have presented an approach for detecting changes in stochastic processes based on comparing observations to Monte Carlo simulations generated under a null hypothesis. While our motivation for developing this approach and our example comes from the the study of geopolitical tendencies, the methods are applicable to a wide range of problems. For example, examining

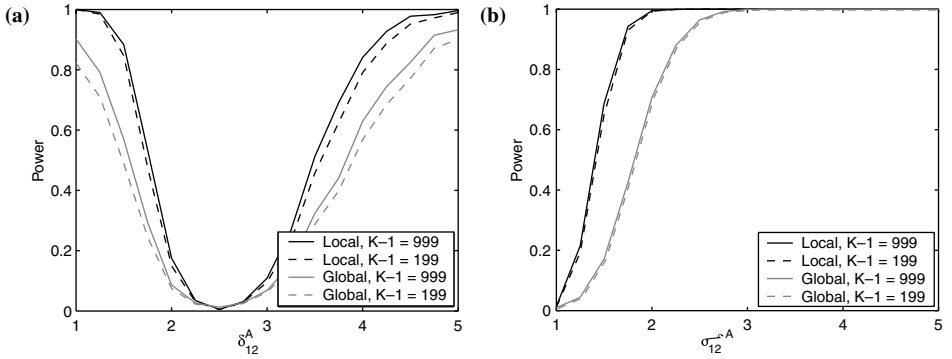


*Figure 4*. Power curves with $\alpha = 0.01$ for (a) two-sided tests with $\delta_{12}^0 = 2.5$, and (b) one-sided tests with $\delta_{12}^0 = 1$; $K - 1 = 199$ versus $K - 1 = 999$.

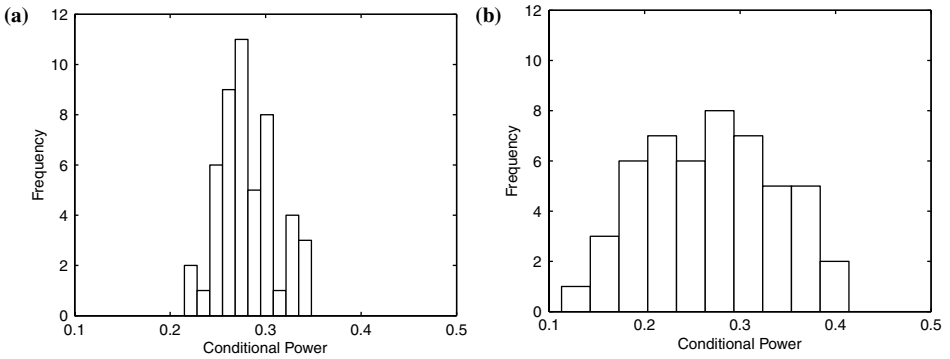


*Figure 5*. Histograms of the conditional power values for the two-sided local test statistic with $\alpha = 0.05$, $\delta_{12}^0 = 2.5$, and $\delta_{12}^A = 3$ for (a) $K - 1 = 999$, and (b) $K - 1 = 199$.

the effects of protection actions on threatened species, such as the recent attempts at relocation of wolves from Canada into the American Rockies, could be carried out using these methods.

The kernel-smoothing techniques proposed for describing the intensity function of the underlying point process are nonparametric. Potentially, this is a great advantage in the mobile-launcher problem, since intelligence information often will not be good enough to supply a parametric model for the intensity function. Also, the smoothed intensity function is defined across the whole spatial domain $D$, and hence any summary functional can be the basis for inference.

A Monte Carlo approach for testing for changes in a given stochastic process or between two stochastic processes, has been proposed. Importantly, since the test is based on simulation under the null hypothesis of no change, complicated approximations to the null distribution of the test statistic are not needed. While any functional could be used to develop a test statistic, careful thought is required to achieve a powerful test. As seen in the power study in Section 4, the global test statistic given by the spatial variance is less powerful than the local test statistic given by the mean intensity ratio. This result is to be expected, since the alternative hypotheses examined in the study were based on the mean intensity ratio.

The Monte Carlo approach proposed in this paper can have power to detect small deviations from the null hypothesis. In the example presented, while the point pattern (and its smoothed intensity function) generated under the intermediate-threat condition ($\delta_{12} = 2.75$) does not look different than the point pattern (and its smoothed intensity function) generated under the low-threat condition ($\delta_{12} = 2.5$), both tests provide reasonable evidence (based on their $p$-values) that a change has occured. One of the most compelling features of this approach for detecting changes, is in settings

where the stochastic process being considered may have properties that are difficult to derive, but simulation from it is straightforward.

## Acknowledgement

This research was partially supported by SSCSD, San Diego, under Contract No. N66001-97-D-5038 with San Diego State University Research Foundation, and partially by the Office of Naval Research under grant N00014-02-1-0052. The authors would like to thank John McDonnell and John Custy of SSCSD for helpful discussions.

## References

Allard, D. and Fraley, C.: Nonparametric maximum likelihood estimation of features in spatial point processes using Voronoi tessellation, *J. Am. Stat. Assoc.* **92** (1997), 1485–1493.

Alm, S. E.: On the distribution of scan statistics of a two-dimensional Poisson process, *Adv. Appl. Probab.* **29** (1997), 1–18.

Barnard, G.: Comment on 'The spectral analysis of point processes' by M. S. Bartlett. *J. Royal Stat. Soc. B* **25** (1963), 294.

Besag, J. E. and Diggle, P. J: Simple Monte Carlo tests for spatial pattern, *Appl. Stat.* **26** (1977), 327–333.

Birnbaum, Z. W.: Testing for intervals of increased mortality. In: Barlow , R. E., Fussell, J. B. and Singpurwalla. N. D. (eds), *Reliability and Fault Tree Analysis*, SIAM, Philadelphia, PA, 1975, 413–426.

Byers, S. and Raftery, A. E.: Nearest-neighbor clutter removal for estimating features in spatial point processes, *J. Am. Stat. Assoc.* **93** (1998), 577–584.

Chen, J. and Glaz, J.: Two-dimensional discrete scan statistics, *Stat. Probab. Lett.* **31** (1996), 59–68.

Cressie, N.: *Statistics for Spatial Data, Revised Edition*, Wiley, New York, 1993.

Dasgupta, A. and Raftery, A. E.: Detecting features in point processes with clutter *via* model-based clustering, *J. Am. Stat. Assoc.* **93** (1998), 294–302.

Friston, K. J.: Statistical parametric mapping. Book manuscript. Introductory chapter available at http://www.fil.ion.ucl.ac.uk/spm/papers/SPM-Chapter.pdf, 2002.

Glaz, J., Naus, J. I. and Wallenstein, S.: *Scan Statistics*, Springer-Verlag, New York, 2001.

Hope, A. C. A.: A simplified Monte Carlo significance test procedure, *J. Royal Stat. Soc. B*, **30** (1968), 582–598.

Kulldorff, M.: A spatial scan statistic. *Communications in Statistics. Theory and Methods* **26** (1997), 1481–1496.

Kulldorff, M. Spatial scan statistics, models, calculations, and applications, In: Glaz, J. and Balakrishnan. N. (eds.), *Scan Statistics and Applications*, Birkhauser, Boston, 1999.

Naus, J. I.: Clustering of random points in two dimensions, *Biometrika* **52** (1965), 263–267.

Priebe, C. E., Olson, T. and Healy, D. M. Jr: A spatial scan statistic for stochastic scan partitions, *J. Am. Stat. Assoc.* **92** (1997), 1476–1484.

Rosenfeld, A. and Kak, A. C.: *Digital Picture Processing*, Academic Press, New York, 1976.

Tubbs, R. N., Baldwin, J. E., Mackay, C. D. and Cox, G. C.: Diffraction-limited CCD imaging with faint reference stars, *Astron. Astrophys.* **387** (2002), L21–L24.

Turnbull, B. W., Iwano, E. J., Burnett, W. S., Howe, H. L. and Clark, L. C.: Monitoring for clusters of disease: application to leukemia incidence in upstate New York, *Am. J. Epidemiol.* **132** (1990), S136–S143.