Annales Academiæ Scientiarum Fennicæ Series A. I. Mathematica Volumen 13, 1988, 259–268

# A RANDOM FIELD APPROACH TO BITTERLICH SAMPLING

## A. Penttinen

#### 1. Introduction

Bitterlich's angle count method, also called "point sampling", is a widely used technique for estimating the proportional basal area of a forest stand. Basal area is the area of the cross section of a tree taken at the height of 1.3 m. Proportional basal area is the basal area per unit area of forest. This quantity is an important parameter in practical forestry, e.g., in the estimation of the volume of trees in a forest stand.

The idea behind Bitterlich's angle count method is simple: Take a sample point, then count the number of trees which the observer, standing in the sample point, sees in an angle less than  $2\alpha$ , where  $\alpha$  is a small angle fixed in advance. Assuming that the cross-sections are circular then number of counts  $\times \sin^2 \alpha$  is an unbiased estimator of the basal area per unit area of forest. This procedure is usually repeated in several sample points. Their locations follow some set of rules called sampling design, e.g., uniform random sampling or systematic sampling. A simple piece of equipment called the relascope is used for the measurement and, therefore, the method is also called relascope sampling.

Since Bitterlich [1] much has been published in the field of the angle count method. The majority of this work considers practical aspects in field work and empirical comparisons, see Kuusela [4] for an extensive study. Mathematical and statistical aspects associated with this method have been considered for example by Holgate [2], Matérn [5], [6] and Ord [8]. This theoretical work is mainly concentrated on the case where the trees are distributed according to a planar Poisson process and the diameters of the trees are independent. Sukwong et al. [14] have carried out a simulation study concerning the relative performance of Bitterlich's method under various conditions. In the literature the discussion and comparisons to other sampling methods are based on the variance of the proportional basal area estimator at a single sample point.

The purpose of the present study is to generalize the earlier theoretical results in two respects. First, we derive an explicit formula for the variance of the angle count estimator based on a single observation, under the condition that the pattern formed by trees is a second order stationary point process and the diameters of the

trees are independent. Second, the spatial covariance function of the observations is derived under the same condition. This allows us to calculate the asymptotic (in the sense defined by Matérn [7]) error variance under different sampling designs even when the observations are dependent. We also discuss the possibility of removing the assumption that the diameters are independent.

Our general approach to these problems is by means of an induced shot-noise random field. This is simply the random field  $\{y(\xi), \xi \in R^2\}$  where  $y(\xi)$  is the number of relascope counts at the sample point  $\xi$ , multiplied by the relascope factor  $\sin^2 \alpha$ . A general discussion of such spatial shot-noise processes can be found in Schmidt [10]. The covariance function of this random field  $\{y(\xi), \xi \in R^2\}$ provides us with all information needed in relascope sampling problems. Therefore, in Chapters 2 and 3, we concentrate on the link between the initial marked point process (describing the forest) and the corresponding shot-noise random field which we call here "relascope process". In the final Chapter 4 we return to the original forestry sampling problem.

#### 2. Preliminaries

Forests are described by a planar marked point process  $\Psi = \{[x_n; m_n]\}$  where the x's are the locations of the trees  $(\in \mathbb{R}^2)$  and the m's are the tree diameters. The distribution of  $\Psi$  is denoted by P. The unmarked point process of tree locations is denoted by  $\Phi = \{[x_n]\}$  which is assumed to be a locally finite and motion-invariant (i.e., stationary and isotropic) planar point process with intensity  $\lambda$ . Throughout this paper  $\Phi(B)$  stands for the number of trees in a Borel set B. The second reduced moment measure of  $\Phi$  is denoted by K. Interpretations of  $\lambda$ and K are:

- $\lambda$  is the expected number of trees in a unit area. We assume that  $0 < \lambda < \infty$ . Especially  $E\Phi(B) = \lambda \nu_2(B)$  where B is a Borel set in  $R^2$  and  $\nu_2$  is the Lebesgue measure on  $R^2$ .
- $\lambda \mathbf{K}(B)$  is the expected number of additional trees in a Borel set B given there is a tree in the origin.

In the case of an isotropic point process **K** is usually replaced by the K-function  $K(r) = \mathbf{K}(b(o,r))$  where b(o,r) is a sphere of radius r (r > 0) and centred at o, or by the g-function, called pair-correlation function in statistical physics,

$$g(r) = \frac{dK(r)}{dr}/(2\pi r), \quad r > 0,$$

if it exists. Other second order characteristics and their interpretations can be found in Ripley [9] and Stoyan et al. [13].

The first order characteristics of  $\Psi$  are the intensity  $\lambda$  and the mark (diameter) distribution M. An interpretation of the mark distribution is as follows: If L is an interval in  $R^+ = \{x \in R \mid x > 0\}$  then M(L) is the probability that the diameter of a "typical tree" belongs to L. The support of M is here typically  $R^+$  or a finite interval.

In order to give a second order characterization of  $\Psi$  we need the second reduced moment measure **K** (or K-function) described above and, in addition, the two-point mark distribution function  $M_{x_1x_2}$  defined by Stoyan [11]. An interpretation of  $M_{x_1x_2}$  is the following: if  $L_1$  and  $L_2$  are two Borel sets in  $R^+$ then  $M_{x_1x_2}(L_1, L_2)$  is the probability that the diameters of two trees at  $x_1$  and  $x_2$  belong to  $L_1$  and  $L_2$ , respectively, given that there are trees at  $x_1$  and  $x_2$ . We assume that  $M_{x_1x_2}$  depends on the distance  $r = |x_1 - x_2|$  only; hence the notation  $M_r$ . A special case is that the point process is independently marked. In that case  $M_r(L_1, L_2) = M(L_1)M(L_2)$ .

We are interested in the proportional basal area of forests. The basal area of a tree [x;m] is  $\pi m^2/4$  and, under the stationarity assumption, the proportional mean basal area of a forest is

$$\tilde{A} = \lambda \int_0^\infty (\pi m^2/4) M(dm)$$

which is assumed to be finite. Although  $\tilde{A}$  is the main parameter in this forestry problem we "mark" the trees by tree diameters for convenience.

## 3. The angle count random field

The relascope angle  $2\alpha$  is assumed to be fixed ( $\alpha$  small). Let us suppose that  $\xi \in \mathbb{R}^2$  is an arbitrary sample point. Then a tree  $[x;m] \in \Psi$  with diameter m will be counted through the relascope if and only if  $|x - \xi| < m/(2\sin\alpha)$  or, equivalently,  $x \in b(o, m/(2\sin\alpha)) + \xi$ . Here  $B + \xi$  stands for the transformed set  $\{y \in \mathbb{R}^2 \mid y = u + \xi, u \in B\}$ . This construction leads to the following definition:

**Definition.** The angle count random field generated by  $\Psi$  and  $\alpha$  is the planar stochastic process  $\{y(\xi), \xi \in \mathbb{R}^2\}$  where

(1) 
$$y(\xi) = \sin^2 \alpha \sum_{[x;m] \in \Psi} 1_{b(o,m/(2\sin \alpha)) + \xi}(x) .$$

Heuristically,  $y(\xi)$  is the number of relascope counts at  $\xi$  multiplied by the factor  $\sin^2 \alpha$ . The process  $\{y(\xi), \xi \in \mathbb{R}^2\}$  is also called simply "relascope process". Note that (1) is of the form

(2) 
$$y(\xi) = \sum_{[x;m] \in \Psi} \vartheta(x - \xi, m)$$

where  $\vartheta: R^2 \times R^+ \to R^+$  is called the response function. According to (2) the process  $\{y(\xi), \xi \in R^2\}$  is a special case of spatial shot-noise processes (shot-noise random fields). This connection is important because it allows us to apply results, e.g., moment formulae concerning shot-noise processes.

**Theorem 1.** (The basic theorem of relascope sampling.) Under conditions  $\lambda < \infty$  and  $\int m^2 M(dm) < \infty$ ,  $Ey(o) = \tilde{A}$ .

Theorem 1 means that the relascope counts multiplied by the factor  $\sin^2 \alpha$  is an unbiased estimator of the proportional basal area  $\tilde{A}$ . This result is due to Bitterlich [1]. The theorem can be proved by applying a simple geometric reasoning, see e.g. Ripley [9], or the Campbell theorem (see Stoyan et al. [13], p. 99). No assumption on independence in marking is needed in the proof of Theorem 1.

The performance of relascope sampling under different sampling designs can be deduced from the second order property of the process  $\{y(\xi), \xi \in \mathbb{R}^2\}$ . Therefore, we derive an expression for E[y(o)y(h)].

As an application of theorems of Campbell type (see Theorems 2.1 and 2.3 in Stoyan [11]) we obtain the following general formula:

$$\begin{split} E[y(o)y(h)] &= \int \sum_{[x_1,m_1] \in \psi} \vartheta(x_1,m_1) \sum_{[x_2,m_2] \in \psi} \vartheta(x_2 - h,m_2) P(d\psi) \\ &= \int \sum_{[x,m] \in \psi} \vartheta(x,m) \vartheta(x - h,m) P(d\psi) + \int \sum_{\substack{[x_1,m_1] \\ [x_2 + h,m_2] \in \psi}} \overset{\neq}{ \vartheta(x_1,m_1) \vartheta(x_2,m_2) P(d\psi) } \\ &= \lambda \int \vartheta(x,m) \vartheta(x - h,m) M(dm) \, dx + \\ &\quad + \lambda^2 \int \vartheta(x,m_1) \vartheta(x - h + u,m_2) M_u(d(m_1,m_2)) \mathbf{K}(du) \, dx \end{split}$$

for  $h \in \mathbb{R}^2$ , where  $\sum^{\neq}$  means summation over pairs of points  $[x_1; m_1]$ ,  $[x_2; m_2]$  of  $\Psi$  for which  $x_1 \neq x_2$ .

Let us assume now that  $\Psi = \{[x_n; m_n]\}$  is independently marked. Then, taking the specific form of the response function into account we obtain first

$$\begin{split} E[y(o)y(h)] &= \\ \lambda \,\sin^4 \alpha \, \int \mathbf{1}_{b(o,m/(2\sin\alpha))\cap(b(o,m/(2\sin\alpha))+h)}(x) M(dm) \, dx \\ &+ \lambda^2 \sin^4 \alpha \, \int \mathbf{1}_{b(o,m_1/(2\sin\alpha))\cap(b(o,m_2/(2\sin\alpha))+h-u)}(x) \mathbf{K}(du) M(dm_1) M(dm_2) \, dx \end{split}$$

A random field approach to Bitterlich sampling

$$= \lambda \sin^2 \alpha \int U(h, m/(2 \sin \alpha), m/(2 \sin \alpha)) M(dm) + \lambda^2 \sin^4 \alpha \iiint \bar{U}_r(h, m_1/(2 \sin \alpha), m_2/(2 \sin \alpha)) dK(r) M(dm_1) M(dm_2),$$

where  $U(h, r_1, r_2) = \nu_2[b(o, r_1) \cap b(h, r_2)]$  (the area of the intersection of two circles with radii  $r_1$  and  $r_2$  and having centres at o and h, respectively), and

$$\bar{U}_r(h, R_1, R_2) = \int_0^{2\pi} U(h - (r, \theta), R_1, R_2) \, d\theta / 2\pi$$

 $((r,\theta)$  is a point of  $R^2$  in polar coordinates). For numerical calculations the expression for  $U(h, r_1, r_2)$  is as follows: denote  $R = \max(r_1, r_2)$ ,  $r = \min(r_1, r_2)$  and  $x = |h|/2 - (R^2 - r^2)/(2|h|)$ . Then

$$\begin{split} &= \pi r^2, \quad 0 < |h| \le R - r, \\ &= R^2 \arccos((|h| - x)/R) + r^2 \arccos(x/r) - r|h| \sqrt{1 - (x/r)^2}, \ R - r < |h| < R + r, \\ &= 0, \qquad |h| > R + r. \end{split}$$

As a conclusion we obtain the following theorem:

**Theorem 2.** Under the assumption of independence in marking the variance and covariance function of the induced random field  $\{y(\xi), \xi \in \mathbb{R}^2\}$  are

$$\begin{aligned} \operatorname{Var}\{y(o)\} &= \tilde{A} \, \sin^2 \alpha \\ &+ \lambda^2 \sin^4 \alpha \int_0^\infty \int_0^\infty \int_0^\infty U(r, m_1/(2\sin\alpha), m_2/(2\sin\alpha)) dK(r) M(dm_1) M(dm_2) - \tilde{A}^2 \end{aligned}$$

and for h > 0,

$$\begin{split} c_y(h) &= \lambda \sin^4 \alpha \int_0^\infty U(h, m/(2\sin\alpha), m/(2\sin\alpha)) M(dm) \\ &+ \lambda^2 \sin^4 \alpha \int_0^\infty \int_0^\infty \int_0^\infty \bar{U}_r(h, m_1/(2\sin\alpha), m_2/(2\sin\alpha)) dK(r) M(dm_1) M(dm_2) - \\ &- \tilde{A}^2 \ . \end{split}$$

**Example.** If  $\Psi = \{[x_n; m_n]\}$  is an independently marked planar Poisson process then

$$\operatorname{Var}\{y(o)\} = \tilde{A}\sin^2\alpha$$

(see Holgate [2]) and

$$c_y(h) = \lambda \sin^4 \alpha \int_0^\infty U(h,m/(2\sin\alpha),m/(2\sin\alpha)) M(dm), \quad h > 0 \ ,$$

cf. (3.13) in Schmidt [10].

We now discuss briefly the previous formulae of Theorem 2. Their use in sampling situations is considered in Chapter 4.

In the general case where the assumption of independence in marking has not been made, the derivation of the covariance function of the relascope process  $\{y(\xi), \xi \in \mathbb{R}^2\}$  presumes knowledge of the K-function (or the g-function), the mark distribution M and the two-point mark distribution function  $M_r$ . Although one is able to calculate the covariance function, numerically at least, it may be difficult to construct two-point distribution functions which would be relevant for practical purposes. This problem will be bypassed if one is able to assume independence in marking; indeed, the density  $\lambda$ , the K-function and the mark distribution function M are then sufficient for the application of Theorem 2. The forestry literature may also be useful here: Some examples of modelling point patterns and diameter distributions of trees can be found in Tomppo [15] and Kilkki and Päivinen [3], respectively.

Assume now that the K-function describing the pattern of the location of trees, the mark distribution and, in addition, the  $k_{mm}$ -function

$$k_{mm}(r) = \int \int m_1 m_2 M_r(d(m_1, m_2)),$$

introduced by Stoyan [12], are known. The formula above indicates that this is not sufficient for determining the covariance function  $c_y(h)$ . For practical sampling design purposes one can proceed in two ways in order to determine a (not unique) solution. First, it is possible to construct a two-point distribution function leading approximately to the given mark correlation function, and then apply the general formulae derived above. Second, one can try to construct a marked point process  $\Psi = \{[x_n; m_n]\}$  having the given K- and  $k_{mm}$ -functions, and then apply simulation in order to obtain a Monte Carlo solution for  $c_y(h)$ . Here sample points are located in a regular lattice and the usual covariance function estimators developed in time series analysis can be applied. It is an open problem which one of the two approaches is better.

#### 4. Application of the moment formula to forestry sampling

In this chapter we assume that  $W \subset R^2$  is a bounded homogeneous forest area which is large enough for the border effects to be negligible. The forest is described by the motion-invariant marked point process  $\Psi = \{[x_n; m_n]\}$  restricted to the set W. We denote by

$$a(W) = \sum_{[x_n;m_n] \in \Psi} (\pi m^2/4) \ \mathbf{1}_W(x)$$

the (realization dependent) basal area and by  $\tilde{a}(W) = a(W)/\nu_2(W)$  the proportional basal area of the forest W.

In applying the method of angle counts we do not observe the process  $\Psi = \{[x_n; m_n]\}$  but instead the induced random field  $\{y(\xi), \xi \in W\}$  defined in Chapter 3. The first order characteristic of this random field corresponding to  $\tilde{a}(W)$  is

$$\tilde{y}(W) = \int_{W} y(\xi) d\xi / \nu_2(W).$$

Indeed, if we neglect the boundary effects then

$$\tilde{y}(W) \approx \int_{W} \sum_{[x;m] \in \psi} \sin^2 \alpha \ \mathbf{1}_{b(o,m/(2\sin\alpha)) + \xi}(x) \mathbf{1}_{W}(x) d\xi / \nu_2(W)$$
$$= \sum_{[x;m] \in \psi} (\pi m^2/4) \mathbf{1}_{W}(x) / \nu_2(W)$$
$$= \tilde{a}(W).$$

The small deviance is caused by the "smoothing" property of summation in the construction of the induced random field where trees outside W have also been counted. Hence, for large W the integral  $\tilde{y}(W)$  approximates  $\tilde{a}(W)$  and it is sufficient to consider the first quantity.

Let  $\xi_1, \xi_2, ..., \xi_n$  be *n* sample points in *W* in which the angle counts are measured. The observations are  $y(\xi_1), y(\xi_2), ..., y(\xi_n)$ . The organization of the sample points is called sampling design. Examples of spatial sampling designs are described e.g. by Matérn [7] and Ripley [9]. In the following we consider two of them: uniform random sampling where the *n* points are located independently and uniformly in *W*, and systematic sampling where, for simplicity, the *n* points form a square grid with side length *d*. In both cases, the unweighted mean

$$\bar{y}_n = \frac{1}{n} \sum_{1}^n y(\xi_i)$$

serves as an estimator of  $\tilde{y}(W)$  (and of  $\tilde{a}(W)$  in the case when W is large).

There are two sources of randomness in this plan: the first one is due to the marked point process model assumption and the second one due to random sampling. The expectations in the following are taken with respect to both of them. Now  $\bar{y}_n$  is unbiased if a sufficient guard area around W is allowed for, otherwise only unbiased in the limit when W grows to  $R^2$ . Therefore, for large W it can be considered as an unbiased estimator.

### 4.1. Comparison of sampling designs

If an estimator like  $\bar{y}_n$  is efficient it should produce small sampling error  $\bar{y}_n - \tilde{y}(W)$  and, in addition, it should not be costly when applied in practice. We shall consider here only the precision of the estimator.

In the comparison we use the error variance  $E\{[\bar{y}_n - \tilde{y}(W)]^2\}$  the size of which is affected by the forest process  $\Psi$ , the relascope angle  $\alpha$ , the particular sampling design applied, and, for small W, also the size and form of W. Therefore, following Matérn [7] we consider the error variance per sample point defined by

$$\sigma_p^2 = \lim_{n \to \infty} n \operatorname{Var}\{\bar{y}_n - \tilde{y}(W_n)\} ,$$

where  $\{W_n\}$  is an increasing sequence of forest areas such that  $n^{-1}\nu_2(W_n)$  is constant. For large n,  $n\operatorname{Var}\{\bar{y}_n - \tilde{y}(W_n)\}$  is nearly constant and adequate for comparisons.

For simple random sampling, according to Ripley [9],

$$n \operatorname{Var}\{\bar{y}_n - \tilde{y}(W)\} = \sigma^2 - E\{c_y(U, V)\}$$

where U and V are two independent uniformly distributed points in W,  $c_y$  is the covariance function of the random field  $\{y(\xi), \xi \in \mathbb{R}^2\}$  and  $\sigma^2 = \operatorname{Var}\{y(o)\}$ .

In the case of systematic sampling, where for simplicity, the sample points are located according to a square grid with side length d and, again, W is large compared to the range of  $c_y(r)$ , we obtain

$$n \operatorname{Var}\{\bar{y} - \tilde{a}(W)\} \approx n \operatorname{Var}\{\bar{y} - \tilde{y}(W)\}$$
$$\approx \sum_{\substack{u,v \\ integers}} c_y(d|u-v|) - \frac{1}{d^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_y(s,t) \, ds \, dt \, .$$

The first sum on the right hand side is over the lattice points (ud, vd),  $u, v = 0, \pm 1, \pm 2, \ldots$ . Usually  $c_y(r)$  decreases rapidly to zero; therefore, only few terms in the summation are non-negligible.

**Remark.** Note that in earlier work concerning Bitterlich sampling the variance  $\operatorname{Var}\{y(o)\}\$  was derived. This is enough under uniform random sampling for large areas, and also sufficient for systematic sampling if the grid formed by the sample points is sparse compared to the range of the covariance function  $c_y$ . In general, the comparison of the two observation methods under systematic sampling requires the covariance function.

## 4.2. Estimation of the sampling error in systematic sampling with angle count observations

If the sampling design is uniform random sampling then

$$\frac{1}{n(n-1)}\sum_{1}^{n} \left(y(\xi_i) - \bar{y}_n\right)^2$$

#### A random field approach to Bitterlich sampling

serves as an unbiased estimator for the error variance. For systematic sampling, on the other hand, no precise method for evaluation of the sampling error exists if we do not have a priori information about the covariance function  $c_y$ . A suggestion presented in Matérn [7] is the use of local quadratic forms. As an example we consider the following choice: four adjacent sample points, say  $\xi_1$ ,  $\xi_2$ ,  $\xi_3$  and  $\xi_4$ form a group (see Figure 1 below for notation). The quadratic formula

$$T_1 = \frac{1}{4} [y(\xi_1) - y(\xi_2) - y(\xi_3) + y(\xi_4)]^2$$

is used as an estimator for the error variance. Now

$$E(T_1) = \sigma_y^2 - c_y(d) + \frac{1}{2}c_y(d\sqrt{2}).$$

In the case where the sampling distance d is large compared to the range of  $c_y(r)$  (so that the adjacent sample points can be considered uncorrelated)  $T_1$  is unbiased for error variance per sample unit. That is not true in general: typically  $c_y(r)$  is convex and strictly decreasing and in this case  $T_1$  seems to over-estimate the error variance per sample point. In practice, several (say p) not necessarily non-overlapping groups of adjacent observations and corresponding quadratic forms  $T_1, T_2, ..., T_p$  are calculated (see Figure 1) and  $(\sum_{j=1}^{p} T_j)/Np$  serves as an estimator of the error variance per sample point.





The evidence reported in Matérn [7] shows that this is conservatively biased if the covariance function  $c_y$  is convex and strictly increasing. This method gives a rough idea of the precision of  $\bar{y}_n$  under systematic sampling. However, in many forestry applications this is adequate for practical purposes.

The role of the moment formulas in Theorem 2 is the following: The covariance function  $c_y(r)$  allows us to calculate the expectation  $E\{T_1\}$  and the error variance  $\operatorname{Var}\{\bar{y} - \tilde{a}(W)\}$  under different model assumptions. This gives further the possibility to calculate the size of the bias.

#### References

[1]	BITTERLICH, W.: Die Winkelzählprobe Allg. Forst. und Holzwirtsch. Zeitung 59, 1948, 4-5.
[2]	HOLGATE, P.: The angle-count method Biometrika 54, 1967, 615–623.
[3]	KILKKI, P., and R. PÄIVINEN: Weibull function in the estimation of the basal area dbh- distribution Silva Fennica 20, 1986, 149-156.
[4]	KUUSELA, K.: A basal mean area tree method in forest inventory Communicationes Instituti Forestalis Fenniae 61(2), 1966, 1-32.
[5]	MATÉRN, B.: Wie gross ist die "Relaskop-Fläche" Allgemeine Forstzeitung 79(2), 1969, 21–22.
[6]	MATÉRN, B.: The precision of basal area estimates Forest Science 18, 1972, 123-127.
[7]	MATÉRN, B.: Spatial variation Lecture Notes in Statistics 36, Springer-Verlag, Berlin, 1986, 1–151.
[8]	ORD, K.: How many trees in a forest? - Math. Scientist 3, 1978, 23-33.
[9]	RIPLEY, B.D.: Spatial statistics Wiley, Chichester, 1981.
[10]	SCHMIDT, V.: Qualitative and asymptotic properties of stochastic integrals related to random marked point processes Optimization 18, 1987, 737-759.
[11]	STOYAN, D.: On correlations of marked point processes Math. Nachr. 116, 1984, 197-207.
[12]	STOYAN, D.: Correlations of the marks of marked point process – statistical inference and simple models Elektron. Informationsverarb. Kybernet. 20, 1984, 285–294.
[13]	STOYAN, D., W.S. KENDALL, and J. MECKE: Stochastic geometry and its applications. - Wiley, Chichester, 1987.
[14]	SUKWONG, S., W.G. FRAYER, and E.W. MOGREN: Generalized comparisons of the precision of fixed-radius and variable-radius plots for basal-area estimates Forest Science 17, 1971, 263-271.
[15]	TOMPPO, E.: Models and methods for analysing spatial patterns of trees Communica- tiones Instituti Forestalis Fenniae 138, 1986, 1–65.
University of Jyväskylä Department of Statistics Seminaarinkatu 15 SF-40100 Jyväskylä	

Finland