J. Japan Statist. Soc.
Vol. 33 No. 1 2003 23–37

ON THE THRESHOLD METHOD FOR MARKED SPATIAL POINT PROCESSES

Takayuki Sakaguchi* and Shigeru Mase*

The threshold method in the framework of marked spatial point processes on a continuous space is discussed. The threshold method is a linear prediction of the total sum of marks using only the number of points with marks exceeding a given threshold value. The result is an extension of Mase (1996) to a continuous space and also the independent mark assumption of Mase (1996) is weakened. It is shown that the total sum of the marks is linearly predictable if the number of points has a huge variation and marks satisfy some mixing condition. A simulation study is given to illustrate the theoretical result.

Key words and phrases: Linear prediction, mixing condition, non-ergodicity.

1. Introduction

Let $\Phi = \{(X_i, M_i);\ i = 1, 2, \dots\}$ be a marked spatial point process on $\mathbb{R}^d \times \mathbb{R}$, where X_i represents an observational position and M_i an associated mark. Sometimes it is difficult, inaccurate, or impossible to observe marks which are relatively small. The threshold method tries to predict the total sum of marks in a region $G \subset \mathbb{R}^d$ using only the number of positions with marks exceeding a given threshold value.

The idea of the threshold method was given by Deneaud *et al.* (1984), see also Shimizu (2002). Chiu (1988) remarked the fact that for tropical rain rate data, there is a surprisingly high correlation between the area average rain rate and the fractional area where rain rates exceed a certain threshold value. This phenomenon has been confirmed subsequently in various rainfall data and there are several papers on how to choose an optimal threshold value, see, e.g., Kedem and Pavlopoulos (1991), Shimizu *et al.* (1993) and Short *et al.* (1993). These authors assumed the rainfall distribution to be a mixture of a discrete distribution and a positive continuous distribution. Also Mase (1996) discussed why the threshold method works using a spatial model. Its main conclusion is that the threshold method works fine if a variation of the number of raining sites is dominant.

As is easily conceivable, the threshold method is useful in many spatial problems. However, in general spatial problems, observational locations are not discretized and are both continuously and randomly located. For example, consider animals which move in groups in a region. Small herds are difficult to observe, while larger ones are less difficult to observe. Moreover, even if observed, precise

Received November 15, 2002. Revised March 7, 2003. Accepted March 31, 2003.

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

counting of sizes of herds may be difficult too. In this case, the threshold method amounts to predict the total population of animals in a region using only the number of herds with sizes exceeding a certain threshold value.

In this paper, we study a theoretical basis of the threshold method for marked spatial point processes which generalizes the result in Mase (1996) to continuous spaces. In Mase (1996), discrete spatial marked point processes are assumed to be combinations of non-marked point processes and random fields that are assumed to be mutually independent. Although every marked point processes, continuous or not, can be represented as such a combination, the independence assumption of point processes and random fields may be sometimes dubious and many examples of position-dependent marks such as nearest neighbor distances are used in spatial statistics. The present result also includes this kind of position-dependent mark cases partially.

Let F_G be the total sum of marks M_i with $X_i \in G$ and B_G be the number of points in G with marks greater than or equal to a certain threshold value. We consider a linear predictor $\hat{F}_G = \alpha + \beta B_G$ and give a condition under which the prediction error is small. In Section 2, we review the general theory of marked spatial point processes. In Section 3, formulas for moments of B_G and F_G are derived. In Section 4, asymptotic behaviors of these moments as G expands are given under some mixing type condition. In Section 5, we show our main result that the threshold method can predict F_G well if the variance of $\Psi(G)$, the number of points in G, diverges faster than the volume of G. Finally, we illustrate the main result by a simulation study in Section 6.

2. Marked spatial point processes

In this section we summarize basic results from the theory of marked spatial point processes, see, e.g., Stoyan *et al.* (1995). A marked spatial point process on $\mathbb{R}^d$ with the mark space $\mathbb{R}$ is a collection of pairs

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

X_i is a random position and its mark M_i is an associated random quantity. Φ is simple if $X_i \neq X_j$ $(i \neq j)$. The associated non-marked point process $\Psi = \{X_i\}$ is of n-th order if $\boldsymbol{E}\Psi(A)^n < \infty$ for any bounded Borel set $A \subset \mathbb{R}^d$. Φ is (strongly) stationary if $\Phi_h = \{(X_i + h, M_i)\}$ has the same distribution as Φ for any $h \in \mathbb{R}^d$. We will assume Φ is stationary and Ψ is of 2nd order in the following.

The intensity measure Λ gives the mean number of marked points $\Lambda(A \times L) = \boldsymbol{E}\Phi(A \times L)$ for measurable sets $A \times L \subset \mathbb{R}^d \times \mathbb{R}$. Since Φ is stationary, there exists the intensity λ, the mean number of points X_i per unit volume. The Campbell formula is basic in the theory of point processes and states:

$$\boldsymbol{E} \sum_{(x,m)\in\Phi} f(x,m) = \int f(x,m) d\Lambda(x,m) = \lambda \iint f(x,m) d\mathcal{M}(m) dx \tag{2.1}$$

for arbitrary measurable function $f \geq 0$ on $\mathbb{R}^d \times \mathbb{R}$. The probability measure $\mathcal{M}$ on $\mathbb{R}$ is called the mark distribution. In the following, we assume that $\mathcal{M}$ is supported on $[0, \infty)$ (i.e., marks are non-negative) and is not degenerated to 0.

The Palm distribution $\boldsymbol{P}_{(x,m)}$ is the distribution of Φ under the condition $(x,m) \in \Phi$. From the stationarity, $\boldsymbol{P}_{(x,m)}$ is the translation of $\boldsymbol{P}_{(0,m)}$ by the vector x. The two-point Palm distribution $\boldsymbol{P}_h^{m,l}$ is the distribution of Φ under the condition $(0,m),(h,l) \in \Phi$. Let $\mathbb{C}$ be the set of all configurations (i.e., locally finite subsets) of $\mathbb{R}^d$ and $\mathcal{C}$ be its standard Borel σ-algebra. The Palm distribution $\boldsymbol{P}_0$ of Ψ is of the form

$$\boldsymbol{P}_0(A) = \boldsymbol{P}(\Psi \in A \mid 0 \in \Psi)$$

for any measurable $A \in \mathcal{C}$ and satisfies

$$\int f(\psi) d\boldsymbol{P}_0(\psi) = \frac{1}{\lambda|B|} \boldsymbol{E} \sum_{y \in \Psi \cap B} f(\Psi_{-y}) \tag{2.2}$$

for any $\mathcal{C}$-measurable function $f \geq 0$. Here B is any Borel set of positive Lebesgue measure $|B|$. The reduced Palm distribution $\boldsymbol{P}_0^!$ of Ψ is the conditional distribution

$$\boldsymbol{P}_0^!(A) = \boldsymbol{P}(\Psi \setminus \{0\} \in A \mid 0 \in \Psi).$$

The second-order reduced moment measure $\mathcal{K}$ is defined by

$$\mathcal{K}(A) = \frac{1}{\lambda} \int \psi(A) d\boldsymbol{P}_0^!(\psi)$$

for any bounded Borel set $A \subset \mathbb{R}^d$. The quantity $\lambda\mathcal{K}(A)$ is the mean number of points in $A \setminus \{0\}$ under the condition $0 \in \Psi$.

The second-order factorial moment measure $\Lambda^{(2)}$ gives the mean number of different pairs of marked points in Φ,

$$\Lambda^{(2)}(A_1 \times L_1 \times A_2 \times L_2) = \boldsymbol{E} \sum_{(x,m),(y,n) \in \Phi}^{\neq} \mathbf{1}_{A_1 \times L_1}(x,m) \mathbf{1}_{A_2 \times L_2}(y,n) \tag{2.3}$$

for measurable sets $A_1,\ A_2 \subset \mathbb{R}^d$ and $L_1,\ L_2 \subset \mathbb{R}$, where the summation symbol with $\neq$ means that the sum is taken for all different pairs. Using these measures second-order moments can be expressed as follows:

$$\begin{aligned} &\boldsymbol{E} \int g(x,m) d\Phi(x,m) \int h(x,m) d\Phi(x,m) \\ &\quad = \int g(x,m) h(x,m) d\Lambda(x,m) + \int g(x,m) h(y,l) d\Lambda^{(2)}(x,m,y,l), \end{aligned} \tag{2.4}$$

where g and h are non-negative measurable functions on $\mathbb{R}^d \times \mathbb{R}$.

Finally there is a following relation, the two-point refined Campbell theorem, among these measures:

$$\begin{aligned} &\boldsymbol{E} \sum_{(x,m),(y,l) \in \Phi}^{\neq} f(x,m,y,l,\Phi) \\ &\quad = \lambda^2 \int \cdots \int f(x,m,x+h,l,\phi_x) d\boldsymbol{P}_h^{m,l}(\phi) d\mathcal{M}_h(m,l) d\mathcal{K}(h) dx \end{aligned} \tag{2.5}$$

for arbitrary measurable function $f \geq 0$ on $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{C}$. Here $\mathcal{M}_h(m,l)$ is the two-point mark distribution, that is, the joint distribution of marks m and l under the condition that Φ has points at 0 and h.

3. Moment formulas

Let Φ be a simple and stationary marked point process on $\mathbb{R}^d$ and Ψ be an associated 2nd order non-marked point process. For a region $G \subset \mathbb{R}^d$, define

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

for non-negative measurable functions b and f. Let $c \geq 0$ and consider the case $b(m) = \mathbf{1}_{[c,\infty)}(m)$ 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. The constant c is called a threshold value in this case. We will consider a simple linear predictor of F_G of the form:

$$\widehat{F}_G = \alpha + \beta B_G. \tag{3.1}$$

Here α and β are real constants. It can be shown that the minimum of a prediction error $\boldsymbol{E}|F_G - \alpha - \beta B_G|^2$ over $\mathbb{R}^2$ is given by

$$\mathrm{Var}\,\{F_G\}\left\{1 - \mathrm{Corr}\,\{B_G, F_G\}^2\right\}. \tag{3.2}$$

If this error is small enough, we can predict the total sum of marks using (3.1) with considerable accuracy. Therefore, the question we have to ask is under what condition this error will be small.

Let $\mu_b = \int b(m) d\mathcal{M}(m)$ and $\mu_f = \int f(m) d\mathcal{M}(m)$. Since Φ is stationary, Campbell theorem (2.1) yields relations

$$\boldsymbol{E}\{B_G\} = \lambda|G|\mu_b, \quad \boldsymbol{E}\{F_G\} = \lambda|G|\mu_f, \quad \text{and} \quad \boldsymbol{E}\{\Psi(G)\} = \lambda|G|.$$

Hence we can obtain

$$\mathrm{Var}\,\{\Psi(G)\} = \lambda|G| - \lambda^2|G|^2 + \int \mathbf{1}_G(x)\mathbf{1}_G(y) d\Lambda^{(2)}(x,m,y,l)$$

by putting $g(x,m) = h(x,m) = \mathbf{1}_G(x)$ in (2.4).

Proposition 1. *Define*

$$\mathrm{Cov}_h\,\{b(M), f(L)\} = \int b(m) f(l) d\mathcal{M}_h(m,l) - \mu_b\mu_f$$

and let $G_{-h} = \{x - h;\ x \in G\}$. Then the covariance of B_G and F_G is given by

$$\begin{aligned}\mathrm{Cov}\,\{B_G, F_G\} = \lambda^2 \int |G \cap G_{-h}|\,\mathrm{Cov}_h\,\{b(M), f(L)\}\, d\mathcal{K}(h) \\ + \lambda|G|\,\mathrm{Cov}\,\{b(M), f(M)\} + \mu_b\mu_f\,\mathrm{Var}\,\{\Psi(G)\}.\end{aligned} \tag{3.3}$$

PROOF. Since we have

$$\boldsymbol{E}\{B_G F_G\} = \lambda|G| \int b(m)f(m)d\mathcal{M}(m) + \int \mathbf{1}_G(x)\mathbf{1}_G(y)b(m)f(l)d\Lambda^{(2)}(x,m,y,l)$$

from (2.1) and (2.4), the covariance can be written as

$$\operatorname{Cov}\{B_G, F_G\} = \int \mathbf{1}_G(x)\mathbf{1}_G(y)(b(m)f(l) - \mu_b\mu_f)d\Lambda^{(2)}(x,m,y,l) + \lambda|G| \operatorname{Cov}\{b(M), f(M)\} + \mu_b\mu_f \operatorname{Var}\{\Psi(G)\}. \tag{3.4}$$

From (2.3) and (2.5), the first term on the right-hand side of (3.4) is equal to

$$\lambda^2 \int |G \cap G_{-h}| \left\{ \int b(m)f(l)d\mathcal{M}_h(x,m) - \mu_b\mu_f \right\} d\mathcal{K}(h).$$

This completes the proof of (3.3).

4. Asymptotic behaviors of moments

To get our main result, we need to know the asymptotic behavior of moments. In the following proposition, we assume the existence of a function ξ, which corresponds to a mixing coefficient of a random field, see Bolthausen (1982).

We partition $\mathbb{R}^d$ into congruent cubes

$$\Delta_i = \left\{ x = (x_1, \dots, x_d) \in \mathbb{R}^d;\, 2i_l - 1 \le x_l < 2i_l + 1,\, l = 1, \dots, d \right\}$$

for $i = (i_1, \dots, i_d) \in \mathbb{Z}^d$ and define $d(\Delta_i, \Delta_j) = \max_{1 \le k \le d} |i_k - j_k|$.

PROPOSITION 2. *Assume there exist a non-negative and non-increasing function ξ on the set of all nonnegative integers and some constants $\delta, k_1 > 0$ such that*

$$\sum_{n=1}^{\infty} n^{d-1} \xi(n)^{\delta/(2+\delta)} < \infty,$$

$$\|b(M)\|_{2+\delta} = \left(\int b^{2+\delta}(m)d\mathcal{M}(m) \right)^{1/(2+\delta)} < \infty,$$

$$\|f(M)\|_{2+\delta} = \left(\int f^{2+\delta}(m)d\mathcal{M}(m) \right)^{1/(2+\delta)} < \infty,$$

and

$$|\operatorname{Cov}_h\{b(M), f(L)\}| \le k_1 \|b(M)\|_{2+\delta} \|f(M)\|_{2+\delta} \xi(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)} \tag{4.1}$$

for any $h \in \Delta_i$. Then, as $G \uparrow \mathbb{R}^d$,

$$\operatorname{Cov}\{B_G, F_G\} = O(|G|) + \mu_b\mu_f \operatorname{Var}\{\Psi(G)\}. \tag{4.2}$$

PROOF. There exists a constant $K_1 > 0$ such that

$$\lambda^2 \int |G \cap G_{-h}| \operatorname{Cov}_h \{b(M), f(L)\} \, d\mathcal{K}(h) + \lambda |G| \operatorname{Cov} \{b(M), f(M)\}$$
$$\leq K_1 |G| \left\{ \sum_i \int_{\Delta_i} |\operatorname{Cov}_h \{b(M), f(L)\}| d\mathcal{K}(h) + |\operatorname{Cov} \{b(M), f(M)\}| \right\}.$$

From the present assumption, $|\operatorname{Cov}\{b(M), f(M)\}|$ is finite. In order to prove the assertion, it is enough to show

$$\sum_i \int_{\Delta_i} |\operatorname{Cov}_h \{b(M), f(L)\}| d\mathcal{K}(h) < \infty. \tag{4.3}$$

There exists a constant $K_2 > 0$ such that

$$\sum_i \int_{\Delta_i} |\operatorname{Cov}_h \{b(M), f(L)\}| d\mathcal{K}(h) \leq K_2 \sum_i \xi(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)} \int_{\Delta_i} d\mathcal{K}(h)$$

from the assumption (4.1). Since we have the inequality

$$\mathcal{K}(\Delta_i) \leq K_3 \tag{4.4}$$

for some constant $K_3 > 0$ from Lemma 1 below, there exist a constant $K_4 > 0$ such that

$$\begin{aligned} \sum_i \xi(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)} \mathcal{K}(\Delta_i) &\leq K_3 \sum_i \xi(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)} \\ &\leq K_4 \sum_{n=1}^{\infty} n^{d-1} \xi(n)^{\delta/(2+\delta)} < \infty. \end{aligned}$$

This shows (4.3) and the assertion follows.

LEMMA 1. *Under the same assumptions as in Proposition* 2, *we can show* (4.4).

PROOF. Let $\Delta_i \oplus \{x\} = \{z + x; z \in \Delta_i\}$, $x \in \mathbb{R}^d$ and D_i, $i \in \mathbb{Z}^d$, be

$$D_i = \Delta_i \oplus b(0, \sqrt{d}) = \left\{ a + b;\ a \in \Delta_i,\, b \in b(0, \sqrt{d}) \right\}$$

where $b(0, \sqrt{d})$ is the closed disk with center at 0 and radius $\sqrt{d}$. From the definition of $\mathcal{K}, \boldsymbol{P}_0^!$ and $\boldsymbol{P}_0$, we have

$$\begin{aligned} \mathcal{K}(\Delta_i) &= \frac{1}{\lambda} \int \psi(\Delta_i) d\boldsymbol{P}_0^!(\psi) \\ &\leq \frac{1}{\lambda} \int \psi(\Delta_i) d\boldsymbol{P}_0(\psi) = \frac{1}{\lambda^2 |\Delta_0|} \boldsymbol{E} \sum_{x \in \Psi \cap \Delta_0} \Psi(\Delta_i + x). \end{aligned}$$

The last equation is derived from (2.2). Since the inequality $\Psi(\Delta_i + x) \leq \Psi(D_i)$ always holds for arbitrary $x \in \Delta_0$ and Ψ is a stationary and 2nd order point process, we have

$$\boldsymbol{E} \sum_{x \in \Psi \cap \Delta_0} \Psi(\Delta_i + x) \leq \boldsymbol{E}\Psi(D_i)\Psi(D_0) \leq \boldsymbol{E}\Psi(D_0)^2 < \infty$$

by the Schwarz inequality. Hence the assertion follows.

Remark 1. Consider the case $f(m) = b(m)$. Then, by (3.3), the variance of B_G is written as

$$\begin{aligned} \operatorname{Var}\{B_G\} =& \lambda^2 \int |G \cap G_{-h}| \operatorname{Cov}_h \{b(M), b(L)\}\, d\mathcal{K}(h) \\ &+ \lambda |G| \operatorname{Var}\{b(M)\} + \mu_b^2 \operatorname{Var}\{\Psi(G)\}. \end{aligned}$$

The variance of B_G is also written as

$$\operatorname{Var}\{B_G\} = O(|G|) + \mu_b^2 \operatorname{Var}\{\Psi(G)\} \quad \text{as} \quad G \uparrow \mathbb{R}^d$$

provided that the assumptions in Proposition 2 are satisfied.

Remark 2. The function $\xi(n)$ is a kind of mixing coefficients which measures the dependency between random variables. In fact, it is a measure of the dependency between marks of points with distance n.

Remark 3. Let $X = \{X_i\}$ be a non-marked point process, $S = \{S(x);\ x \in \mathbb{R}^d\}$ be a stationary random field which is independent of X. As a special case of our model, we can construct the marked point process $\Phi = \{(X_i, S(X_i))\}$ as in Mase (1996). Let σ_i be the σ-algebra generated by $\{S(x);\ x \in \Delta_i\}$, and define a mixing coefficient

$$\begin{aligned} &\eta(n) \\ &= \sup \left\{ |\boldsymbol{P}(A_1 \cap A_2) - \boldsymbol{P}(A_1)\,\boldsymbol{P}(A_2)| :\ A_1 \in \sigma_0,\ A_2 \in \sigma_i,\ \sup_{1 \leq k \leq d} |i_k| \geq n \right\} \end{aligned}$$

as in Bolthausen (1982). Then, if there exists a constant $\delta > 0$ such that

$$\sum_{n=1}^{\infty} n^{d-1} \eta(n)^{\delta/(2+\delta)} < \infty, \quad \int b^{2+\delta}(m) d\rho(m) < \infty, \ \text{and} \ \int f^{2+\delta}(m) d\rho(m) < \infty,$$

where ρ is the distribution of $S(0)$, we have the inequality

$$\begin{aligned} &|\operatorname{Cov}_h \{b(M), f(L)\}| \\ &\quad \leq k_2\, \eta(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)} \\ &\qquad \times \left(\int b^{2+\delta}(m) d\rho(m) \right)^{1/(2+\delta)} \left(\int f^{2+\delta}(m) d\rho(m) \right)^{1/(2+\delta)} \end{aligned}$$

for some $k_2 > 0$, see Bolthausen (1982). Note that $\mathrm{Cov}_h\{b(M), f(L)\} = \mathrm{Cov}\{b(S(0)), f(S(h))\}$ in this case. Hence we can show that

$$\begin{aligned}
\mathrm{Cov}\{B_G, F_G\} &= O(|G|) + \left(\int b d\rho\right)\left(\int f d\rho\right)\mathrm{Var}\{\Psi(G)\},\\
\mathrm{Var}\{B_G\} &= O(|G|) + \left(\int b d\rho\right)^2 \mathrm{Var}\{\Psi(G)\},\\
\mathrm{Var}\{F_G\} &= O(|G|) + \left(\int f d\rho\right)^2 \mathrm{Var}\{\Psi(G)\}.
\end{aligned}$$

Remark 4. In order to get covariance estimates like (4.2), generally, it seems necessary to use mixing-type conditions. But, for particular problems, we may get direct evaluations. For example, let consider a stationary Poisson point process X with the intensity μ and let the associated mark M_x for $x \in X$ be its nearest neighbor distance $\min\{|x-y|; y \in X, y \neq x\}$. Fix two points $x, y \in X$ with $h = |x-y|$. It is known that, given x, y, $X\backslash\{x, y\}$ is also a stationary Poisson point process with the same mean μ. The distribution F_h of M_x, M_y is a mixture of the form $\mathbf{1}_{[0,h)}(r)p(r)dr + \exp(-\mu\pi h^2)\delta_h(dr)$, where $p(r) = 2\pi\mu r\exp(-\mu\pi r^2)$ and δ_h is the Dirac measure. Let $\mathcal{M}_h$ be the joint distribution of (M_x, M_y). $\mathcal{M}_h$ has the density $p(r)p(s)$ if $r + s < h$. Assume that

$$\int b^4(r)p(r)dr < \infty, \quad \text{and} \quad \int f^4(r)p(r)dr < \infty,$$

and

$$b(r) = O(\exp(\mu\pi r^2/4)), \quad \text{and} \quad f(r) = O(\exp(\mu\pi r^2/4)).$$

Then we have

$$\begin{aligned}
|\mathrm{Cov}_h\{b(M_x), f(M_y)\}| &= \int_{r\geq h/2 \text{ or } s\geq h/2} b(r)f(s)[d\mathcal{M}_h(r,s) - dF_h(r)dF_h(s)]\\
&\leq k\exp(-\mu\pi h^2/8),
\end{aligned}$$

for some constant $k > 0$, since

$$\begin{aligned}
\left|\int_{s\geq h/2} b(r)f(s)d\mathcal{M}_h(r,s)\right|^4 &\leq \left|\int_{s\geq h/2} b^2(r)d\mathcal{M}_h(r,s)\int_{s\geq h/2} f^2(s)d\mathcal{M}_h(r,s)\right|^2\\
&\leq \int b^4(r)dF_h(r)\int f^4(s)dF_h(s)\left(\int_{s\geq h/2} dF_h(s)\right)^2,
\end{aligned}$$

and

$$\begin{aligned}
&\left|\int_{s\geq h/2} b(r)f(s)dF_h(r)dF_h(s)\right|^4 \\
&\qquad \leq \left|\int b^2(r)dF_h(r)\int_{s\geq h/2} f^2(s)dF_h(s)\int_{s\geq h/2} dF_h(r)\right|^2 \\
&\qquad \leq \int b^4(r)dF_h(r)\int f^4(r)dF_h(r)\left(\int_{s\geq h/2} dF_h(r)\right)^2,
\end{aligned}$$

by the Schwarz inequality, similar inequalities also hold for integrals over the domain $\{r; r \geq h/2\}$, and

$$\begin{aligned}
&\int b^4(r)dF_h(r)\int f^4(s)dF_h(s)\left(\int_{s\geq h/2} dF_h(s)\right)^2 \\
&\quad \leq \left(\int b^4(r)p(r)dr + b^4(h)\exp(-\mu\pi h^2)\right) \\
&\qquad \times \left(\int f^4(r)p(r)dr + f^4(h)\exp(-\mu\pi h^2)\right) \times \exp(-\mu\pi h^2/2).
\end{aligned}$$

Therefore, for some $K > 0$,

$$\sum_i \int_{\Delta_i} |\operatorname{Cov}_h\{b(M_x), f(M_y)\}|d\mathcal{K}(h) \leq K\sum_i \exp(-\mu\pi d(\Delta_0,\Delta_i)^2/8) < \infty$$

holds.

One important model for which we can get the estimate (4.2) directly is the model $\Phi = \{(X_i, S(X_i))\}$ where S is a stationary Gaussian random field. For details about Gaussian random fields, see, e.g., Cressie (1993).

PROPOSITION 3. *Let S be a stationary Gaussian random field and its correlation function be the $c(h)$. We assume that*

$$\int b^2(m)d\rho(m) < \infty, \quad \text{and} \quad \int f^2(m)d\rho(m) < \infty, \tag{4.5}$$

and $c^*(r) = \sup_{|h|=r} c(h)$ is a non-increasing function such that

$$\sum_{n=1}^{\infty} n^{d-1}c^*(n) < \infty. \tag{4.6}$$

Then, as $G \uparrow \mathbb{R}^d$,

$$\operatorname{Cov}\{B_G, F_G\} = O(|G|) + \left(\int b d\rho\right)\left(\int f d\rho\right)\operatorname{Var}\{\Psi(G)\}.$$

PROOF. For simplicity, let $\boldsymbol{E}S(0) = 0$ and $\mathrm{Var}\,\{S(0)\} = 1$. By the Schwarz inequality,

$$
\begin{aligned}
&|\,\mathrm{Cov}\,\{b(S(0)), f(S(h))\}\,|^2 \\
&= \left| \int\int b(x)f(y)\frac{e^{-R/2}}{2\pi}\left[\frac{1}{\sqrt{1-c^2(h)}}e^{-(Q-R)/2} - 1\right] dxdy \right|^2 \\
&\le \int b^2 d\rho \int f^2 d\rho \int\int \frac{e^{-R/2}}{2\pi}\left[\frac{1}{\sqrt{1-c^2(h)}}e^{-(Q-R)/2} - 1\right]^2 dxdy \\
&= \int b^2 d\rho \int f^2 d\rho \times \frac{c^2(h)}{1-c^2(h)}
\end{aligned}
$$

where $Q = (x^2 - 2c(h)xy + y^2)/(1 - c^2(h))$ and $R = x^2 + y^2$. Since $c(h) \to 0$ as $|h| \to \infty$, there exists a finite set $I \subset \mathbb{Z}^d$ such that

$$
\frac{c^2(h)}{1-c^2(h)} \le 2c^2(h) \quad \text{if} \quad h \in \bigcup_{i \in \mathbb{Z}^d \setminus I} \Delta_i.
$$

Then we have

$$
\begin{aligned}
&\int |\,\mathrm{Cov}\,\{b(S(0)), f(S(h))\}\,| d\mathcal{K}(h) \\
&\quad \le \sum_{i \in I} \int_{\Delta_i} |\,\mathrm{Cov}\,\{b(S(0)), f(S(h))\}\,| d\mathcal{K}(h) \\
&\qquad + K_5 \sum_{i \in \mathbb{Z}^d \setminus I} c^*(d(\Delta_0, \Delta_i))\mathcal{K}(\Delta_i) < \infty
\end{aligned}
$$

for some constant $K_5 > 0$. Hence, the assertion follows.

5. Main result

Now we are in position to state a sufficient condition under which the threshold method works fine.

PROPOSITION 4. *Assume there exists a non-negative and non-increasing function ξ on the set of all nonnegative integers and some constants $\delta, k_4 > 0$ such that*

$$
\begin{aligned}
&\sum_{n=1}^{\infty} n^{d-1}\xi(n)^{\delta/(2+\delta)} < \infty, \\
&\|M\|_{2+\delta} = \left(\int m^{2+\delta} d\mathcal{M}(m)\right)^{1/(2+\delta)} < \infty,
\end{aligned}
$$

and

$$
|\,\mathrm{Cov}_h \left\{\mathbf{1}_{[c,\infty)}(M), L\right\}| \le k_4 \|\mathbf{1}_{[c,\infty)}(M)\|_{2+\delta} \|L\|_{2+\delta} \xi(d(\Delta_0, \Delta_i))^{\delta/(2+\delta)}
$$

for any $h \in \Delta_i$. Then,

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

if the conditions

$$\frac{1}{|G|} \operatorname{Var}\{\Psi(G)\} \to \infty \quad as \quad G \uparrow \mathbb{R}^d \tag{5.1}$$

and $\int_c^\infty d\mathcal{M} \neq 0$ are satisfied.

Proof. The mean squared prediction error is given by

$$\min_{\alpha,\beta} \boldsymbol{E} \left| \frac{F_G}{\sqrt{\operatorname{Var}\{F_G\}}} - \alpha - \beta \frac{B_G}{\sqrt{\operatorname{Var}\{B_G\}}} \right|^2 = 1 - \operatorname{Corr}\{B_G, F_G\}^2$$

from (3.2). Since $\operatorname{Corr}\{B_G, F_G\}$ is of the form

$$\frac{O(|G|) + \int_c^\infty d\mathcal{M} \int_0^\infty m d\mathcal{M}(m) \operatorname{Var}\{\Psi(G)\}}{\sqrt{O(|G|) + (\int_c^\infty d\mathcal{M})^2 \operatorname{Var}\{\Psi(G)\}} \sqrt{O(|G|) + (\int_0^\infty m d\mathcal{M}(m))^2 \operatorname{Var}\{\Psi(G)\}}}$$

by Proposition 2, the assertion follows.

Remark 5. Consider the model $\Phi = \{(X_i, S(X_i))\}$ where S is a stationary Gaussian random field. We can show the same assertion as in Proposition 4 without using the function ξ if the assumptions (4.6) and (5.1) are satisfied for $b(m) = \mathbf{1}_{[c,\infty)}(m)$ and $f(m) = m$.

Remark 6. From Proposition 4, we can see that the threshold method works fine if the condition (5.1) is satisfied. Furthermore, we can predict $F_G/|G|$ from $B_G/|G|$ if $\operatorname{Var}\{\Psi(G)\} = O(|G|^2)$ since

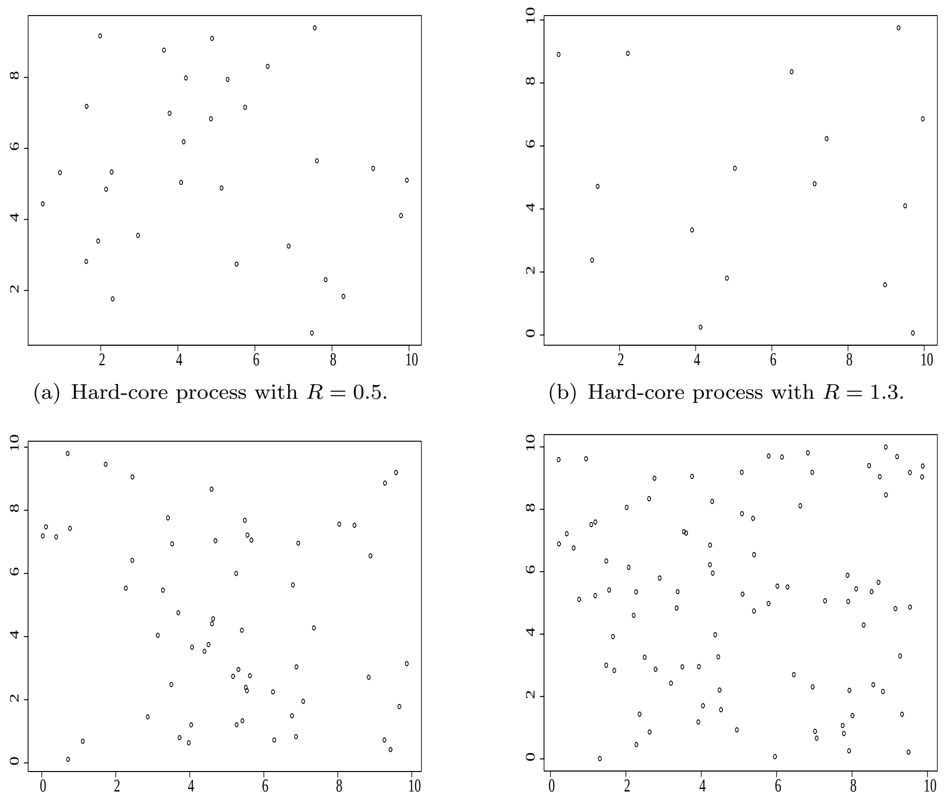
$$\min_{\alpha,\beta} \boldsymbol{E} \left| \frac{F_G}{|G|} - \alpha - \beta \frac{B_G}{|G|} \right|^2 = \frac{\operatorname{Var}\{F_G\}}{|G|^2} \left\{ 1 - \operatorname{Corr}\{B_G, F_G\}^2 \right\} \to 0.$$

These results extend those of Mase (1996). Note that the condition $\operatorname{Var}\{\Psi(G)\} = O(|G|^2)$ implies the non-ergodicity of Ψ.

6. Simulation study

In this section, we illustrate the previous theoretical result by numerical experiments. We assume the distribution $\boldsymbol{P}$ of X is a mixture

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



(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. Realizations of four point processes in a 10×10 square region.

and distributions $\boldsymbol{P}_i$ are:

$\boldsymbol{P}_1$: 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 point process that distances of arbitrary pairs of points are larger than $2R$. Figure 1 shows their realizations. The process X can be considered to model a phenomenon that the degree of interactions between points varies. As to algorithms of generating X, see, e.g., van Lieshout (2000).

Let S be a stationary and isotropic Gaussian random field. Its mean vector is $(20, 20, \ldots, 20)^t$ and the covariance matrix is given as

$$\begin{pmatrix} 10 & 10-\gamma(x_1-x_2) & \ldots & 10-\gamma(x_1-x_n) \\ 10-\gamma(x_2-x_1) & 10 & \ldots & 10-\gamma(x_2-x_n) \\ \vdots & \vdots & \ddots & \vdots \\ 10-\gamma(x_n-x_1) & 10-\gamma(x_n-x_2) & \ldots & 10 \end{pmatrix}.$$

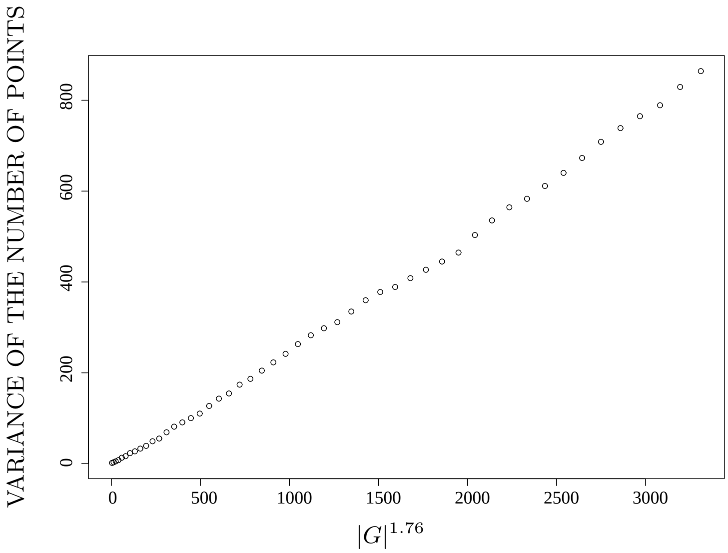


Figure 2. Var $\{\Psi(G)\}$ for the mixture process versus $|G|^{1.76}$.

Table 1. Threshold values c and corresponding coefficient of determination R^2-value for the mixture process.

c	13	14	15	16	17	18	19	20
R^2	0.9599	0.9644	0.9618	0.9473	0.9225	0.8755	0.7860	0.6353

Here, γ is an exponential semivariogram

$$\gamma(h) = 10(1 - \exp(-|h|/10)),$$

see, e.g., Cressie (1993) for details about semivariograms and their corresponding Gaussian random fields. Note that this model satisfies (4.5) and (4.6) for $b(m) = \mathbf{1}_{[c,\infty)}(m)$ and $f(m) = m$. In total, 100 marked point processes are generated.

Figure 2 is the graph of Var $\{\Psi(G)\}$ versus $|G|^{1.76}$. It is seen that Var $\{\Psi(G)\}$ $\simeq |G|^{1.76}$ and, hence, the data satisfy the condition (5.1). Therefore, the threshold method should work fine.

Table 1 is the result of the threshold method. We can see that the sum of marks can be predicted well even for a wide range of threshold values.

Next, we also generate ergodic Poisson processes with intensity 0.5. Figure 3 is the graph of Var $\{\Psi(G)\}$ versus $|G|^{0.93}$ and shows Var $\{\Psi(G)\} \simeq |G|^{0.93}$. The condition (5.1) is not satisfied.

Table 2 seems to show that the threshold method may work even for the ergodic process $\boldsymbol{P}_1$ if we choose an appropriate threshold value, say $c = 14$. But almost all marks are larger than 14 in this case and this is nothing but the effect of the law of large numbers.

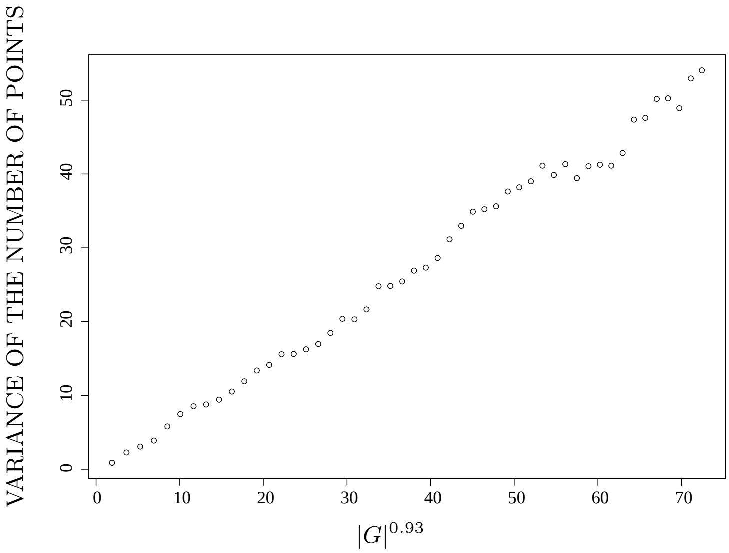


Figure 3. Var $\{\Psi(G)\}$ for the Poisson process with $\lambda = 0.5$ versus $|G|^{0.93}$.

Table 2. Threshold values c and corresponding coefficient of determination R^2-values for the Poisson process with $\lambda = 0.5$.

c	13	14	15	16	17	18	19	20
R^2	0.6447	0.7039	0.6956	0.6641	0.6456	0.6474	0.6535	0.6455

7. Conclusions and remarks

In this paper, the threshold method was studied in the framework of the theory of marked point processes on $\mathbb{R}^d$. We showed that the threshold method works fine if the correlation of marks becomes weaker as the distance of points becomes larger and (5.1) is satisfied.

We did not discuss how to estimate regression coefficients. Of course, if complete training data are available, we can estimate the coefficients directly. Even if we can only get thresholded data as assumed in this paper, we can still employ, e.g., the method of Shimizu *et al.* (1993), or Short *et al.* (1993).

Acknowledgements

Authors express their thanks to the referees for their kind and detailed comments. Their suggestions were very useful to refine 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, Jan. 31–Feb. 4, 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, *J. Clim. Appl. Meteor.*, **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. (1996). The threshold method for estimating total rainfall, *Ann. Inst. Statist. Math.*, **48**, 201–213.

Shimizu, K. ed. (2002). *Global Environmental Data* (in Japanese), Kyoritsu Publishing Co., Tokyo.

Shimizu, K., Short, D. A. and Kedem, B. (1993). Single- and double-threshold methods for estimating the variance of area rain rate, *J. Meteor. Soc. 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, *J. Appl. Meteor.*, **32**, 182–192.

Stoyan, D. (1984). On correlations of marked point processes, *Math. Nachr.*, **116**, 197–207.

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

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*, Wiley, Chichester.

van Lieshout, M. N. M. (2000). *Markov Point Process and Their Applications*, Imperial College Press.