J. Japan Statist. Soc.
Vol. 36 No. 1 2006 1–15

PREDICTION OF THE SAMPLE VARIANCE OF MARKS FOR A MARKED SPATIAL POINT PROCESS BY THE THRESHOLD METHOD

Takayuki Sakaguchi* and Shigeru Mase**

We discuss the prediction of the sample variance of marks of a marked spatial point process on a continuous space by the threshold method. The threshold method is a statistical prediction using only the number of points with marks exceeding a given threshold value. Mase (1996) considered the method in the framework of spatial point processes on a discrete space and Sakaguchi and Mase (2003) extended the results of Mase (1996) to a continuous space. They considered the prediction of the sum of marks. In the present paper, it is shown that the sample variance of marks can be also predicted well if a point process is non-ergodic and marks satisfy some mixing-type condition. A simulation study is given to confirm the theoretical result.

Key words and phrases: Marked spatial point process, mean square error, mixing condition, non-ergodicity, threshold method.

1. Introduction

We discuss the threshold method in the framework of a marked spatial point process $\Phi = \{(X_i, M_i)\}$ on $\mathbb{R}^d$, which is a set of pairs of an observational position $X_i \in \mathbb{R}^d$ and a corresponding random variable M_i called the mark. Sometimes it may be difficult or impossible to observe comparatively small marks and this causes a trouble for calculating the sample mean or the sample variance of marks. The threshold method is a statistical prediction using only binary data,

$$\begin{cases} 1 & \text{for} \quad m_i \geq c, \\ 0 & \text{for} \quad m_i < c, \end{cases}$$

for a mark m_i and a given threshold value $c > 0$. Thus, even if complete observation of marks is difficult, we can still predict the sample mean or the sample variance of marks by the threshold method.

There are several papers concerning the threshold method. Deneaud *et al.* (1984) gave the original idea of this method. For tropical rain rate data, Chiu (1988) pointed out that there is a high correlation between the area-average rain rate and the fractional area with the intensity higher or equal to a threshold value. This relation was confirmed for other regional rainfall data after that, see,

Received December 28, 2004. Revised May 16, 2005. Accepted July 14, 2005.

*Center for Advanced Research in Finance, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

**Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, O-okayama 2-12-1, Meguro-ku, Tokyo 152-8552, Japan.

e.g., the references of Shimizu *et al.* (1993) or Oki *et al.* (1997) which has shown this relation for radar-AMeDAS precipitation data in Japan. The empirical fact implies the availability of the threshold method.

Also, some theoretical studies on the threshold method have been done. Kedem and Pavlopoulos (1991) and Short *et al.* (1993) modeled a distribution of rainfall intensity as a mixture of a discrete part and a continuous part and discussed how to choose the optimal threshold value. Shimizu *et al.* (1993) proposed single- and double-threshold methods for the estimation of the area-average rain rate variance and gave a statistical explanation for selecting optimal threshold levels. They stated that some knowledge of the variance of rain rate could be useful from a meteorological view point because, for example, the knowledge would make it possible to distinguish between the convective and the stratiform conditions of rain.

All those studies do not take into account the spatial aspect of the rainfall. Mase (1996) modeled rainfall by a spatial point process on $\mathbb{Z}^2$ and studied the prediction of total rainfall. His main conclusion is that the threshold method would work well if a point process is non-ergodic and corresponding marks satisfy some mixing conditions. Sakaguchi and Mase (2003) considered the prediction of the sum of marks for a marked spatial point process on $\mathbb{R}^d$ and extended the results of Mase (1996).

In this paper, we consider the prediction of the sample variance of marks of Φ in almost the same statistical setup as in Sakaguchi and Mase (2003). The present study is motivated by Shimizu *et al.* (1993) mentioned above. We reformulate the problem of the variance estimation in the framework of a marked spatial point process on $\mathbb{R}^d$ and examine the prediction of the sample variance of marks by the threshold method. The convergence of the mean square error to 0 as an observational region expands to $\mathbb{R}^d$ will be shown. It is concluded that non-ergodicity of a point process and some mixing properties of corresponding marks are important conditions for accuracy of the threshold method.

In Section 2, basic facts on the general theory of marked spatial point processes and basic assumptions are introduced. In Section 3, we recall the results of Sakaguchi and Mase (2003) briefly and in Section 4, the prediction of the sample variance of marks is discussed. In Section 5, we confirm the theoretical results by a simulation study.

2. Preliminary

We explain a marked spatial point process model and its basic assumptions in this section. For the general theory, see Stoyan *et al.* (1995).

A marked spatial point process on $\mathbb{R}^d$ is a model of random point patterns on $\mathbb{R}^d$ and defined by a set

$$\Phi = \{(X_i, M_i)\}, \qquad (X_i, M_i) \in \mathbb{R}^d \times \mathbb{R},$$

where X_i is a random position and M_i is an associated random value called its mark. The non-marked point process Ψ corresponding to Φ is $\{X_i\}$, $X_i \in \mathbb{R}^d$, and $\Psi(A)$ means the number of points of Ψ in a Borel set $A \subset \mathbb{R}^d$.

In this paper, it is supposed that a marked spatial point process is a combination of Ψ and S as in Mase (1996). Here $S = \{S(x); x \in \mathbb{R}^d\}$ is a random field on $\mathbb{R}^d$ and Φ is given by

$$\Phi = \{(X_i, S(X_i))\}, \qquad X_i \in \mathbb{R}^d. \tag{2.1}$$

A non-marked point process Ψ is said stationary if $\Psi_h = \{X_i + h\}$ has the same distribution as Ψ for any $h \in \mathbb{R}^d$. Analogously, S is stationary if, for any $x_1, \ldots, x_n \in \mathbb{R}^d$ and any $h \in \mathbb{R}^d$, the distribution of $(S(x_1 + h), \ldots, S(x_n + h))$ is equal to that of $(S(x_1), \ldots, S(x_n))$. Let us assume that Ψ and S are independent and both stationary. When Ψ is stationary, there exists the constant λ called the intensity which is the mean number of points of Ψ per unit volume. We also suppose $\mathcal{M}([0, \infty)) = 1$ and Ψ is a fifth order simple point process, that is, $\boldsymbol{E}\Psi^5(A) < \infty$ for any bounded Borel set $A \subset \mathbb{R}^d$ and $X_i \neq X_j$ $(i \neq j)$.

The results of Mase (1996) and Sakaguchi and Mase (2003) indicate that non-ergodicity of Ψ and a mixing property of marks are key conditions for the accuracy of the threshold method for predicting the sum of marks. This suggests that the sample variance of marks will be also predicted well by the threshold method under those conditions.

Let C be the set of all configurations (i.e., locally finite subsets) of $\mathbb{R}^d$, $\mathcal{C}$ be its standard Borel σ-algebra and $\mathfrak{T}$ be the σ-algebra generated by translation invariant measurable sets in $\mathcal{C}$. We suppose Ψ is non-ergodic in the sense of Nguyen and Zessin (1979), that is,

$$\boldsymbol{E}(\Psi(\Delta_0) \mid \mathfrak{T}) \neq \lambda, \qquad \boldsymbol{P}\text{-a.s.}$$

is satisfied. Here Δ_0 is the unit cube including the origin.

Mixing conditions are defined by means of a mixing coefficient as in Bolthausen (1982). We partition $\mathbb{R}^d$ into congruent unit cubes

$$\Delta_i = \left\{ x = (x_1, \ldots, x_d) \in \mathbb{R}^d; i_l - \frac{1}{2} \leq x_l < i_l + \frac{1}{2}, l = 1, \ldots, d \right\}, \qquad i = (i_1, \ldots, i_d) \in \mathbb{Z}^d.$$

Let $d_1(\Delta_i, \Delta_j) = \max_{1 \leq k \leq d} |i_k - j_k|$, $d_2(I, J) = \min\{d_1(\Delta_i, \Delta_j) : i \in I, j \in J\}$ for $I, J \subset \mathbb{Z}^d$ and σ_I denote the σ-algebra generated by $\{S(x); x \in \cup_I \Delta_i\}$. For $k, l \in \mathbb{N}$, define the mixing coefficient as

$$\xi_{k,l}(n) = \sup\{|\boldsymbol{P}(A \cap B) - \boldsymbol{P}(A)\boldsymbol{P}(B)| : A \in \sigma_I, B \in \sigma_J, |I| \leq k, |J| \leq l, d_2(I, J) \geq n\}.$$

Here $|I|$ is the number of elements in I. $\boldsymbol{P}$ satisfies a mixing condition if $\xi_{k,l}(n) \to 0$ as $n \to \infty$. Later we need a more stringent assumption on the rate of convergence of $\xi_{k,l}(n)$.

3. Review of the preceding study

In the present section, we summarize the results of Sakaguchi and Mase (2003) for the model (2.1) although they discussed in the framework of general marked spatial point processes.

A threshold value $c \geq 0$ with $\mathcal{M}([c,\infty)]) > 0$ is fixed and Ψ is supposed to be second order in this section. For an observational region $G \subset \mathbb{R}^d$, define

$$B_G = \sum_{x\in\Psi} \mathbf{1}_G(x) b(S(x)),$$
$$F_G = \sum_{x\in\Psi} \mathbf{1}_G(x) f(S(x))$$

for non-negative measurable functions b and f. If we take $b(m) = \mathbf{1}_{[c,\infty)}(m)$ and $f(m) = m$, then B_G is the number of points in G with marks $\geq c$, and F_G is the total sum of marks in G.

Consider the prediction of $F_G/\sqrt{\mathrm{Var}\{F_G\}}$ by a simple linear predictor of $B_G/\sqrt{\mathrm{Var}\{B_G\}}$ of the form:

$$\alpha + \beta \frac{B_G}{\sqrt{\mathrm{Var}\{B_G\}}}. \tag{3.1}$$

Here α and β are constants. The minimum of the mean square error is given by

$$1 - \mathrm{Corr}\{B_G, F_G\}^2,$$

from elementary calculations. If this error is small enough, we can predict $F_G/\sqrt{\mathrm{Var}\{F_G\}}$ using (3.1) with great accuracy.

Sakaguchi and Mase (2003) calculated the theoretical moment of F_G, B_G and gave the following asymptotic behaviour under the mixing condition of a random field.

Proposition 1. *Assume the following conditions are satisfied*:

$$\sum_{n=1}^{\infty} n^{d-1} \xi_{1,1}^{\delta_1/(2+\delta_1)}(n) < \infty,$$
$$\left(\int b^{2+\delta_1}(m) d\mathcal{M}(m)\right)^{1/(2+\delta_1)} < \infty,$$
$$\left(\int f^{2+\delta_1}(m) d\mathcal{M}(m)\right)^{1/(2+\delta_1)} < \infty,$$

for some constant $\delta_1 > 0$. *Then, as* $G \uparrow \mathbb{R}^d$,

$$\mathrm{Cov}\{B_G, F_G\} = O(|G|) + \mu_b \mu_f \,\mathrm{Var}\{\Psi(G)\},$$
$$\mathrm{Var}\{B_G\} = O(|G|) + \mu_b^2 \,\mathrm{Var}\{\Psi(G)\},$$
$$\mathrm{Var}\{F_G\} = O(|G|) + \mu_f^2 \,\mathrm{Var}\{\Psi(G)\},$$

where $\mu_b = \int b d\mathcal{M}$ *and* $\mu_f = \int f d\mathcal{M}$.

The next proposition is derived immediately if we apply Proposition 1 to the correlation of B_G and F_G.

PROPOSITION 2. *Let us assume the conditions of Proposition* 1. *Then, we have*

$$\min_{\alpha,\beta} \boldsymbol{E} \left| \frac{F_G}{\sqrt{\mathrm{Var}\{F_G\}}} - \alpha - \beta \frac{B_G}{\sqrt{\mathrm{Var}\{B_G\}}} \right|^2 \to 0 \qquad as \quad G \uparrow \mathbb{R}^d,$$

if the condition

$$\frac{1}{|G|} \mathrm{Var}\{\Psi(G)\} \to \infty \qquad as \quad G \uparrow \mathbb{R}^d, \tag{3.2}$$

is satisfied. Moreover, when $\mathrm{Var}\{\Psi(G)\}$ *is of the order* $|G|^2$, *we can show*

$$\min_{\alpha,\beta} \boldsymbol{E} \left| \frac{F_G}{|G|} - \alpha - \beta \frac{B_G}{|G|} \right|^2 \to 0 \qquad as \quad G \uparrow \mathbb{R}^d,$$

similarly.

From Proposition 2, the threshold method will work fine if Ψ is non-ergodic and dependency of marks becomes weaker as the distance of points becomes larger. This result is an extension of Mase (1996) to a continuous space.

Although Sakaguchi and Mase (2003) did not discuss the coefficients of the best linear predictor, the following result is easily obtained.

PROPOSITION 3. *Suppose that* S *satisfies the conditions of Proposition* 1 *and let*

$$\alpha_G^* + \beta_G^* \frac{B_G}{|G|}, \tag{3.3}$$

be the best linear predictor of $F_G/|G|$. *Then,*

$$\beta_G^* \to \frac{\mu_f}{\mu_b}, \qquad \alpha_G^* \to 0 \qquad as \quad G \uparrow \mathbb{R}^d,$$

are shown provided that the condition (3.2) *and* $\mu_b \neq 0$ *are satisfied.*

PROOF. Elementary calculations give us relations

$$\alpha_G^* = \frac{1}{|G|} \boldsymbol{E} F_G - \beta_G^* \frac{1}{|G|} \boldsymbol{E} B_G, \tag{3.4}$$

$$\beta_G^* = \mathrm{Corr}\{B_G/|G|, F_G/|G|\} \sqrt{\frac{\mathrm{Var}\{F_G/|G|\}}{\mathrm{Var}\{B_G/|G|\}}}.$$

Also, $\boldsymbol{E} B_G = \lambda \mu_b |G|$ and $\boldsymbol{E} F_G = \lambda \mu_f |G|$. Therefore, the proof is straightforward from Proposition 1.

Note that when $b(m) = \mathbf{1}_{[c,\infty)}(m)$ and $f(m) = m$, we have $\beta_G^* \to \mu/\nu_c$ as $G \uparrow \mathbb{R}^d$, where $\mu = \int m d\mathcal{M}(m)$ and $\nu_c = \int_c^\infty d\mathcal{M}$.

4. Prediction of the sample variance

In this section, the prediction of the sample variance of marks by the threshold method is discussed. It is assumed that a sequence $\{G_n, n = 1, 2, \ldots\}$ of observational regions expands to $\mathbb{R}^d$ monotonically and there exists a sequence of regions $G'_n = \cup_{i \in I_{G_n}} \Delta_i$ with an index set I_{G_n} satisfying

$$G_n \subset G'_n, \qquad \sup_n \frac{|G'_n|}{|G_n|} < \infty.$$

Then, G_n is regular in terms of Nguyen and Zessin (1979) and we have

$$\frac{1}{|G_n|^2} \mathrm{Var}\{\Psi(G_n)\} \to \boldsymbol{E}|\boldsymbol{E}(\Psi(\Delta_0) \mid \mathfrak{T}) - \lambda|^2 \quad \text{as} \quad n \to \infty,$$

by Lemma 2 below. Therefore, the assumption that Ψ is non-ergodic is equivalent to supposing $\mathrm{Var}\{\Psi(G_n)\}$ is strictly of the order $|G_n|^2$ as $n \to \infty$.

Define the sample variance of marks as follows:

$$V_{G_n} = \frac{1}{|G_n|} F_{2G_n} - \left(\frac{1}{|G_n|} F_{G_n}\right)^2, \qquad F_{2G_n} = \sum_{x \in \Psi} \mathbf{1}_{G_n}(x) S^2(x),$$

where $f(m) = m$. Fix a threshold value $c \geq 0$ with $\mathcal{M}([c, \infty)) \neq 0$ as in the previous section.

If we consider the cases $f(m) = m^2$ and $f(m) = m$ in Proposition 2, we can see that F_{2G_n} and F_{G_n} will be predicted well by a linear function of $B_{G_n} = \sum_{x \in \Psi} \mathbf{1}_{G_n}(x) \mathbf{1}_{[c,\infty)}(S(x))$ respectively. Therefore, V_{G_n} would be predicted by a quadratic expression of B_{G_n},

$$\widehat{V}_{G_n} = \alpha + \beta \frac{B_{G_n}}{|G_n|} + \gamma \left(\frac{B_{G_n}}{|G_n|}\right)^2. \tag{4.1}$$

Shimizu *et al.* (1993) proposed the double-threshold method for the estimation of the area-average rain rate variance. We consider only the single-threshold case.

Our goal is to show that the minimum of the mean square error of (4.1) converges to 0 as G_n expands to $\mathbb{R}^d$. This implies that a quadratic predictor (4.1) will predict the sample variance with great accuracy if G_n is large enough.

First we introduce two lemmas for a non-marked point process Ψ. Lemma 1 treats the asymptotic order of the moments of $\Psi(G_n)$ and Lemma 2 gives the limit of $\boldsymbol{E}|\Psi(G_n) - \lambda|^m / |G_n|^m$.

Lemma 1. *Assume that Ψ is an m-th order point process. Then,*

$$\boldsymbol{E}\Psi^n(G_n) = O(|G_n|^n) \qquad (n = 1, \ldots, m),$$

holds.

PROOF. By Hölder's inequality and the stationarity of Ψ, we have, for some constant $k > 0$,

$$\begin{aligned}\boldsymbol{E}\Psi^n(G_n) &\le \sum_{i_1,\dots,i_n \in I_{G_n}} (\boldsymbol{E}\Psi^n(\Delta_{i_1}) \cdots \boldsymbol{E}\Psi^n(\Delta_{i_n}))^{1/n} \\ &\le k|G_n|^n \boldsymbol{E}\Psi^n(\Delta_0).\end{aligned}$$

Since Ψ is an m-th order point process from the assumption, $\boldsymbol{E}\Psi^n(\Delta_0) < \infty$ follows and the assertion is proved.

LEMMA 2. *Let us suppose Ψ is of $(m+1)$-th order. Then we have*

$$\frac{1}{|G_n|^m}\boldsymbol{E}Z_{G_n}^m \to \boldsymbol{E}|\boldsymbol{E}(\Psi(\Delta_0) \mid \mathfrak{T}) - \lambda|^m \quad \text{as} \quad n \to \infty,$$

where $Z_{G_n} = \Psi(G_n) - \lambda|G_n|$.

PROOF. We shall show that $Z_{G_n}^m/|G_n|^m$ is uniformly integrable. Then we have

$$\lim_{n\to\infty} \boldsymbol{E}\left|\frac{1}{|G_n|}\Psi(G_n) - \lambda\right|^m = \boldsymbol{E}\lim_{n\to\infty}\left|\frac{1}{|G_n|}\Psi(G_n) - \lambda\right|^m. \tag{4.2}$$

$Z_{G_n}^m/|G_n|^m$ is uniformly integrable if $\boldsymbol{E}(Z_{G_n}^m/|G_n|^m)^p$ is finite for some $p > 1$. We consider the case $p = (m+1)/m$. There exists a constant $M > 0$ such that

$$\boldsymbol{E}\left|\frac{1}{|G_n|}\Psi(G_n) - \lambda\right|^{m+1} < M,$$

from Lemma 1 and the assumption that Ψ is of $(m+1)$-th order. Therefore the equation (4.2) is derived.

By the spatial ergodic theorem due to Nguyen and Zessin (1979), the right-hand side of (4.2) can be written as $\boldsymbol{E}|\boldsymbol{E}(\Psi(\Delta_0) \mid \mathfrak{T}) - \lambda|^m$. Hence, the assertion follows.

Now we can show the following main result. Proposition 4 asserts that the threshold method for predicting the sample variance of marks works fine if G_n is large enough provided that marks satisfy the mixing condition.

PROPOSITION 4. *Under the assumptions*

$$\sum_{n=1}^{\infty} n^{d-1}\xi_{1,3}^{\delta_2/(4+\delta_2)}(n) < \infty, \qquad \text{and}$$

$$\left(\int s^{4+2\delta_2} d\mathcal{M}(s)\right)^{1/(4+2\delta_2)} < \infty,$$

for some constant $\delta_2 > 0$, we have

$$\min_{\alpha,\beta,\gamma} \boldsymbol{E}\left|V_{G_n} - \widehat{V}_{G_n}\right|^2 \to 0 \quad \text{as} \quad n \to \infty.$$

Proof. Let

$$\alpha^{**}_{G_n} + \beta^{**}_{G_n} \frac{B_{G_n}}{|G_n|}, \tag{4.3}$$

be the best linear predictor of $F_{2G_n}/|G_n|$ and

$$\begin{aligned} A^{*}_{G_n} &= \frac{1}{|G_n|} F_{G_n} - \alpha^{*}_{G_n} - \beta^{*}_{G_n} \frac{1}{|G_n|} B_{G_n}, \\ A^{**}_{G_n} &= \frac{1}{|G_n|} F_{2G_n} - \alpha^{**}_{G_n} - \beta^{**}_{G_n} \frac{1}{|G_n|} B_{G_n}, \end{aligned}$$

be the corresponding prediction errors of the predictors (3.3) and (4.3) respectively. The inequality

$$\min_{\alpha,\beta,\gamma} \boldsymbol{E}|V_{G_n} - \widehat{V}_{G_n}|^2 \le \boldsymbol{E}\left| A^{**}_{G_n} - {A^{*}_{G_n}}^2 - 2\alpha^{*}_{G_n} A^{*}_{G_n} - 2\beta^{*}_{G_n} \frac{B_{G_n}}{|G_n|} A^{*}_{G_n} \right|^2, \tag{4.4}$$

holds since the right-hand side of (4.4) is the mean square error of the predictor

$$(\alpha^{**}_{G_n} - {\alpha^{*}_{G_n}}^2) + (\beta^{**}_{G_n} - 2\alpha^{*}_{G_n}\beta^{*}_{G_n}) \frac{B_{G_n}}{|G_n|} + (-{\beta^{*}_{G_n}}^2) \left(\frac{B_{G_n}}{|G_n|} \right)^2.$$

By Minkowski's inequality, we have

$$\begin{aligned} &\left(\boldsymbol{E}\left| A^{**}_{G_n} - {A^{*}_{G_n}}^2 - 2\alpha^{*}_{G_n} A^{*}_{G_n} - 2\beta^{*}_{G_n} \frac{B_{G_n}}{|G_n|} A^{*}_{G_n} \right|^2 \right)^{1/2} \\ &\quad \le (\boldsymbol{E}{A^{**}_{G_n}}^2)^{1/2} + (\boldsymbol{E}{A^{*}_{G_n}}^4)^{1/2} + (4{\alpha^{*}_{G_n}}^2 \boldsymbol{E}{A^{*}_{G_n}}^2)^{1/2} \\ &\qquad + \left(4{\beta^{*}_{G_n}}^2 \frac{1}{|G_n|^2} \boldsymbol{E} B^2_{G_n} {A^{*}_{G_n}}^2 \right)^{1/2}. \end{aligned}$$

Furthermore, this is bounded from above by

$$(\boldsymbol{E}{A^{*}_{G_n}}^4)^{1/2} + 2\beta^{*}_{G_n} \left(\left(\frac{1}{|G_n|^4} \boldsymbol{E} B^4_{G_n} \right)^{1/2} (\boldsymbol{E}{A^{*}_{G_n}}^4)^{1/2} \right)^{1/2} + o(1),$$

from Hölder's inequality and the result of the previous section. Since $\boldsymbol{E}B^4_{G_n} \le \boldsymbol{E}\Psi(G_n)^4$, we can see $\boldsymbol{E}B^4_{G_n} = O(|G_n|^4)$ by Lemma 1. Therefore, it is enough to show $\boldsymbol{E}{A^{*}_{G_n}}^4 \to 0$ as $n \to \infty$ in order to prove the assertion.

Let $S_{G_n} = F_{G_n} - \boldsymbol{E}F_{G_n}$ and $T_{G_n} = B_{G_n} - \boldsymbol{E}B_{G_n}$. Then the equality

$$\begin{aligned} \boldsymbol{E}{A^{*}_{G_n}}^4 = \frac{1}{|G_n|^4} (&\boldsymbol{E}S^4_{G_n} - 4\beta^{*}_{G_n} \boldsymbol{E}S^3_{G_n} T_{G_n} + 6{\beta^{*}_{G_n}}^2 \boldsymbol{E}S^2_{G_n} T^2_{G_n} \\ &- 4{\beta^{*}_{G_n}}^3 \boldsymbol{E}S_{G_n} T^3_{G_n} + {\beta^{*}_{G_n}}^4 \boldsymbol{E}T^4_{G_n}), \end{aligned} \tag{4.5}$$

holds from (3.4). The first and the second term in the parenthesis of the right-hand side of (4.5) can be written as

$$\begin{aligned}\boldsymbol{E}S_{G_n}^4 &= \boldsymbol{E}(X_{G_n}+\mu Z_{G_n})^4\\ &= \boldsymbol{E}X_{G_n}^4 + 4\mu\boldsymbol{E}X_{G_n}^3 Z_{G_n} + 6\mu^2\boldsymbol{E}X_{G_n}^2 Z_{G_n}^2 + 4\mu^3\boldsymbol{E}X_{G_n}Z_{G_n}^3 + \mu^4\boldsymbol{E}Z_{G_n}^4,\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{E}S_{G_n}^3 T_{G_n} &= \boldsymbol{E}(X_{G_n}+\mu Z_{G_n})^3(Y_{G_n}+\nu_c Z_{G_n})\\ &= \boldsymbol{E}X_{G_n}^3 Y_{G_n} + 3\mu\boldsymbol{E}X_{G_n}^2 Y_{G_n}Z_{G_n} + 3\mu^2\boldsymbol{E}X_{G_n}Y_{G_n}Z_{G_n}^2 + \mu^3\boldsymbol{E}Y_{G_n}Z_{G_n}^3\\ &\quad + \nu_c\boldsymbol{E}X_{G_n}^3 Z_{G_n} + 3\mu\nu_c\boldsymbol{E}X_{G_n}^2 Z_{G_n}^2 + 3\mu^2\nu_c\boldsymbol{E}X_{G_n}Z_{G_n}^3 + \mu^3\nu_c\boldsymbol{E}Z_{G_n}^4,\end{aligned}$$

respectively, where

$$X_{G_n} = \sum_{x\in\Psi}\mathbf{1}_{G_n}(x)(S(x)-\mu), \qquad Y_{G_n} = \sum_{x\in\Psi}\mathbf{1}_{G_n}(x)(\mathbf{1}_{[c,\infty)}(S(x))-\nu_c).$$

By Hölder's inequality, Lemma 2 and Lemma 3 below, we have

$$\begin{aligned}\frac{1}{|G_n|^4}\boldsymbol{E}S_{G_n}^4 &= \frac{1}{|G_n|^4}(\mu^4\boldsymbol{E}Z_{G_n}^4 + o(|G_n|^4)) \to \mu^4\sigma^*,\\ \frac{1}{|G_n|^4}\boldsymbol{E}S_{G_n}^3 T_{G_n} &= \frac{1}{|G_n|^4}(\mu^3\nu_c\boldsymbol{E}Z_{G_n}^4 + o(|G_n|^4)) \to \mu^3\nu_c\sigma^* \quad \text{as} \quad n\to\infty.\end{aligned}$$

Here $\sigma^* = \boldsymbol{E}|\boldsymbol{E}(\Psi(\Delta_0)\mid\mathfrak{T})-\lambda|^4$. Similarly,

$$\frac{1}{|G_n|^4}\boldsymbol{E}S_{G_n}T_{G_n}^3 \to \mu\nu_c^3\sigma^*, \qquad \frac{1}{|G_n|^4}\boldsymbol{E}T_{G_n}^4 \to \nu_c^4\sigma^* \quad \text{as} \quad n\to\infty,$$

can be shown. Thus, we obtain

$$\begin{aligned}\boldsymbol{E}A_{G_n}^{*\ 4} \le \frac{1}{|G_n|^4}\{&\boldsymbol{E}S_{G_n}^4 - 4\beta_{G_n}^*\boldsymbol{E}S_{G_n}^3T_{G_n} + 6\beta_{G_n}^{*\ 2}(\boldsymbol{E}S_{G_n}^4)^{1/2}(\boldsymbol{E}T_{G_n}^4)^{1/2}\\ &- 4\beta_{G_n}^{*\ 3}\boldsymbol{E}S_{G_n}T_{G_n}^3 + \beta_{G_n}^{*\ 4}\boldsymbol{E}T_{G_n}^4\}\\ \to \mu^4\sigma^* - 4\mu^4\sigma^* + 6\mu^4\sigma^* - 4\mu^4\sigma^* + \mu^4\sigma^* = 0 &\quad \text{as} \quad n\to\infty,\end{aligned}$$

from Proposition 3 and the assertion follows.

LEMMA 3. *Under the same assumptions as in Proposition* 4, *we have*

$$\boldsymbol{E}X_{G_n}^4 = O(|G_n|^3), \qquad \boldsymbol{E}Y_{G_n}^4 = O(|G_n|^3).$$

PROOF. Let $R(x) = S(x)-\mu$ and $\mu^{(4)}$ denotes the fourth order moment measure:

$$\mu^{(4)}(B_1\times B_2\times B_3\times B_4) = \boldsymbol{E}\Psi(B_1)\Psi(B_2)\Psi(B_3)\Psi(B_4),$$

for Borel sets $B_1, B_2, B_3, B_4 \subset \mathbb{R}^d$, see Stoyan *et al.* (1995) for details. Since Ψ and S are independent each other,

$$\begin{aligned} \boldsymbol{E} X_{G_n}^4 &= \boldsymbol{E} \sum_{(x)_4 \in \Psi} \mathbf{1}_{G_n \times \cdots \times G_n}((x)_4) R(x_1) \cdots R(x_4) \\ &= \int_{G_n \times \cdots \times G_n} \boldsymbol{E} R(x_1) \cdots R(x_4) d\mu^{(4)}((x)_4), \end{aligned} \tag{4.6}$$

where $(x)_4$ is an abbreviation for (x_1, x_2, x_3, x_4). The right-hand side of (4.6) is bounded above by

$$\begin{aligned} \sum_{i,j,k,l \in I_{G_n}} \int_{\Delta_i \times \Delta_j \times \Delta_k \times \Delta_l} & [\mathrm{Cum}\{S(x_1), \ldots, S(x_4)\} \\ & + \boldsymbol{E} R(x_1) R(x_2) \boldsymbol{E} R(x_3) R(x_4) \\ & + \boldsymbol{E} R(x_1) R(x_3) \boldsymbol{E} R(x_2) R(x_4) \\ & + \boldsymbol{E} R(x_1) R(x_4) \boldsymbol{E} R(x_2) R(x_3)] d\mu^{(4)}((x)_4). \end{aligned} \tag{4.7}$$

Here $\mathrm{Cum}\{S(x_1), \ldots, S(x_4)\}$ is the fourth order joint cumulant of S:

$$\begin{aligned} \mathrm{Cum}\{S(x_1), \cdots, S(x_4)\} = & \boldsymbol{E} R(x_1) R(x_2) R(x_3) R(x_4) \\ & - \boldsymbol{E} R(x_1) R(x_2) \boldsymbol{E} R(x_3) R(x_4) \\ & - \boldsymbol{E} R(x_1) R(x_3) \boldsymbol{E} R(x_2) R(x_4) \\ & - \boldsymbol{E} R(x_1) R(x_4) \boldsymbol{E} R(x_2) R(x_3). \end{aligned}$$

From Theorem 1 and Theorem 2 of Mase (1982), the fourth order joint cumulants are bounded as

$$|\mathrm{Cum}\{S(x_1), \ldots, S(x_4)\}| \le c \xi_{1,3}^{\delta_2/4+\delta_2}(d_2(\{i\}, \{j,k,l\})),$$

for $x_1 \in \Delta_i$, $x_2 \in \Delta_j$, $x_3 \in \Delta_k$, $x_4 \in \Delta_l$ and some constant $c > 0$. Also, there exists a constant $M > 0$ such that $\mu^{(4)}(\Delta_i \times \Delta_j \times \Delta_k \times \Delta_l) < M$ by the assumption that Ψ is a fifth order point process and Lemma 1. Thus, we have

$$\begin{aligned} & \sum_{i,j,k,l \in I_{G_n}} \int_{\Delta_i \times \Delta_j \times \Delta_k \times \Delta_l} \mathrm{Cum}\{S(x_1), \ldots, S(x_4)\} d\mu^{(4)} \\ & \quad \le c_1 \sum_{i,j,k,l \in I_{G_n}} \xi_{1,3}^{\delta_2/4+\delta_2}(d_2(\{i\}, \{j,k,l\})) \\ & \quad \le c_2 \left(\sum_{m=0}^{\infty} m^{d-1} \xi_{1,3}^{\delta_2/4+\delta_2}(m) \right) |G_n|^3 \\ & \quad = O(|G_n|^3) \end{aligned}$$

for some constants $c_1, c_2 > 0$.

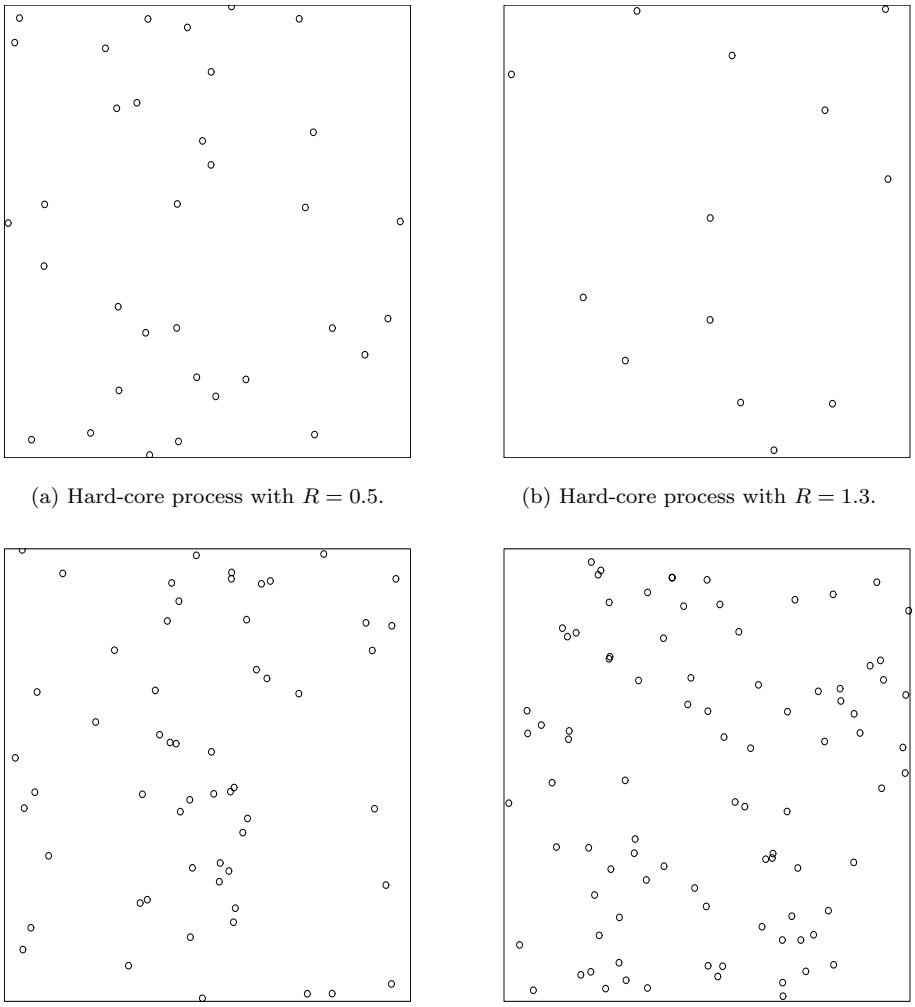
Since the second order cumulants can be also bounded as

$$|\operatorname{Cov}\{S(x_1), S(x_2)\}| \le c' \xi_{1,1}^{\delta_2/2+\delta_2}(d_1(\Delta_i, \Delta_j)),$$

for $x_1 \in \Delta_i$, $x_2 \in \Delta_j$ and some constant $c' > 0$ from Lemma 1 of Bolthausen (1982). Hence, the inequality

$$\begin{aligned}
&\sum_{i,j,k,l \in I_{G_n}} \int_{\Delta_i \times \Delta_j \times \Delta_k \times \Delta_l} \boldsymbol{E}R(x_1)R(x_2)\boldsymbol{E}R(x_3)R(x_4)d\mu^{(4)}((x)_4) \\
&\quad \leq c_2' \left(\sum_{i,j \in I_{G_n}} \xi_{1,1}^{\delta_2/2+\delta_2}(d_1(\Delta_i, \Delta_j)) \right) \left(\sum_{k,l \in I_{G_n}} \xi_{1,1}^{\delta_2/2+\delta_2}(d_1(\Delta_k, \Delta_l)) \right) \\
&\quad = O(|G_n|^2),
\end{aligned}$$

holds for some $c_2' > 0$. Therefore, the expression (4.7) is of order $|G_n|^3$ and $\boldsymbol{E}X_{G_n}^4 = O(|G_n|^3)$ follows.



(a) Hard-core process with $R = 0.5$.

(b) Hard-core process with $R = 1.3$.

(c) Poisson process with $\lambda = 0.5$.

(d) Poisson process with $\lambda = 1$.

Figure 1. A sample of realizations of four point processes in a 10×10 square region.

$\boldsymbol{E}Y_{G_n}^4 = O(|G_n|^3)$ can be shown in the same way. Hence, the assertion is proved.

5. Simulation study

In this section, we simulate non-ergodic marked spatial point processes on $\mathbb{R}^2$ and confirm the previous theoretical result. Let Ψ be the mixture of four point processes and $\boldsymbol{P}$ be its distribution:

$$\boldsymbol{P} = 0.1\boldsymbol{P}_1 + 0.4\boldsymbol{P}_2 + 0.35\boldsymbol{P}_3 + 0.15\boldsymbol{P}_4,$$

where each distribution $\boldsymbol{P}_i$ is

$\boldsymbol{P}_1$: Pure hard-core process with hard-core distance $R = 0.5$,
$\boldsymbol{P}_2$: Pure hard-core process with hard-core distance $R = 1.3$,
$\boldsymbol{P}_3$: Poisson process with intensity $\lambda = 0.5$,
$\boldsymbol{P}_4$: Poisson process with intensity $\lambda = 1$.

A pure hard-core process is a model of a random point pattern generated by hard spheres with radius $R/2$. A Poisson process is a basic model and can be considered as a pure hard-core process with $R = 0$. Thus, the process Ψ models a phenomenon that the degree of interactions between points varies. A realization of each point process appears in Figure 1.

Marks are given by simulating a random field $S = \{S(x) = \exp(T(x)); x \in \mathbb{R}^2\}$ where T is a stationary and isotropic Gaussian random field. We let $\boldsymbol{E}T(0) =$

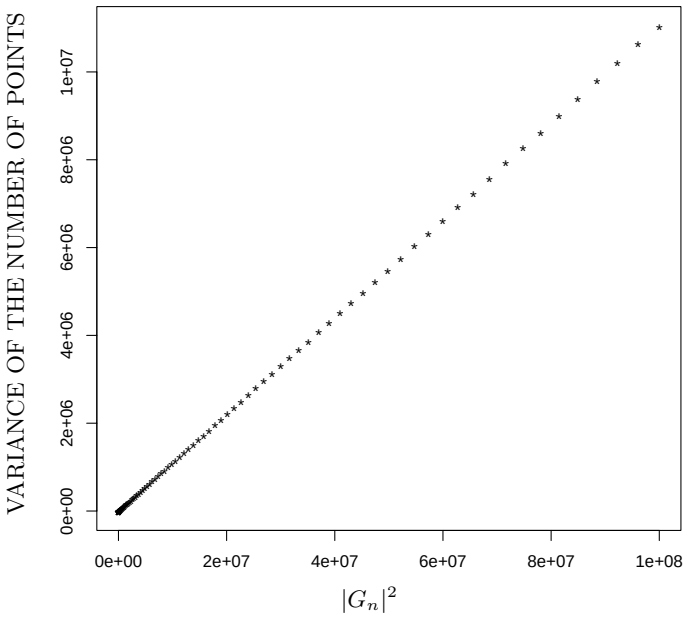


Figure 2. $\mathrm{Var}\{\Psi(G_n)\}$ for the mixture process versus $|G_n|^2$.

0 and the covariance $\mathrm{Cov}\{T(x_1), T(x_2)\} = 1 - \gamma(x_1 - x_2)$ for $x_1, x_2 \in \mathbb{R}^2$. Here, γ is an exponential semivariogram $\gamma(h) = 1 - \exp(-|h|)$. For reference about semivariograms and their corresponding Gaussian random fields, see, e.g., Cressie (1993). In total, 100 marked spatial point processes are generated in a rectangle window.

Figure 2 is the plot of $\mathrm{Var}\{\Psi(G_n)\}$ versus $|G_n|^2$ showing $\mathrm{Var}\{\Psi(G_n)\} = O(|G_n|^2)$. Therefore, the threshold method should work fine.

Table 1 shows the coefficients of determination R^2 when the size of G_n and a threshold value vary. We can see that the threshold method gives higher R^2

Table 1. Coefficient of determination R^2 values for the mixture process.

G_n	$c = 0$	$c = 10$	$c = 20$	$c = 30$	$c = 40$
10×10	0.059	0.149	0.233	0.243	0.297
20×20	0.450	0.523	0.617	0.637	0.453
30×30	0.583	0.619	0.700	0.711	0.609
40×40	0.789	0.812	0.838	0.846	0.716
50×50	0.864	0.890	0.899	0.903	0.831
60×60	0.901	0.917	0.932	0.944	0.888
70×70	0.925	0.933	0.944	0.952	0.930
80×80	0.922	0.935	0.951	0.960	0.931
90×90	0.934	0.947	0.962	0.964	0.958
100×100	0.949	0.959	0.969	0.972	0.962

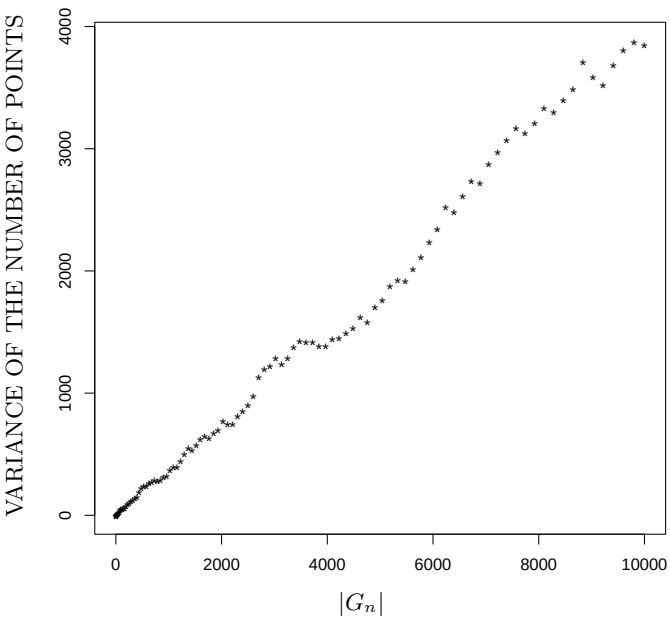


Figure 3. $\mathrm{Var}\{\Psi(G_n)\}$ for the Poisson process versus $|G_n|$.

Table 2. Coefficient of determination R^2 values for the Poisson process.

G_n	$c = 0$	$c = 10$	$c = 20$	$c = 30$	$c = 40$
10×10	0.028	0.436	0.644	0.701	0.550
20×20	0.016	0.145	0.349	0.487	0.584
30×30	0.037	0.221	0.424	0.568	0.544
40×40	0.037	0.325	0.474	0.590	0.533
50×50	0.017	0.195	0.302	0.457	0.446
60×60	0.014	0.233	0.354	0.517	0.473
70×70	0.006	0.234	0.328	0.532	0.522
80×80	0.008	0.215	0.330	0.611	0.569
90×90	0.003	0.334	0.315	0.444	0.467
100×100	0.009	0.381	0.356	0.435	0.434

values as G_n becomes larger for each threshold value. The highest R^2 value is given around $c = 26$ for various sizes of G_n in our simulation study.

For comparison, let us consider an ergodic case that Ψ is a Poisson process with intensity 0.5. Figure 3 is the plot of $\mathrm{Var}\{\Psi(G_n)\}$ versus $|G_n|$. Clearly, $\mathrm{Var}\{\Psi(G_n)\} = O(|G_n|)$.

Table 2 lists R^2 values. Note that R^2 value need not become larger as G_n expands.

6. Concluding remarks

In this paper, we have considered the threshold method for predicting the sample variance of marks of $\Phi = \{(X_i, S(X_i))\}$, $X_i \in \mathbb{R}^d$. It is shown that the mean square error will converge to 0 as an observational region expands to $\mathbb{R}^d$ if Ψ is non-ergodic and S satisfies a mixing-type condition.

For a general marked spatial point process model $\Phi = \{(X_i, M_i)\}$, $(X_i, M_i) \in \mathbb{R}^d \times \mathbb{R}$, it may be possible to get almost the same results as those in this paper. However, more complex mathematical formulations will be needed. Also, the prediction of higher dimensional moments by the threshold method would be considered, but we do not treat it in this paper.

Acknowledgements

The authors would like to thank the referees for their comments and suggestions which helped to improve the paper.

References

Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields, *Ann. Probab.*, **10**, 1047–1050.

Chiu, L. S. (1988). Rain estimation from satellites: Areal rainfall-rain area relations, Third Conference on Satellite Meteorology and Oceanography, American Meteorological Society, Anaheim.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*, Wiley, New York.

Deneaud, A. A., Niscov, S. I., Priegnitz, D. L. and Smith, P. L. (1984). The area-time integral as an indicator for convective rain volumes, *Journal of Climate and Applied Meteorology*, **23**, 555–561.

Kedem, B. and Pavlopoulos, H. (1991). On the threshold method for rainfall estimation: Choosing the optimal threshold level, *J. Amer. Statist. Assoc.*, **86**, 626–633.

Mase, S. (1982). Properties of fourth-order strong mixing rates and its application to random set theory, *J. Multivariate Anal.*, **12**, 549–561.

Mase, S. (1996). The threshold method for estimating total rainfall, *Ann. Inst. Statist. Math.*, **48**, 201–213.

Nguyen, X. X. and Zessin, H. (1979). Ergodic theorems for spatial processes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **48**, 133–158.

Oki, R., Sumi, A. and Short, D. A. (1997). TRMM sampling of Radar-AMeDAS Rainfall using the threshold method, *Journal of Applied Meteorology*, **36**, 1480–1492.

Sakaguchi, T. and Mase, S. (2003). On the threshold method for marked spatial point processes, *J. Japan Statist. Soc.*, **33**, 23–37.

Shimizu, K., Short, D. A. and Kedem, B. (1993). Single- and double-threshold methods for estimating the variance of area rain rate, *Journal of the Meteorological Society of Japan*, **71**, 673–683.

Short, D. A., Shimizu, K. and Kedem, B. (1993). Optimal thresholds for the estimation of area rain-rate moments by the threshold method, *Journal of Applied Meteorology*, **32**, 182–192.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and its Applications*, Second Edition, Wiley, New York.