Ann. Inst. Statist. Math.
33 (1981), Part B, 315-338

ESTIMATION OF INTERACTION POTENTIALS OF SPATIAL POINT PATTERNS THROUGH THE MAXIMUM LIKELIHOOD PROCEDURE

YOSIHIKO OGATA AND MASAHARU TANEMURA

(Received Oct. 11, 1980)

Summary

A homogeneous spatial point pattern is regarded as one of thermal equilibrium configurations whose points interact on each other through a certain pairwise potential. Parameterizing the potential function, the likelihood is then defined by the Gibbs canonical ensemble. A Monte Carlo simulation method is reviewed to obtain equilibrium point patterns which correspond to a given potential function. An approximate log likelihood function for gas-like patterns is derived in order to compute the maximum likelihood estimates efficiently. Some parametric potential functions are suggested, and the Akaike Information Criterion is used for model selection. The feasibility of our procedure is demonstrated by some computer experiments. Using the procedure, some real data are investigated.

1. Introduction

Recently the study of statistical analysis for multidimensional point process (or random point field) has been rapidly developing (see Fisher [7] and Ripley [18] for extensive references). The test of the randomness of spatial patterns (i.e. test of Poisson random field as the null hypothesis) has been discussed by many authors through some nonparametric approaches such as quadrat and distance methods for quite long time. The second order analysis, which was extensively developed recently (Bartlett [2], Ripley [18], Vere-Jones [22]), would also be among the familiar available existing methods. Modelling of clustering Poisson fields is also developed by making use of the probability generating functional (Vere-Jones [21]). However, this does not seem to be well linked with the likelihood procedure. In fact the likelihood function of the clustering Poisson field is too complicated to manage unless the cluster centres are known (Baudin [3]).

Then, what is a key to the likelihood analysis? For one dimen-

sional point process it is a conditional intensity function, through which a variety of statistical models can be proposed (see Ogata [15] and [17], for example). It is suggested in this paper that a pairwise potential function could play a similar role to the conditional intensity. This is because the likelihood can be defined, as we see in Section 2, through the Gibbs distribution for the configuration of points, which is characterized by the potential function. Thus we parameterize the function so as to obtain the maximum likelihood estimate. Modelling of spatial patterns which takes interaction among points into account was considered, for example, by Matérn [10], Ripley [18], Besag and Diggle [4], Hasegawa and Tanemura [8], Diggle [5] and Tanemura and Hasegawa [20]. Nevertheless, the maximum likelihood procedure has not so far been developed.

In the next section the likelihood function is introduced. Then, in Section 3, the Metropolis' simulation method is briefly reviewed in order to obtain the equilibrium spatial point patterns, as well as to estimate the value of the partition function. In Section 4 an approximate log likelihood function is derived under the assumption that an observed spatial pattern belongs to gaseous equilibrium configurations with low density. Some examples of parameterized pairwise potential functions are suggested in Section 5 and the model selection procedure by using AIC is explained in Section 6. Then, in Section 7, some numerical examples are given to illustrate our procedure. Some real observations are considered using our method in Section 8. The final section is allotted for further remarks.

2. Likelihood

Suppose that we have a homogeneous spatial point pattern in a finite region V. We regard this as one of the thermal equilibrium configurations under some interaction potential energy which may or may not be realistic in the sense of statistical mechanics. Further it is assumed here that the equilibrium distribution of the configuration is characterized by a pairwise potential function. We are, then, interested in estimating the form of the function from an observed spatial point pattern.

Consider a family of parameterized pairwise potential functions

$$\{\Phi_\theta(r);\ \theta \in \Theta\} \tag{1}$$

of the Euclidean distance r. Let a finite set of points $\boldsymbol{X}=\{X_n;\ n=1, 2,\cdots, N\}$ be observed in the region V. Then the likelihood of the potential function $\Phi_\theta(r)$ is given by the Gibbs canonical ensemble (see Feynman [6], for example)

$$L(\boldsymbol{X};\theta)=\exp\{-U(\boldsymbol{X};\theta)\}/Z_N(\theta)\,, \tag{2}$$

where $U(\boldsymbol{X};\theta)$ is the total potential energy which is assumed to be given as a sum of pairwise potentials, i.e.

$$U(\boldsymbol{X};\theta)=\sum_{n=1}^{N-1}\sum_{m=n+1}^{N}\Phi_\theta(|X_n-X_m|)\,, \tag{3}$$

and

$$Z_N(\theta)=\int_{V^N}\exp\{-U(\boldsymbol{X};\theta)\}dX_1\cdots dX_N \tag{4}$$

is the configurational part of the classical partition function of N interacting particles. Here we have set $\Phi_\theta(r)$ instead of $\Phi_\theta(r)/(k_BT)$ where T is the temperature and k_B is the Boltzmann's constant. This means that the potential function here is to be estimated together with the effect of the temperature. To avoid the boundary effect, we assume that points are distributed under the periodic boundary conditions (i.e. on the two dimensional torus) from now on.

3. Monte Carlo simulation

To give some feeling of the relation between a pairwise potential and its equilibrium point patterns, let us review a simulation procedure which uses a particular type of random walk known as a Markov chain. The simulation was originally devised by Metropolis et al. [12] and developed by Wood [23] and others for the study of a liquid phase.

Consider a set of particles, interacting according to a certain potential function, on a square V with periodic boundary (i.e. V is identified with a torus). In order to make the situation simple, assume that the square is discretized into an $M\times M$ square lattice on which N particles are attached. Then each configuration of the particles on the lattice is counted as one of the states of a Markov chain where the total number of possible states is $B=\binom{M^2}{N}$. The Gibbs distribution (2) is also discretized into probabilities $\{u_i;\ i=1,2,\cdots,B\}$. In essence, then, the Monte Carlo procedure here is simply to select the transition probabilities (P_{ij}) which satisfy $\sum_{i=1}^{B}u_iP_{ij}=u_j$ for all $j=1,2,\cdots,B$ and that the n-step transition probability $P_{ij}^{(n)}$ converges to the given equilibrium probability u_i. The most commonly used transition probability is

$$P_{ij}=\begin{cases} 0 & \text{if } j \text{ not in } W(i), \\ 1/K & \text{if } j\in W(i),\ j\neq i, \text{ and } u_j\geqq u_i, \\ (u_j/u_i)/K & \text{if } j\in W(i) \text{ and } u_j<u_i, \end{cases} \tag{5}$$

and

$$P_{ii}=1-\sum_{j\neq i} P_{ij},$$

where $W(i)$ denotes a specified set of neighbour states j of state i, such that $i\in W(i)$, $i\in W(j)$ if and only if $j\in W(i)$, and K denotes the number of states in $W(i)$.

The Markov chain defined by (5) is realized for our purpose in the following manner: Assume that at time t, the state of the N particle system is $\boldsymbol{X}(t)=\{X_n(t)=(x_n(t), y_n(t))\in V;\ n=1,\cdots,N\}$. A trial state $\boldsymbol{X}'(t)$ is then chosen randomly and uniformly (i.e. with equal probability $1/K$ for each state) from the neighbour set of states $W(\boldsymbol{X}(t))=W_t$. For example, the sets W_t might consist of all those states j for which the coordinates $(x_n'(t), y_n'(t))$ of a randomly chosen particle n $(n=1, 2, \cdots, N)$ lie in the set $\{(x_n(t)+k, y_n(t)+m);\ |k|, |m|=0, 1,\cdots, k_{\max}\}$ where $k_{\max}$ is a certain integer, while all other $N-1$ particles have the same position as in state $\boldsymbol{X}(t)$ ($K=(2k_{\max}+1)^2$ in this case). The corresponding value of the probability $u'(t)=L(\boldsymbol{X}'(t);\theta)$ in (2) is then calculated and compared with $u(t)=L(\boldsymbol{X}(t);\theta)$. If $u'(t)\geqq u(t)$, then without further ado the next state is taken as $\boldsymbol{X}(t+1)=\boldsymbol{X}'(t)$. If $u'(t)<u(t)$ then we obtain a uniform random number $0\leqq\xi<1$ and (i) if $\xi\leqq u'(t)/u(t)$, then $\boldsymbol{X}(t+1)=\boldsymbol{X}'(t)$; (ii) if $\xi>u'(t)/u(t)$, then $\boldsymbol{X}(t+1)=\boldsymbol{X}(t)$. It is quite important to notice that the common normalizing factor $Z_N(\theta)$ in (2) can be cancelled out in the above comparisons. That is to say, we can put $u(t)=\exp\{-U(\boldsymbol{X}(t);\theta)\}$ without loss of generality.

Thus, starting from an initial state $\boldsymbol{X}(0)$ we have a state $\boldsymbol{X}(T)$ for sufficiently large T as a realization of equilibrium under the Gibbs distribution (2) which is characterized by a prescribed potential function (1). Furthermore the time average

$$\frac{1}{T}\sum_{t=1}^{T}\exp\{U(\boldsymbol{X}(t);\theta)\} \tag{6}$$

can be a good estimate of $|V|^N/Z_N(\theta)$, since

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}F(\boldsymbol{X}(t))=\int_{V^N}F(\boldsymbol{X})L(\boldsymbol{X};\theta)dX_1\cdots dX_N \tag{7}$$

for suitable function F of the configurations. Therefore an estimated likelihood value for a fixed θ is

$$(8) \qquad \tilde{L}(\boldsymbol{X};\theta)=\frac{1}{T}\sum_{t=1}^{T}\exp\{U(\boldsymbol{X}(t);\theta)-U(\boldsymbol{X};\theta)-N\log|V|\}\ ,$$

where $\boldsymbol{X}$ is an observation while $\boldsymbol{X}(t)$, $t=1,\cdots,T$ are the sequence of generated configurations under the potential $\Phi_\theta(\cdot)$.

4. An approximate likelihood

In statistical physics, it seems to be the main theme to obtain the value of partition function as a function of certain parameters such as volume and temperature, because, once the partition function is known, all thermodynamic properties can be derived from it in principle. It is, however, hard in general to calculate the value of partition function analytically because of the high multiplicity of the integral in (4). One way to overcome this difficulty is to compute (4) numerically by brute force. This is one of the roles of the Monte Carlo method described in the previous section (precisely speaking, in the usual computer experiments, thermodynamical quantities are directly computed through (7), not passing through the partition function (see Wood [23]). The other way is to approximate (4). There have been developed a number of approximation methods, most of them working well for the equilibrium configurations of particles in the gaseous and solid phases.

Since one of the purposes of this work is to show the feasibility of the likelihood procedure suggested in Section 2, it is desirable to use an approximate partition function, although in principle the Monte Carlo method might work for the numerical computation of the likelihood function (see (8)) for every set of parameter values.

When the configuration of points can be regarded as in the phase of low density gas, the cluster expansion method (for example, see Mayer and Mayer [11]) is applied. Here, we use the second cluster approximation, that is, an approximation which includes only two body interactions and neglects three body and higher order collisions. In that case, the partition function (4) is given in the following form (Feynman [6]):

$$(9) \qquad Z_N(\theta)=|V|^N(1-a(\theta)/|V|)^{N(N-1)/2}$$

where $|V|$ denotes the volume of the region V, and

$$(10) \qquad a(\theta)=\int_0^\infty[1-\exp\{-\Phi_\theta(r)\}]c(r)dr$$

($c(r)=2\pi r$ or $4\pi r^2$ for R^2 or R^3 respectively). A brief derivation of (9) is given in Appendix I. It is emphasized here that the influential region of the potential should be small enough compared with the specific

volume $|V|/N$. Thus, from (2) and (9), we obtain an approximate log likelihood

$$\log L(\boldsymbol{X};\theta)=-\sum_{n=1}^{N-1}\sum_{m=n+1}^{N}\Phi_\theta(|X_n-X_m|) \tag{11}$$
$$-\frac{1}{2}N(N-1)\log\left\{1-\frac{a(\theta)}{|V|}\right\},$$

where the common constant $N\log|V|$ has been neglected.

5. Models

Let us list a few examples of parameterized potential functions for later use.

PF I ; $\Phi_\theta(r)=-\log\{1+(\alpha r-1)e^{-\beta r^2}\}$, $\theta=(\alpha,\beta)$, $\alpha\geqq 0$, $\beta>0$,

PF II ; $\Phi_\theta(r)=-\log\{1+(\alpha-1)e^{-\beta r^2}\}$, $\theta=(\alpha,\beta)$, $\alpha\geqq 0$, $\beta>0$,

PF III ; $\Phi_\theta(r)=\beta(\sigma/r)^n-\alpha(\sigma/r)^m$, $\theta=(\alpha,\beta,\sigma)$, $n>m$, $\beta>0$.

PF I has both ranges of repulsive and attractive force except the case

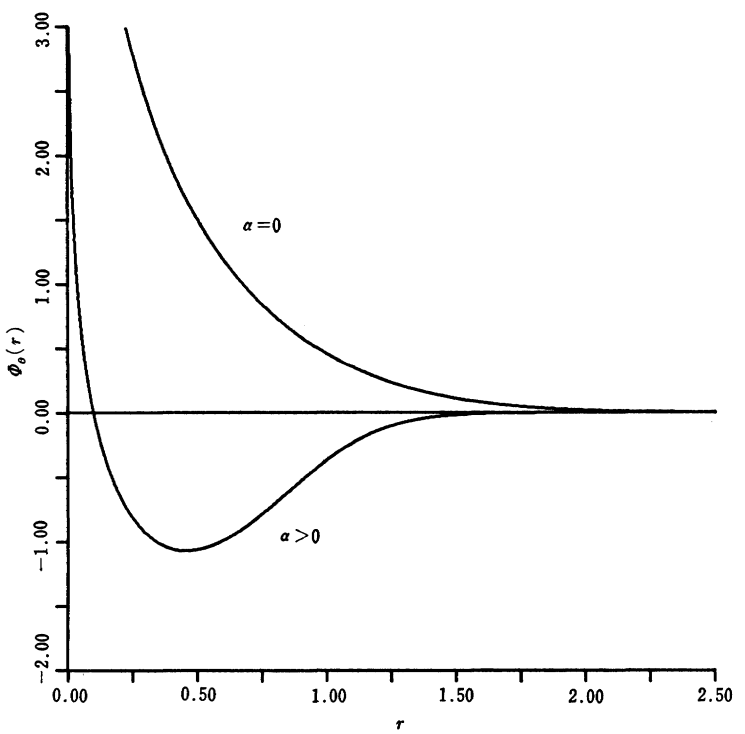

Fig. 1. Potential curves for the model PF I. The case $\alpha=0$ represents purely repulsive potential, while the case $\alpha>0$ has both repulsive and attractive ranges.

$\alpha=0$ (see Fig. 1). PF II provides a variety of type, that is, repulsive for $0\leqq\alpha<1$, ideal gas (Poisson) for $\alpha=1$, and attractive for $\alpha>1$ (see Fig. 2). PF III is known as one of the realistic potentials in statistical mechanics, being called the Lennard-Jones type potential, and has something like an inhibitory region (see Fig. 3). The second cluster integral $a(\theta)$ defined at (10) in the preceding section can be obtained analytically. The corresponding values are, in order,

PF I ; $a(\alpha, \beta)=(\pi/\beta)(1-\alpha\sqrt{\pi/\beta}/2)$,

PF II ; $a(\alpha, \beta)=\pi(1-\alpha)/\beta$,

PF III ; $a(\alpha, \beta, \sigma)=-\frac{\pi}{m}\beta^{1/m}\sigma^2\sum_{k=0}^{\infty}\frac{1}{k!}\Gamma\left(\frac{mk-2}{2m}\right)\alpha^k\beta^{-k/2}$, $m>2$, for $n=2m$.

The derivation of $a(\theta)$ for PF III is reviewed in Appendix II. Actually we consider only the case of $n=12$ and $m=6$ in this paper.

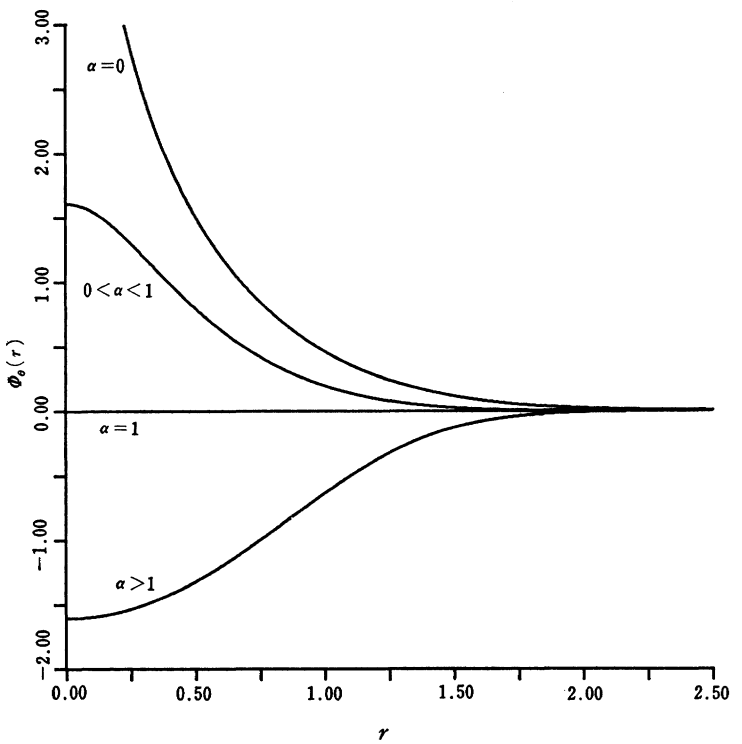

Fig. 2. Potential curves for the model PF II. The cases $\alpha=0$ and $0<\alpha<1$ correspond to repulsive potential, the case $\alpha=1$ to ideal gas, i.e. Poisson, and the case $\alpha>1$ to purely attractive potential.

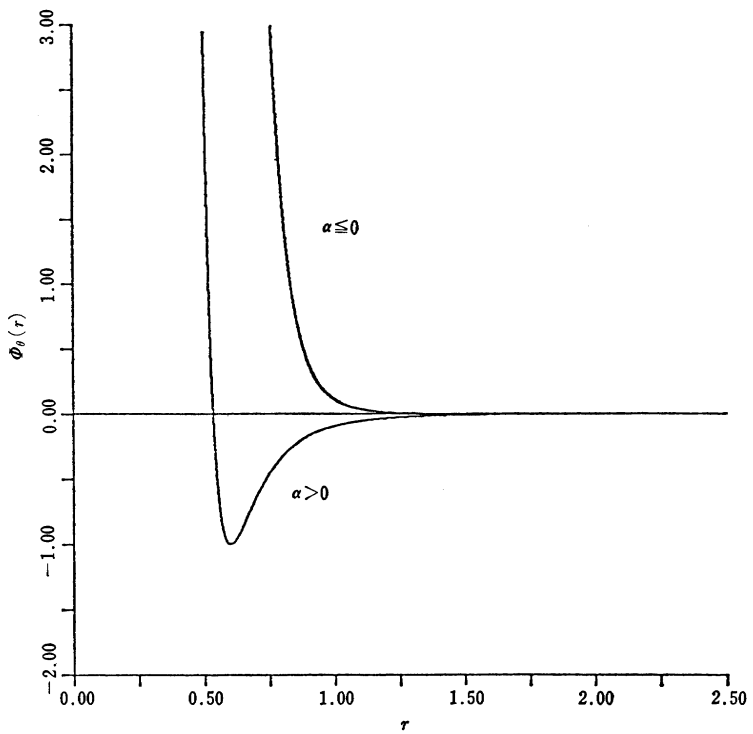

Fig. 3. Potential curves for the model PFIII ($n=12$, $m=6$). The case $\alpha \leqq 0$ corresponds to purely repulsive potential, while the case $\alpha > 0$ has an attractive part.

6. Model selection

The Akaike Information Criterion (AIC) is defined by AIC$=(-2)$ ·(max log likelihood)$+2\cdot$(number of adjusted parameters). This is an estimate of the expected negentropy which is a natural measure of discrimination between the true and estimated distributions of the data. We adopt the model which minimizes the value of AIC among the above competing models. Another selection of models could be performed by the likelihood ratio test of the null hypothesis model H_0 against the alternative model H_1, provided that the model H_1 contains H_0 as a special case (i.e. nested sequence of models). The relationship between the AIC and the likelihood ratio statistic is given by

$$\Lambda(H_0; H_1) = \mathrm{AIC}\,(H_0) - \mathrm{AIC}\,(H_1) + 2k\ ,$$

where k is the difference between the degrees of freedom of H_0 and H_1 (Akaike [1]). Notice that models PF II and III contain the Poisson case (i.e. $\Phi_\theta(r) \equiv 0$), and also that models PF I and III contain a repulsive potential as a special case. For such nested situations the mini-

mum AIC procedure is nothing but a χ^2-test, taking the variance (i.e. $2k$) as the level of significance. It is emphasized that we can further extend the comparison among non-nested models by using AIC, which was originally derived for such a purpose (Akaike [1]). (See Ogata [16] for a detailed review on the derivation of the AIC, and also see Sakamoto and Akaike [19] for a comparative study between the minimum AIC and the likelihood ratio test procedure.)

7. Computer experiments

Firstly we fitted the models PF I–III to a Poisson pattern which is illustrated in Fig. 4, and the corresponding values of the maximum log likelihood (11) and AIC are listed in Table 1. It should be noticed here that the value of both statistics are always zero for the Poisson model, since it corresponds to the null potential function ($\Phi_0(r)\equiv 0$). Thus Table 1 shows that the Poisson model is the best, according to the minimum AIC procedure.

Secondly, using the algorithm given in Section 3, we independently generated three equilibrium patterns of the model PF I with its parameters $\alpha=5.0$ and $\beta=2.0$. After this the model PF I was fitted twice (once with restriction $\alpha=0$) to each data. Furthermore we generated

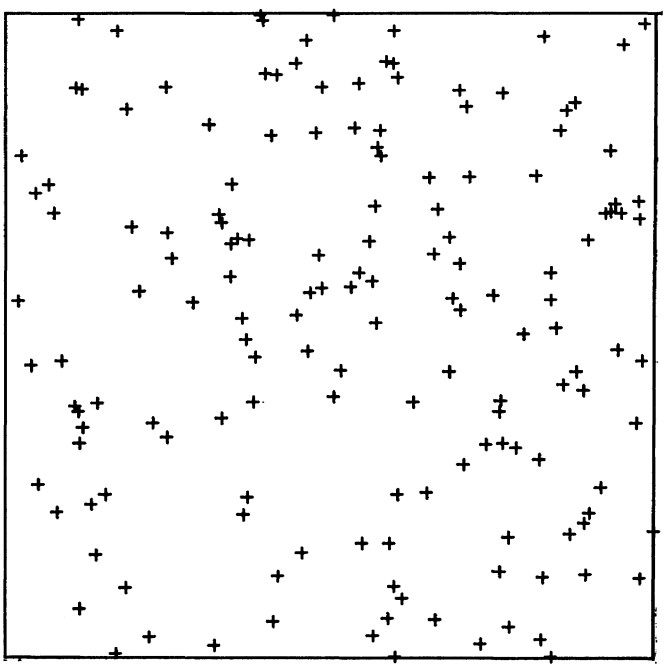

Fig. 4. Sample of Poisson pattern used for the analysis. This is only a part of the pattern of 500 points.

Table 1. Fitting of models to a Poisson pattern (N=500)

Model	Estimated parameters			log $\hat{L}$	AIC
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$		
PF I	13.25	77.28	—	−0.17	3.7
PF II	0.00	5.54×10^4	—	−0.00	4.0
PF III	2.06	0.65	0.15	−1.88	2.2
Poisson	—	—	—	0.00	0.0

Table 2. Simulations for PF I and fitting of the same model

True parameters		AIC(T)	Estimated parameters		AIC(E)
α_0	β_0		$\hat{\alpha}$	$\hat{\beta}$	
5.0	2.0	−20.9	3.97	1.52	−17.8
			0.0 (fixed)	37.62	1.4
5.0	2.0	−14.7	7.34	3.56	−15.1
			0.0 (fixed)	37.77	0.6
5.0	2.0	7.4	2.22	1.50	−11.2
			0.0 (fixed)	43.18	0.9
0.0	1.0	−24.0	0.00	1.11	−20.1
			0.0 (fixed)	1.11	−22.1
0.0	1.0	−36.6	0.00	0.86	−33.1
			0.0 (fixed)	0.86	−35.1
0.0	1.0	−30.5	0.00	1.05	−26.5
			0.0 (fixed)	1.05	−28.5

For all simulated patterns, the number of points is N=500. The second line in each data gives the estimates with restriction α=0.0. In this and the following tables, AIC(T) and AIC(E) represent the values of AIC for the true and the estimated parameters.

Table 3. Simulations for PF II and the fitting of the same model

Number of points N	True parameters		AIC(T)	Estimated parameters		AIC(E)
	α_0	β_0		$\hat{\alpha}$	$\hat{\beta}$	
270	3.0	1.0	−12.2	3.18	1.53	−8.9
				0.0 (fixed)	563.09	2.0
270	3.0	1.0	−54.8	12.11	3.78	−72.0
				0.0 (fixed)	565.23	2.0
270	3.0	1.0	−30.1	3.70	0.85	−28.8
				0.0 (fixed)	224.73	1.9
500	0.0	1.0	−20.7	0.13	0.87	−18.1
				0.0 (fixed)	1.12	−18.9
500	0.0	1.0	−35.4	0.00	0.87	−31.8
				0.0 (fixed)	0.87	−33.8
500	0.0	1.0	−29.4	0.00	0.90	−25.6
				0.0 (fixed)	0.90	−27.6

Table 4. Simulations for PF III and the fitting of the same model

Number of points N	True parameters α_0	β_0	σ_0	AIC (T)	Estimated parameters $\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	AIC (E)
429	0.0	10.0	1.0	−78.7	0.00	9.88	0.93	−79.6
					0.0 (fixed)	9.78	0.94	−81.6
429	0.0	10.0	1.0	−105.0	1.53	10.35	1.08	−101.7
					0.0 (fixed)	10.02	1.04	−102.7
429	0.0	10.0	1.0	−103.0	3.99	22.18	1.04	−102.0
					0.0 (fixed)	10.01	1.03	−99.9
500	4.0	2.0	0.2	−18.2	4.41	2.11	0.21	−14.7
					0.0 (fixed)	0.50	0.17	3.3
500	4.0	2.0	0.2	−28.0	4.34	1.75	0.21	−25.2
					0.0 (fixed)	0.01	0.23	2.4
500	4.0	2.0	0.2	−17.7	4.27	2.01	0.21	−13.1
					0.0 (fixed)	2.00	0.15	2.3

three equilibrium patterns of the same model with $\alpha=0$ and $\beta=1.0$. Then the model PF I was fitted again in the same way. The values of the true parameters, maximum likelihood estimates and AIC of each experiment are listed in Table 2. Similar experiments are performed for the models PF II and III and corresponding values are listed in Tables 3 and 4 respectively. These tables show that AIC $(T)<$AIC (E) holds for most cases, which is expected if the simulation of the patterns was performed adequately. Note here that AIC (T) is nothing but (-2) log (likelihood at the true parameter), since we have no adjusted parameters in this case. These tables also show that the minimum AIC procedure can distinguish the purely repulsive case $\alpha=0$ from other cases $\alpha>0$ within each model.

Lastly we generated three spatial patterns (Figs. 5 (a)–7 (a)) whose

Table 5. Cross fitting of models to simulated data

Data	N	True parameters α_0	β_0	σ_0	Fitted model	AIC	Number of parameters k
PF I	500	5.0	2.0	—	PF I	−9.1	2
					PF II	−6.1	2
					PF III	4.2	3
PF II	500	1.8	1.0	—	PF I	−4.1	2
					PF II	−9.8	2
					PF III	1.8	3
PF III	500	4.0	2.0	0.2	PF I	−4.1	2
					PF II	−5.5	2
					PF III	−22.9	3

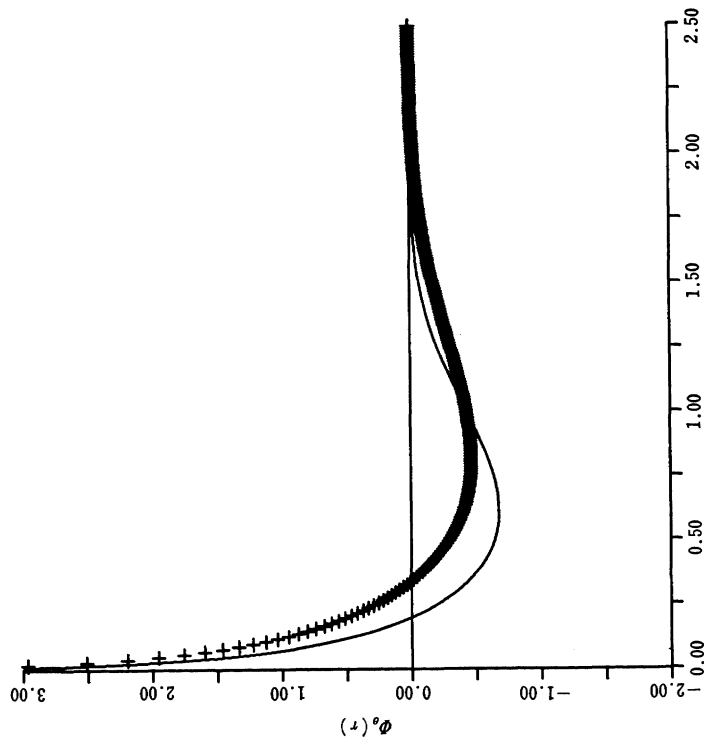

Fig. 5(b). Potential curves for true parameters (presented by a solid line) and for estimated parameters (presented by the symbol +), both of which belong to the model PF I.

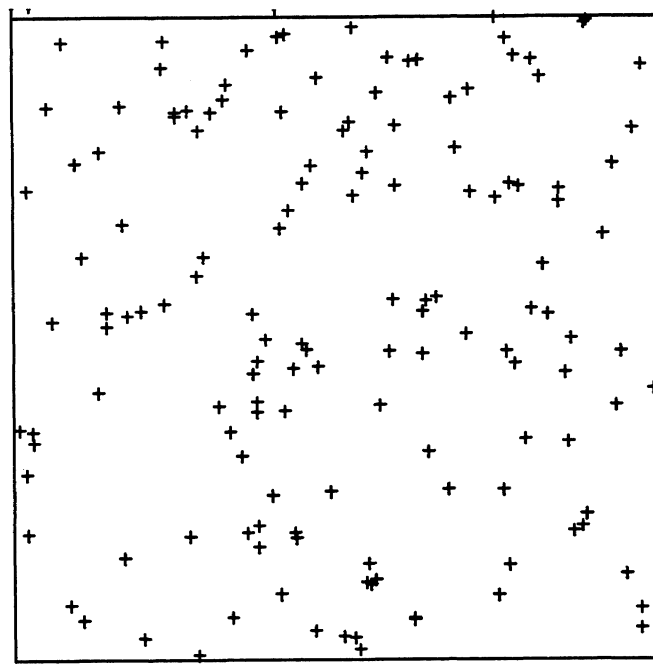

Fig. 5(a). Sample pattern of the simulation for PF I, where $\alpha=5$, $\beta=2$. This is a part of the pattern of 500 points.

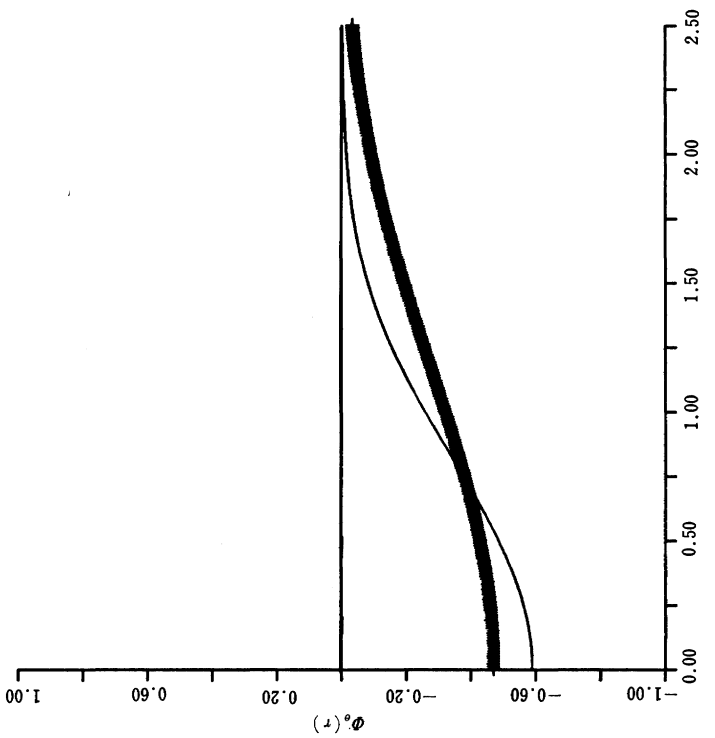

Fig. 6(b). Potential curves for true parameters corresponding to the model PF II (presented by a solid line) and for estimated parameters corresponding to PF II (presented by a series of the symbol +).

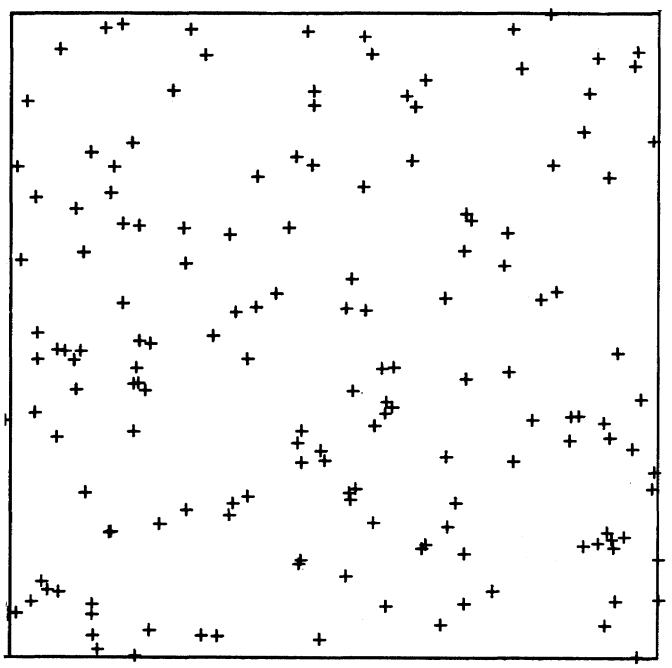

Fig. 6(a). Sample pattern of the simulation for PF II, where $\alpha = 1.8$, $\beta = 1.0$. The pattern presented here is a part of the pattern of 500 points.

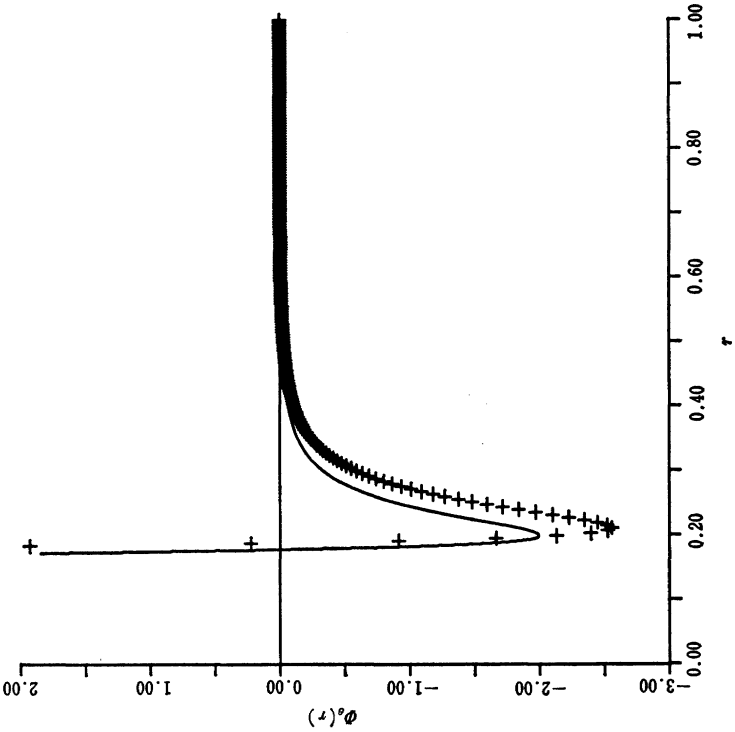

Fig. 7(b). The true and the estimated potential curves. The solid line represents a curve of PF III for true parameters and the series of symbol + represent a curve of the same model for estimated parameters.

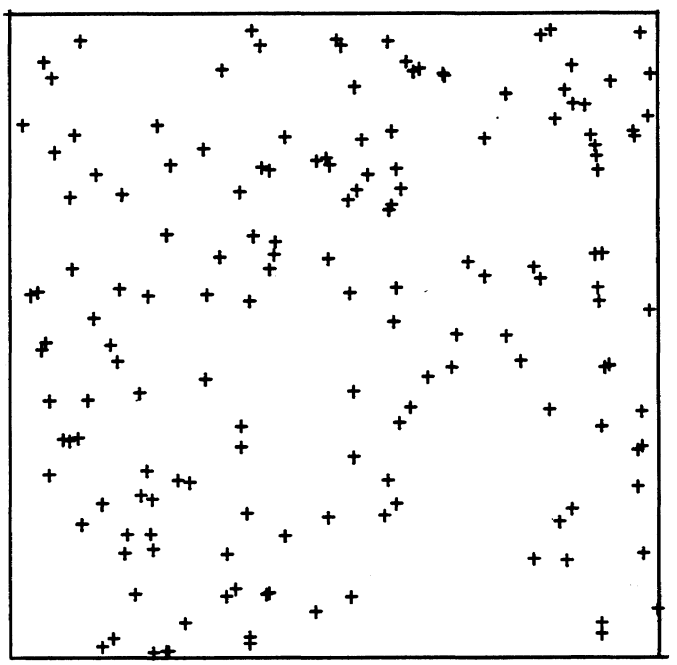

Fig. 7(a). Sample pattern of the simulation for PF III, where $\alpha=4$, $\beta=2$, $\sigma=0.2$. This is a part of the pattern of 500 points.

corresponding potential functions, respectively PF I, II and III, are illustrated in Figs. 5 (b)–7 (b). Then all models PF I–III were fitted to all of the three sets of data. Each AIC value is recorded in Table 5, which shows that we can discriminate the true potential function from others reasonably well with a moderate sample size. The best fitted potential functions are also given in Figs. 5 (b)–7 (b).

8. Some examples of spatial patterns

Next, all of the models PF I–III were fitted to some real data. Figures 8 and 9 (a) show respectively the natural stands of saplings ($N=65$) and of seedlings and saplings ($N=204$) of Japanese black pine, *Pinus Thunbergii*. These data were obtained by Numata [13], [14]. The AIC values for these data for respective models are given in Tables 6 and 7. As for the first data, the AIC values in Table 6 indicate that the Poisson model is better than any of the available models in this paper. This result coincides with the original analysis by Numata [13] through quadrat method and with other non-parametric analyses by Bartlett [2], Besag and Diggle [4] and Diggle [5]. For the second data the model PF III was selected to be the best with AIC$=-17.9$ (see Table 7), whose estimated potential function is illustrated in Fig. 9 (b).

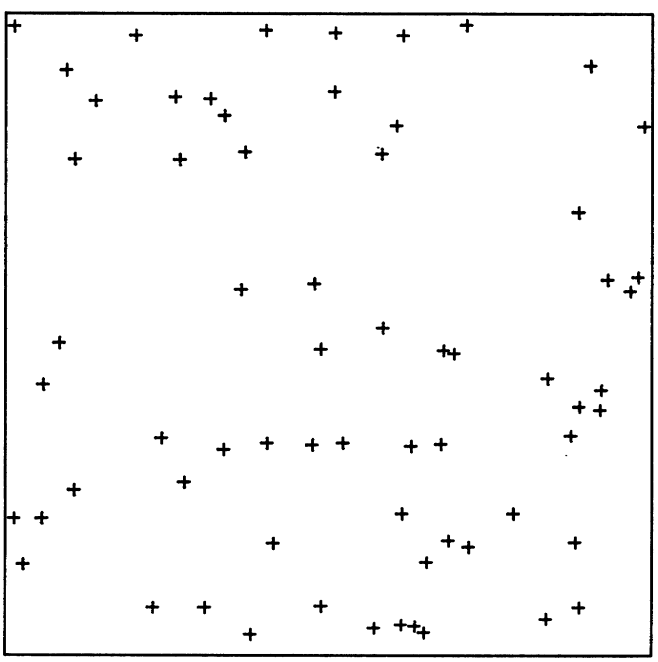

Fig. 8. The map of saplings of Japanese black pine, *Pinus Thunbergii* ($N=65$ in a 5 metres$\times$5 metres area, Numata, [13]).

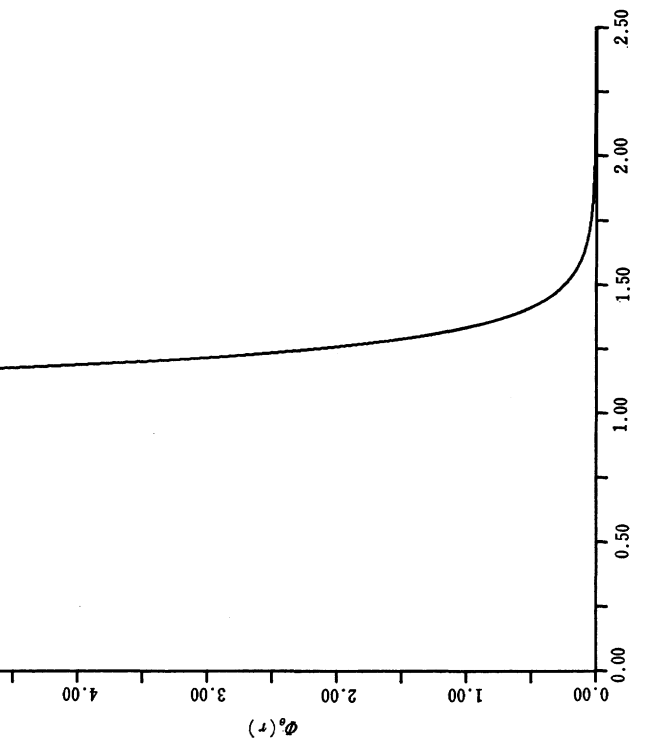

Fig. 9 (b). The curve of the potential PF III which was estimated from the data (a). The unit of length here corresponds to 0.05 metres.

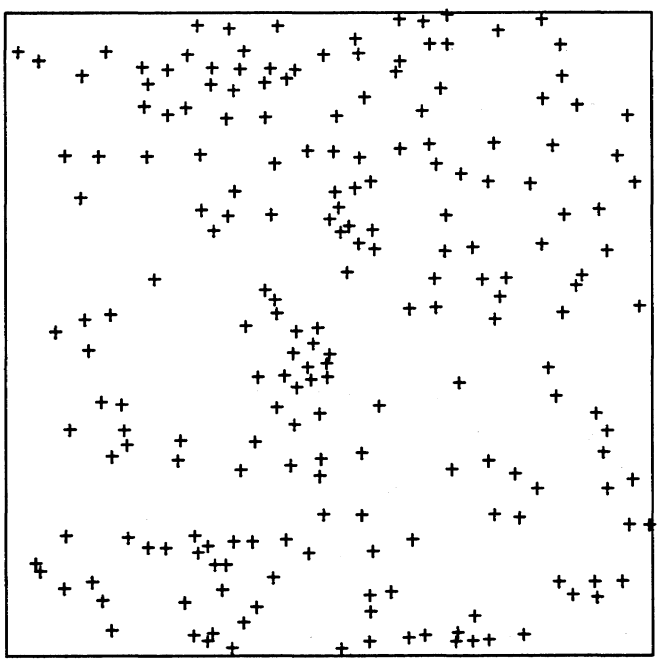

Fig. 9 (a). The map of seedlings and saplings of Japanese black pine (N=204 in a 10 metres×10 metres area, Numata [14]).

Table 6. Model fitting to Japanese black pine saplings (N=65, Numata [13])

Model	Estimated parameters			AIC	Number of parameters
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$		k
PF I	0.00	0.59	—	2.8	2
	0.0 (fixed)	0.59	—	0.8	1
PF II	0.00	0.59	—	2.8	2
	0.0 (fixed)	0.59	—	0.8	1
PF III	1.71	0.89	1.73	2.1	3

Table 7. Model fitting to Japanese black pine seedlings and saplings (N=204, Numata [14])

Model	Estimated parameters			AIC	Number of parameters
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$		k
PF I	0.00	0.42	—	−11.9	2
	0.0 (fixed)	0.42	—	−13.9	1
PF II	0.00	0.42	—	−11.9	2
	0.0 (fixed)	0.42	—	−13.9	1
PF III	0.24	3.42	1.22	−15.9	3
	0.0 (fixed)	2.20	1.25	−17.9	2

This indicates that something like hard-core (or self-inhibiting) interaction is exerted by the seedlings and saplings. In other words, it means that the effect of spacing between pine trees is found, whereas Numata [14] himself suggests a contagious tendency after quadrat analysis. In our opinion, this does not indicate contradiction between the two results, but is due to the difference of scale between the two analyses. That is to say, our approach is restricted to be local, whereas Numata's is rather global.

Next, we analyzed a nesting pattern of Gray gull, *Larus modestus*, in a 100 metres square area (Howell, Araya and Millie [9]). The distribution of nests (N=110) is shown in Fig. 10 (a). Table 8 gives the

Table 8. Model fitting to Gray Gull's nesting pattern (N=110, Howell, Araya and Millie [9])

Model	Estimated parameters			AIC	Number of parameters
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$		k
PF I	0.00	0.14	—	−8.6	2
	0.0 (fixed)	0.14	—	−10.6	1
PF II	0.00	0.14	—	−8.6	2
	0.0 (fixed)	0.14	—	−10.6	1
PF III	0.00	3.40	2.01	−11.5	3
	0.0 (fixed)	1.99	2.10	−13.5	2

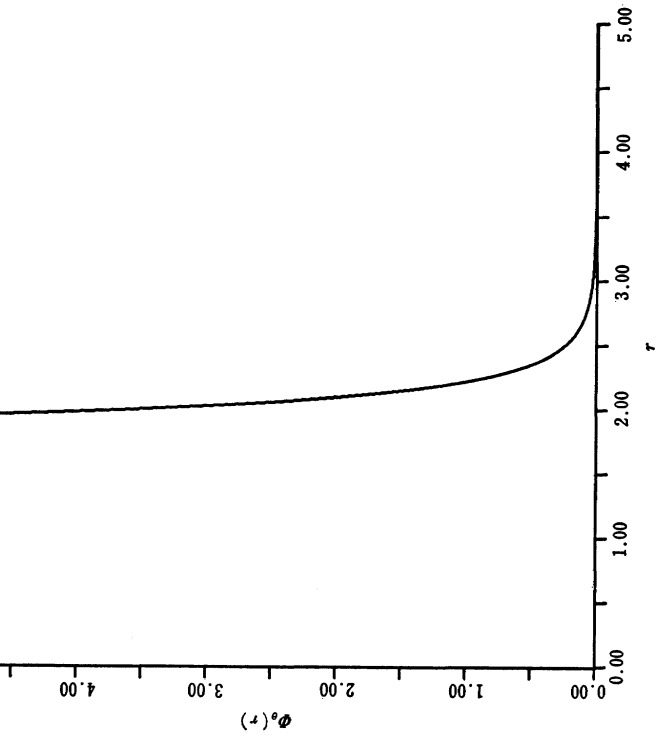

Fig. 10 (b). The curve of the potential PF III which was estimated from the data (a). The unit of length here corresponds to 1 metre.

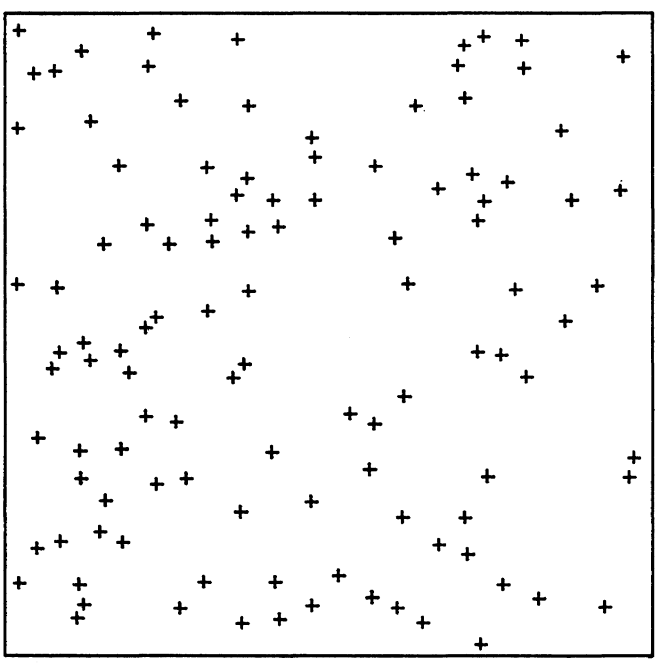

Fig. 10 (a). The map of the nests of the Gray Gull, *Larus modestus* (N=110 in a 100 metres × 100 metres area, Howell, Araya and Millie [9]).

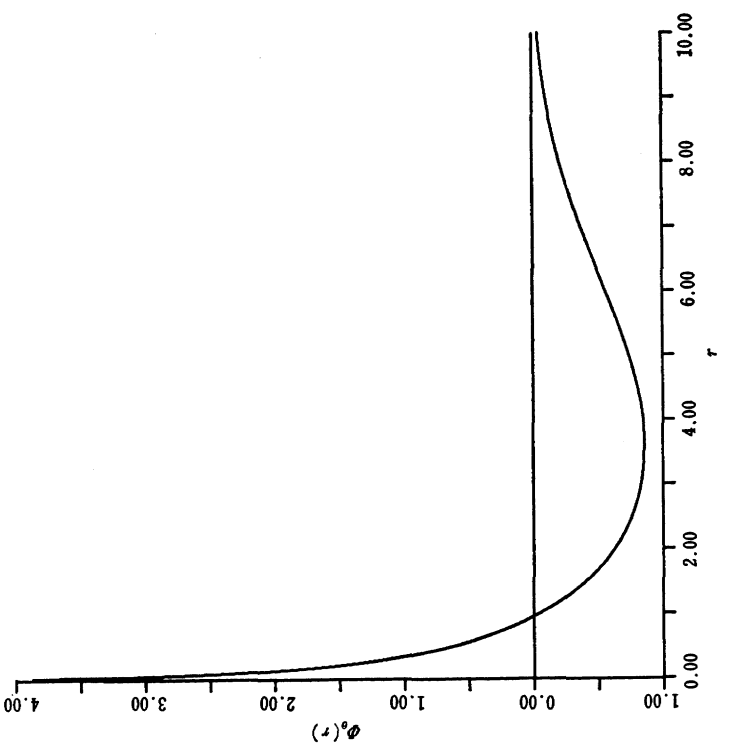

Fig. 11 (b). The curve of the potential PF I which was estimated from the data (a). The unit of length here corresponds to 1.11 km.

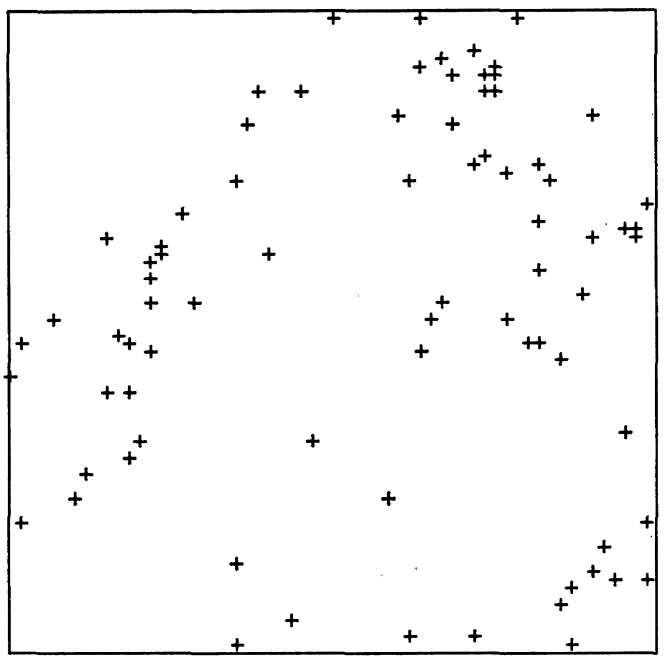

Fig. 11 (a). The map of aftershock epicentres of the Off Tokachi earthquake within the region from 40°N to 41°N and from 142°41′E to 144°E. This is nearly a 111 km × 111 km square.

AIC values of this data for respective models. In this case, as before, PF III was selected as the best model with AIC$=-13.5$. The estimated potential function is indicated in Fig. 10 (b). This result suggests that some spacing mechanisms also works among nests due to the territorial aggression of individuals against intruders and that no indication of attractive potential can be found, whereas Howell et al. [9] concludes that the distribution of nests does not differ from a Poisson point pattern.

Finally we consider a spatial pattern of aftershock hypocentres of the Off Tokachi earthquake in Japan which occurred on May 16, 1968. Based on the observation by the Japan Meteorological Agency, we restricted here the observed region to a certain area which contains the epicentre of the main shock, and then plotted the epicentres of earthquakes in Fig. 11 (a) during the six months' interval immediately after the main shock. The AIC values of the models PF I–III are shown in Table 9, from which PF I was chosen to be the best. The estimated potential function is given in Fig. 11 (b), which would indicate a spatial mechanism of stress release.

Table 9. Model fitting to the spatial aftershock pattern of the Off Tokachi earthquake (N=80)

Model	Estimated parameters			AIC	Number of parameters
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$		k
PF I	1.04	0.05	—	−16.4	2
PF II	2.72	0.03	—	−15.1	2
PF III	6.49	6.92	1.16	−5.3	3

9. Further remarks

It was assumed in the analysis of the preceding sections that the given spatial patterns are one of the gas-like configurations. Otherwise the approximations (9) and (11) do not hold. It is not easy to evaluate the accuracy of the approximate likelihood theoretically at this stage.

Other approximations of the partition function (4) exist for the solid-like configurations. Nevertheless the most difficult but quite interesting case is a liquid phase, where the partition function can usually be estimated by the simulation (see (8)). A Bayesian estimation of the potential function, therefore, would be useful as a general method for any of those phases. That is to say, for some prior $\pi(\theta)$,

$$\hat{\Phi}(r; \boldsymbol{X})=\int \Phi(r; \theta)L(\boldsymbol{X}; \theta)\pi(\theta)d\theta \Big/ \int L(\boldsymbol{X}; \theta)\pi(\theta)d\theta$$

$$\simeq \sum_{i=1}^{M} \Phi(r;\theta_i)\tilde{L}(\boldsymbol{X};\theta_i) \Big/ \sum_{i=1}^{M} \tilde{L}(\boldsymbol{X};\theta_i)$$

is an estimate of the potential function, where θ_i, $i=1,\cdots,M$, are to be generated according to the distribution $\pi(\theta)$ and then for each θ_i, $L(\boldsymbol{X};\theta_i)$ is replaced by the estimate $\tilde{L}(\boldsymbol{X};\theta_i)$ of (8) in Section 3. These remarks will be developed elsewhere.

Further details such as the asymptotic properties of the maximum likelihood estimates of (2) and (11), possible biases caused by edge effects or approximation of the likelihood we expect to study in the future. It is of course desirable that further suitable parametric potential models should be exploited according to the purpose of the analysis.

Acknowledgement

The authors would like to express their gratitudes to Mr. K. Katsura for his generous help in computing the data; to Dr. M. Hasegawa for kindly informing us of the Gray Gull's nesting pattern; and to the Japan Meteorological Agency for providing the earthquakes data. Finally Y. O. would like to thank to Mr. M. Baudin for stimulating discussions of the subject.

APPENDIX I. Derivation of Approximate Partition Function (9)

In the following, we derive (9) in a somewhat different way from that of Feynman [6].

Let us first rewrite (4), using (3), as follows:

$$\text{(A.1)} \qquad Z_N(\theta) = \int_{V^N} \prod_{n=1}^{N-1} \prod_{m=n+1}^{N} \exp\{-\Phi_\theta(X_{nm})\} dX_1 \cdots dX_N$$
$$= \int_{V^N} \prod_{n=1}^{N-1} \prod_{m=n+1}^{N} \{1+f(X_{nm})\} dX_1 \cdots dX_N ,$$

where $X_{nm}=|X_n-X_m|$ and

$$f(X_{nm}) = \exp\{-\Phi_\theta(X_{nm})\} - 1 ,$$

which is called the Mayer function. Note that $f(r)\equiv 0$ for the ideal gas, i.e. the case $\Phi_\theta(r)=0$ for $r>0$, and that $f(r)\to 0$ as $r\to\infty$ for physically reasonable potentials, i.e. $\Phi_\theta(r)\to 0$ as $r\to\infty$. If the influential range of the potential $\Phi_\theta(r)$ is short enough, $f(r)$ tends to zero rapidly as r increases. This fact will make it reasonable to expand the product in the integrand of (A.1) in terms of the Mayer function.

For example, if $N=3$, the expansion of the integrand of (A.1) has the form:

$$1+f(X_{12})+f(X_{13})+f(X_{23})+f(X_{12})f(X_{13})+f(X_{12})f(X_{23}) \\ +f(X_{13})f(X_{23})+f(X_{12})f(X_{13})f(X_{23}) \,. \tag{A.2}$$

Each term in (A.2) constitutes a "cluster" and such an expansion is called a cluster expansion. The first term is the ideal gas term and contributes a factor $|V|^3$ to the partition function. Each of the following three terms such as $f(X_{12})$ is the second cluster. These terms give to the integral (A.1) a contribution of $-3|V|^2a(\theta)$, where

$$a(\theta)=\int_V [1-\exp\{-\Phi_s(X_{12})\}]dX_1 \,, \tag{A.3}$$

which is called the second cluster integral. If the potential range is small enough compared with the mean volume per point, as already assumed in the text, the range of the integration in (A.3) can be safely extended to infinity, and then (10) results. The second cluster integral represents a contribution of two body collision to the partition function. The last four terms in (A.2) represent the third clusters. It is easy to show that, among them, three terms such as $f(X_{12})f(X_{13})$ contribute to the integral (A.1) $3|V|\{a(\theta)\}^2$. The integral of the final term, however, cannot be reduced to a multiple of the second cluster integrals. Such an integral is called an irreducible integral. (Note that $a(\theta)$ is the simplest irreducible integral.) In this case, it represents a contribution of three body collision. In other words, the value of $f(X_{12})\cdot f(X_{13})f(X_{23})$ is large only when the points 1, 2 and 3 are within the potential range of each other at the same time.

Here, we introduce an approximation. Namely, we replace the integral of $f(X_{12})f(X_{13})f(X_{23})$ by $\left\{\int_V f(X_{12})dX_1\right\}^3=-\{a(\theta)\}^3$. It means that the (irreducible) cluster of three points 1, 2 and 3 is assumed as if it is composed of three independent second clusters. Thus, within the framework of this approximation, (A.1) turns out, for $N=3$, to be of the form

$$|V|^3-3|V|^2a(\theta)+3|V|\{a(\theta)\}^2-\{a(\theta)\}^3=|V|^3\left(1-\frac{a(\theta)}{|V|}\right)^{3(3-1)/2} \,.$$

For a general N, by introducing a similar approximation to every irreducible integrals are shown above, (A.1) gives

$$Z_N(\theta)=|V|^N\left(1-\frac{a(\theta)}{|V|}\right)^{N(N-1)/2} \,,$$

which is equivalent to the second cluster approximation (9) of the partition function as given in the text.

APPENDIX II. Derivation of $a(\alpha, \beta, \sigma)$

We show here a derivation of the second cluster integral $a(\alpha, \beta, \sigma)$ for the Lennard-Jones type potential PF III of the form

$$\text{(A.4)} \qquad \Phi_\theta(r)=\beta\left(\frac{\sigma}{r}\right)^{2m}-\alpha\left(\frac{\sigma}{r}\right)^{m}, \qquad m>2,\ \theta=(\alpha, \beta, \sigma)\,.$$

By integration by parts, (10) becomes for R^2

$$a(\theta)=-\pi\int_0^\infty r^2\frac{d\Phi_\theta(r)}{dr}\exp\{-\Phi_\theta(r)\}dr\,.$$

Substituting (A.4) in the above expression, we obtain

$$a(\theta)=-m\pi\alpha\sigma^2\int_0^\infty x^{m-3}\left(1-\frac{2\beta}{\alpha}x^m\right)\exp(-\beta x^{2m}+\alpha x^m)dx\,, \qquad (x\equiv\sigma/r)\,.$$

Expanding the factor $\exp(\alpha x^m)$ in the integrand into a series, we obtain an expression,

$$\text{(A.5)} \qquad a(\theta)=-m\pi\alpha\sigma^2\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\int_0^\infty x^{mk+m-3}\left(1-\frac{2\beta}{\alpha}x^m\right)\exp(-\beta x^{2m})dx\,.$$

This reminds us of an integral formula

$$\text{(A.6)} \qquad \int_0^\infty x^{b-1}e^{-ax^c}dx=\Gamma\left(\frac{b}{c}+1\right)\Big/a^{b/c}b\,, \qquad \text{for } a, b, c>0\,,$$

where $\Gamma(\cdot)$ is the gamma function. Applying (A.6) to (A.5), it follows that

$$a(\theta)=-m\pi\alpha\sigma^2\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\left\{\frac{\Gamma((mk+3m-2)/2m)}{mk+m-2}\beta^{-(mk+m-2)/2m}\right.$$
$$\left.-\frac{2\beta}{\alpha}\frac{\Gamma((mk+4m-2)/2m)}{mk+2m-2}\beta^{-(mk+2m-2)/2m}\right\}\,.$$

If we use the property of gamma function, i.e. $\Gamma(z)=(z-1)\Gamma(z-1)=(z-1)(z-2)\Gamma(z-2)$ for respective terms in the parenthesis, we finally obtain a more simplified form

$$a(\theta)=-\frac{\pi}{m}\beta^{1/m}\sigma^2\sum_{k=0}^{\infty}\frac{1}{k!}\Gamma\left(\frac{mk-2}{2m}\right)\alpha^k\beta^{-k/2}\,, \qquad m>2\,.$$

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Akaike, H. (1977). On entropy maximization principle, *Applications of Statistics* (Ed. P. R. Krishnaiah), North-Holland, Amsterdam, 27-41.

[2] Bartlett, M. S. (1964). The spectral analysis of two-dimensional point processes, *Biometrika*, **44**, 299-311.

[3] Baudin, M. (1980). Likelihood and nearest neighbor distance properties of multidimensional Poisson cluster processes, submitted to *J. Appl. Prob.*

[4] Besag, J. and Diggle, P. J. (1977). Simple Monte Carlo tests for spatial pattern, *Appl. Statist.*, **26**, 327-333.

[5] Diggle, P. J. (1979). On parameter estimation and goodness-of-fit testing for spatial patterns, *Biometrics*, **35**, 87-101.

[6] Feynman, R. P. (1972). *Statistical Mechanics: A Set of Lectures*, Benjamin, Reading.

[7] Fisher, L. (1972). A survey of the mathematical theory of multidimensional point processes, *Stochastic Point Processes: statistical analysis, theory and applications* (ed. P. A. W. Lewis), Wiley, New York, 468-513.

[8] Hasegawa, M. and Tanemura, M. (1978). Mathematical models on spatial patterns of territories, *Proceedings of the international symposium on mathematical topics in biology*, Kyoto, Japan, Sept. 11-12, 1978, 39-48.

[9] Howell, T. R., Araya, B. and Millie, W. R. (1974). Breeding biology of the Gray Gull, *Larus modestus*, *Univ. Calif. Publ. Zool*, **104**, 1-57.

[10] Matérn, B. (1960). Spatial variation, *Meddelanden fran Statens Skogsforskningsinstitut*, **49**, No. 5, 1-144.

[11] Mayer, J. E. and Mayer, M. G. (1940). *Statistical Mechanics*, John Wiley, New York.

[12] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087-1092.

[13] Numata, M. (1961). Forest vegetation in the vicinity of Choshi-coastal flora and vegetation at Choshi, Chiba Prefecture, IV (in Japanese), *Bull. Choshi Marine Laboratory*, No. 3, Chiba University, 28-48.

[14] Numata, M. (1964). Forest vegetation, particularly pine stands in the vicinity of Choshi-flora and vegetation at Choshi, Chiba Prefecture, VI (in Japanese), *Bull. Choshi Marine Laboratory*, No. 6, Chiba University, 27-37.

[15] Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes, *Ann. Inst. Statist. Math.*, **30**, A, 243-261.

[16] Ogata, Y. (1980). Maximum likelihood estimates of incorrect Markov models for time series and the derivation of AIC, *J. Appl. Prob.*, **17**, 59-72.

[17] Ogata, Y. (1981). On Lewis' simulation method for point processes, *IEEE Trans. Inform. Theory*, IT-27, **1**, 23-31.

[18] Ripley, B. D. (1977). Modelling spatial patterns (with discussion), *J. R. Statist. Soc.*, B, **39**, 172-212.

[19] Sakamoto, Y. and Akaike, H. (1978). Analysis of cross classified data by AIC, *Ann. Inst. Statist. Math.*, **30**, B, 185-197.

[20] Tanemura, M. and Hasegawa, M. (1980). Geometrical models of territory. I. Models for synchronous and asynchronous settlement of territories, *J. Theor. Biol.*, **82**, 477-496.

[21] Vere-Jones, D. (1970). Stochastic models for earthquake occurrences (with discussion), *J. R. Statist. Soc.*, B, **32**, 1-62.

[22] Vere-Jones, D. (1978). Space time correlations for microearthquakes—a pilot study, *Suppl. Adv. Appl. Prob.*, **10**, 73-87.

[23] Wood, W. W. (1968). Monte Carlo studies of simple liquid models, *Physics of Simple Liquids* (eds. H. N. V. Temperley, J. S. Rowlinson and G. S. Rushbrooke), Chap. 5, North-Holland, Amsterdam, 115-230.