ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 194, 2005

*Jerzy Korzeniewski**

ANALYSIS OF POINT PROCESSES OBSERVED WITH NOISE WITH APPLICATIONAL EXAMPLE

Abstract

An example of the application of point processes observed with noise are aerial photographs of forests with the aim of estimating the actual number of trees on a given area. Lund and Rudemo (2000) proposed a model useful in this context, basing on the number of "trees candidates" visible on the photograph. The parameters of conditional likelihood function were estimated taking into account such variations of noise as points thinning, points displacement and appearing of extra ghost points. The approach proposed does not solve the problem of the estimation of the actual number of trees.

In this paper a new algorithm to estimate directly the number of actual trees is proposed. The only assumption on which the new measure depends is the natural assumption about forest density being locally constant. The results achieved with the help of the new measure may be assessed as interesting.

Key words: point process, maximum likelihood method, noise, incomplete observation, image data, computer algorithm.

I. INTRODUCTION

Figure 1 depicts a map of a part of a forest with 206 small circles and 171 dots. The dots were found on the basis of an aerial photograph of this part of the forest with the help of a template constructed by Larsen and Rudemo (1998) and they represent candidates for trees (Norway spruce). Basically, the idea of the template construction is to choose pixels from black-and-white photograph the ellipse neighbourhood of which gives suitably high correlation between the shades of the grey colour of the neighbourhood pixels and the shades of grey of the ideal template. The dots represent pixels for which the correlation was high enough. The circles represent true

* Ph.D., Chair of Statistical Methods, University of Łódź.

trees found in the same region of the forest by manual inspection. One pixel on the photograph corresponds to ground area of 0.15×0.15 m^2.

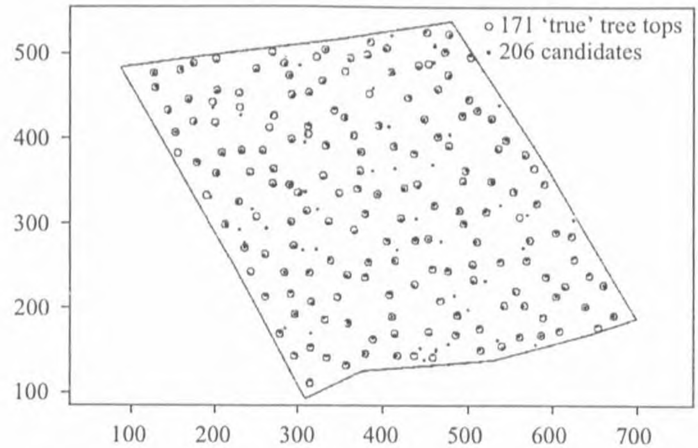

Figure 1. True tree tops (dots) and candidates for trees (circles)

The statistician's task is to investigate all the phenomena influencing the picture that we arrive at in connection with the true number of trees that grow on the area being photographed and, if possible, to estimate the number of trees from the number of candidates. Lund and Rudemo (2000) propose an approach the idea of which is to treat trees candidates as realisations of a point process contaminated by noise of different kinds. The authors found the conditional likelihood function and analysed its behaviour to estimate some interesting parameters. This approach is briefly outlined in the next section. However, the mentioned approach, though mathematically elegant does not answer the question that is most interesting to forest men and ecologists i.e. what is the approximate number of true trees? In the third section we propose a new measure which tries to tackle this problem directly.

II. POINT PROCESSES WITH NOISE

Lund and Rudemo proposed to consider two point processes, one of which i.e. Y is an imperfect observation of the other process X. We assume that X and Y are point processes on a subset A of d-dimensional Euclidean space R^d with a finite number of points, $X = \{X_i : i \in M\}$, $M = \{1, ..., m\}$, $Y = \{Y_j : j \in N\}$, $N = \{1, ..., n\}$. Assume further that A is bounded with a positive d-dimensional volume $|A_d|$. Suppose that Y is generated from the X process by the following disturbance mechanisms.

1. **Thinning.** Each point X_i, for $i \in M$, is thinned with probability $1-p(X_i)$ and retained with probability $p(X_i)$. If an X point is thinned, then there will be any corresponding Y point. Thinings are assumed to be independent for different points.

2. **Displacement.** For each remaining point X_i a corresponding Y_i point is generated by displacement to a position with probability density $k(\cdot|X_i)$ with respect to Lebesgue measure on R^d Given X, the displacements of different points are independent, mutually and of the thinnings.

3. **Censoring.** The displaced points are observed if they are within the observation region A; otherwise they are censored and not observed. Thus censoring of an unthinned point generated by X_i occurs with probability $\int_{A^c} k(y|X_i)dy$. Here A^c denotes the complement of the set A.

4. **Superposition** of ghost points. In addition to the points generated as described above we have superposition of extra "ghost" points. These points are assumed to arise from a Poisson process on A with intensity $(\cdot|X_i)$ where X, as above, denotes the entire X-process.

The initial and basic result is the following theorem which gives formula for the conditional probability of a point process Y given another process X.

Theorem 1. Let X and Y be two finite point processes as specified above, on a bounded set A. Suppose that $g(y|X)$ and $k(y|X_i)$ for $i \in M$, are continuous functions of $y \in A$. Then the conditional likelihood of Y given X is

$$L(Y|X) = \exp\{|A|_d - \int_A g(y|X)dy\} \sum_{\substack{M_1 \subseteq M \\ N_1 \subseteq N \\ |M_1| = |N_1|}} \sum_{\pi \in P(M_1, N_1)} L_1 L_2 L_3, \qquad (1)$$

where

$$L_1 = \prod_{i \in M_1} p(X_i) k(Y_{\pi(i)}|X_i),$$

$$L_2 = \prod_{i \in M|M_1} \{p(X_{ii}) \int_{A^c} k(y|X_i)dy + 1 - p(X_i)\}$$

$$L_3 = \prod_{j \in N|N_1} g(Y_j|X),$$

and the reference measure corresponds to the Poisson process on A with intensity 1. Looking at this formula we can see that all possible noise "combinations" were taken into account, because the symbol $P(M_1, N_1)$ denotes all possible one-to-one mappings from M_1 to N_1.

The formula given in Theorem 1 is too complicated to analyse in order to find its maximums, therefore, we can proceed with a couple of simplifications. First simplification is that of the homogenous intensity of the ghost points i.e. $g(\cdot|X) = \lambda$. Then the likelihood function (1) simplifies to

$$L(Y|X) = \sum_{\substack{M_1 \in M \\ N_1 \in N \\ |M_1| = |N_1|}} \sum_{\pi \in \rho(M_1, N_1)} T(M_1, N_1, \pi) \tag{2}$$

with the summed terms given by

$$T(M_1, N_1, \pi) = p^{|M_1|} \lambda^{|N|N_1|} \exp\{(1-\lambda)|A_d|\} \left\{ \prod_{i \in M_1} k(Y_{\pi(i)}|X_i) \right\}$$
$$\times \prod_{i \in M|M_1} \{p \textstyle\int_{A^c} k(y|X_i)dy + 1 - p\}. \tag{3}$$

The second simplification refers to the fact that nearly all X_i points are so far from the boundary of the observation region A that we can safely assume that they are not censored i.e.

$$\int_{A^c} k(y|X_i)dy = 0. \tag{4}$$

In the first approximation of (3) we assume that (4) holds for all X-points, and thus replace (3) by

$$T(M_1, N_1, \pi) = p^{|M_1|}(1-p)^{|M \setminus M_1|} \lambda^{|N \setminus N_1|} \exp\{(1-\lambda)|A_d|\} \left\{ \prod_{i \in M_1} k(Y_{\pi(i)}|X_i) \right\}. \tag{5}$$

For $s = (M_1, N_1, \pi)$ note that (5) may be considered as a function of the parameter vector

$$\boldsymbol{\theta} = (p, \lambda, \mu_1, \mu_2, \sigma_1, \sigma_2, p) \tag{6}$$

and let us denote that it is maximised by the following vector

$$\hat{\boldsymbol{\theta}}(s) = (\hat{p}, \hat{\lambda}, \hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{p})$$

where

$$\hat{p} = |M_1|/|M|, \quad \hat{\lambda} = |N|N_1|/|A_d|, \quad \text{and} \quad (\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{p})$$

are the standard maximum likelihood estimates of the parameters in a two-dimensional normal distribution based on the sample $(Y_{\pi(i)} - X_i,\ i \in M_1)$.

Search for the maximum is still a problematic task and we cope with it by considering the function value in all possible "neighbours of a considered state". For state (M_1, N_1, π) we define its neighbour (M_1, N_1, π') if it can be obtained from (M_1, N_1, π) in one of the following five ways.

1. Addition of a pair of X-and Y-points: $M'_1 = M_1 \cup \{i'\}$ where $i' \in M | M_1$, $N'_1 = N_1 \cup \{j'\}$, where $j' \in N \setminus N_1, \pi'(i) = \pi(i)$, for $i \in M_1$ and $\pi'(i') = j'$. The number of such neighbours is $|M \setminus M_1||N \setminus N_1|$.

2. Removal of a pair of X-and Y-points: $M'_1 = M_1 \setminus \{i'\}$, where $i' \in M_1$, $N'_1 = N_1 \setminus \{j'\}$, where $j' \in N_1, \pi'(\mathrm{i}) = \pi(i)$, for $i \in M'_1$ and $\pi(i') = j'$. This can be done in $|M_1| = |N_1|$ ways.

3. Swapping an X-point: $M'_1 = (M_1 \setminus \{i'\}) \cup \{i''\}$, where $i' \in M_1$ and $i'' \in M \setminus M_1, N'_1 = N_1 \pi'(i) = \pi(i)$, for $i \in M_1 \setminus \{i'\}$ and $\pi'(i'') = \pi(i')$. There are $|M_1||M \setminus M_1|$ such neighbours.

4. Swapping a Y-point: $M'_1, N'_1 = (N_1 \setminus \{j'\}) \cup \{j''\}$, where $j' \in N_1$ and $j'' \in N \setminus N_1$, $\pi'(i) = \pi(i)$, for $i \cup M_1 \setminus \{i'\}$, where $\pi(i') = j'$ and $\pi'(i') = j''$. Swapping a Y-point can be done in $|N_1||N \setminus N_1|$ ways.

5. Exchange among two pairs: $M'_1 = M_1, N'_1 = N'_1 = N_1, \pi'(i) = \pi(i)$, for $i \in M_1 \setminus \{i', i''\}$, where $i' \in M_1$ and $i'' \in M_1$, for $i' \neq i'', \pi'(i') = \pi(i')$ and $\pi'(i'') = \pi(i')$. The number of such neighbours is $|M_1|(M_1|-1)/2$.

Now we can search for the maximum of function (2) by considering iteratively its value on all possible states each of which is a neighbour of some other state of the previous iteration. In this way Lund and Rudemo found the maximum of the conditional likelihood function and tried to investigate the behaviour of this function. This approach, though quite attractive from the mathematical point of view, does not solve the main problem of estimating the extent of forest depletion on the basis of the possessed noise version i.e. the Y process realisation.

III. DIRECT ASSESSMENT OF FOREST DEPLETION

Assessing intuitively the number of true trees one feels that the position of false trees is not independent of the position of true trees (obviously it cannot be, e.g. both cannot be located in the same positions) or, in other words, the average distance between false tree and the nearest tree (either false or true) is smaller than the same distance for the true trees only. This is probably caused by the fact that false spruces for some reasons happen to be located close to true spruces but are not spruces. Therefore, the quadratic dependence between the forest area and the number of trees comprised by it (if we assume uniform forest density) should be violated for area chosen in some way.

The area on which we will require the quadratic dependence between it and the number of spruces is constructed in the following way. For all pixels representing either true or false spruces we consider circles of the same radius r with centres at the pixels. We let r grow and for each value of r e.g. positive integer we calculate the number of trees which fall within at least one circle. The function describing the dependence of the number of trees on r should be a quadratic one. We cannot go too far with the radius length, because for big values of r the circles overlap one another and thus some trees would be counted twice. In Table 1 we calculated the numbers of trees comprised by circles with radius length equal 20 pixels at the greatest. The particular lengths ending with 0.1 were chosen so as to make the series of the number of trees as smooth as possible. Actually, no matter what the ending of the successive radius values is, the conclusions are exactly the same but for the ending chosen i.e. 0.1 the series are without big "jumps". Next, we calculate the coefficients of determination for the least squares quadratic regression for the number of trees comprised by all circles in dependence on r. The coefficients presented in Table 1 were calculated for the regression lines based on 10 successive values of r. For the reasons mentioned above we cannot go to far with radius length and if we choose other number of observations for quadratic regression e.g. 8, 9, 11 or 12, the results are almost identical (r^2 is the same up to 0.01).

Table 1. Numbers of spruces (true and all) within successive circles and determination coefficient for quadratic regression

Radius length r	Number of trees within circle of radius r		Determination Coefficient	
	all trees	true trees	all trees	true trees
4.1	5	1	0.902	0.803
5.1	9	2	0.930	0.873
6.1	10	2	0.952	0.904
7.1	16	2	0.963	0.914
8.1	24	5	0.975	0.943
9.1	34	6	0.985	0.966
10.1	50	13	0.990	0.979
11.1	69	23	0.990	0.983
12.1	92	36	0.991	0.987
13.1	109	42	0.994	0.992
14.1	140	64		
15.1	169	84		
16.1	197	104		
17.1	234	126		
18.1	261	145		
19.1	297	171		

Source: Author's investigation.

Looking at the numbers presented in Table 1 one can see that the coefficient of determination for true trees only reach the ideal value (i.e. equal to one) "later" than the coefficient for all trees. Later, obviously, refers to bigger values of r and, what follows, to greater number of trees comprised by all circles. To some extent this phenomenon may seem natural because the forest density of true trees is smaller. In our opinion, however, this discrepancy is also caused by particular location of false trees. The value of determination coefficient for all trees which may be appropriate for the "threshold point" should be the one from interval (0.985, 0.990), preferably closer to 0.985, because later the differences between the values of determination coefficient become very small (even less than 0.001), too small to be considered informative signs of the dependence of determination coefficient on radius length. If we assume 0.985 to be the threshold point we can see that for true trees we would have to throw away only 6 trees out of 171, which is a tolerable mistake (3.5%) and for all trees we would have to throw away 34 trees, which is very close to the ideal number of 35 false trees that should be rejected. This method of assessing the number of true trees is not very precise because it leaves some place for deliberate threshold point choice, but it is very simple and free of any assumptions apart from uniform forest density.

REFERENCES

Larsen M., Rudemo M. (1998), Optimizing templates for finding trees in aerial photographs, *Pattern Recognition Letters*, **19**, 1153–1162.

Lund J., Rudemo M. (2000), Models for point processes observed with noise, *Biometrika*, **87**, 235–249.

Jerzy Korzeniewski

ANALIZA PROCESÓW PUNKTOWYCH Z SZUMEM Z PRZYKŁADEM APLIKACYJNYM

Streszczenie

Przykładem zastosowania procesów punktowych obserwowanych wraz z szumem są zdjęcia lotnicze lasów robione w celu oszacowania ubytków leśnych na danym terenie. Rudemo i Lund (2000) zaproponowali model, który może być użyteczny w tym celu, wykorzystujący liczbę „kandydatów na drzewa" widocznych na zdjęciu. Parametry warunkowej funkcji wiarygodności zostały oszacowane z uwzględnieniem takich odmian szumu, jak

znikanie punktów, przemieszczanie się punktów oraz pojawianie się punktów fałszywych. To podejście nie rozwiązuje problemu szacowania faktycznej liczby drzew.

W artykule tym zaproponowano nowy algorytm, który bezpośrednio szacuje faktyczną liczbę prawdziwych drzew. Jedynym koniecznym założeniem jest założenie o stałej gęstości zalesienia na danym obszarze lasu. Rezultaty uzyskane za pomocą nowego algorytmu można ocenić jako interesujące.