ACTA STEREOL 1993; 12/2: 175-180
PROC 6ECS PRAGUE, 1993
ORIGINAL SCIENTIFIC PAPER

SECOND ORDER STEREOLOGY FOR ANISOTROPIC BOOLEAN SEGMENT PROCESSES WITH APPLICATION

Joël Chadœuf, Viktor Beneš*, Pascal Monestiez, André Kretzschmar**

Lab. de Biométrie, Domaine St Paul, BP 91, 84140 MONTFAVET France

* Dept of Mathematics, FSI, Czech Technical University,
Karlovo Nám. 13, 12135 Prague 2, Czech Republic

** Lab. de Zoologie, Domaine St Paul, BP 91, 84140 MONTFAVET France

ABSTRACT

This paper describes statistical properties of various estimators of intensity of anisotropic boolean segment processes. In order to quantify the estimation variances, the pair correlation function of appropriate random measures has to be evaluated. The projections of the process on R^1 and intersections with a system of $(d-1)$-dimensional parallel hyperplanes were studied. Some results of Beneš *et al* 1993 are used, where the second order stereological formula for the pair correlation function of the projection measure of anisotropic fibre processes was derived and estimation variances compared.

These results were applied to a study of soil porosity, where the earthworm burrow system were modelled by a segment process.

The variances of these estimators were compared and the convergence of the serial section estimator to the projection estimator illustrated. The variance of the serial section estimator decreases rapidly when the number of sections increases and flattens out to the variance of the projection estimator. This approximation was used to discuss the effect of the sample shape onto the variance of length intensity estimation by using serial sections.

KEY-WORDS: anisotropy, estimation variance, length density, boolean segment process.

INTRODUCTION

The length density is a basic characteristic of fibre processes. To estimate it, the first method is to measure all the fibres in a given volume. However, this is generally impossible and people use serial sections to estimate it (Hilliard 1967, Kanatani 1984) needing IUR series of parallel planes. Similarly, estimators based on second order stereology, needing in practice IUR random sections or slices through fixed points, have been introduced (see Jensen & Gundersen 1989, Kieu & Vedel Jensen 1993). More recently, specific methods have been developed when the projection of the fibres through thick sections is possible (see Gokhale 1993, Cruz-Orive & Howard 1991 for example).

However, IUR probes or projection of thick sections may not be available. In such cases, sampling methods based on sections with given orientations are the last remaining methods and an unbiased estimator is built if the rose of direction is known.

In this paper, the variance of estimators based on sections were compared to the projection estimator in the case of a stationary boolean anisotropic segment process with independent length and orientation, supposing the distributions of segment length and orientation to be known. This study was applied to the estimation of the length density of earthworm burrow systems. Earthworm burrows are an essential part of the biological porosity and play an important role in soil fertility. For experimental reasons, only horizontal or vertical sections of parallelepipeds of soil can be practically performed. The method was then used to evaluate the influence of the parallelepiped shape.

NOTATION

Let $x = (r, l)$ be a system of polar coordinates in R^3, $l = (\theta, \phi)$, and let $dl = \sin(\theta)d\theta d\phi$, θ being the colatitude with respect to the vertical axis and ϕ the longitude.
Let Φ be a stationary anisotropic boolean segment process in R^3, with independent length distribution $\mathcal{H}$ and orientation distribution $\mathcal{R}$ and denote λ the intensity of the Poisson point process of segment centers. Let us suppose that $\mathcal{H}$ and $\mathcal{R}$ admit continuous densities h and ρ. The mean fibre length per unit 3-dimensional volume in R^3 is $L = \lambda\bar{H} = \lambda\int_R h(x)dx$ and its pair correlation function p is:

$$p(x) = 1 + \frac{2\rho(l)}{\lambda\bar{H}^2 r^2}\int_r^\infty (y-r)h(y)dy \tag{1}$$

MEAN LENGTH ESTIMATOR

Let B be a parallelepiped having one of its faces horizontal and let u denote the vertical unit vector. The first estimator of L is built using the definition of L. Suppose all the fibres can be measured in B, it is

$$L_1 = \frac{\Phi(B)}{\nu(B)} \tag{2}$$

and its variance is:

$$\begin{aligned}\mathrm{var}(L_1) &= \frac{L^2}{\nu(B)^2}\int_{R^3} g_B(x)(p(x)-1)dx \\ &= \frac{2\lambda}{\nu(B)^2}\int_{(r,l)} g_B(r,l)\rho(l)\int_r^\infty (y-r)h(y)dydrdl\end{aligned} \tag{3}$$

where $g_B(x) = \nu(B \cap B_{-x})$ and $p(x)$ is the pair correlation function of Φ.

SERIAL SECTION ESTIMATOR

Consider a series (H_i) of parallel horizontal hyperplanes of vertical coordinates ia. Let n be the number of hyperplanes intersecting B. Let $N_i = \nu(B \cap H_i \cap \Phi)$ be the number of intersection points of the fibre process with $B \cap H_i$. An unbiased estimator of L is

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\frac{N_i}{\nu(H_i \cap B)\mathcal{F}_u} \tag{4}$$

where $\mathcal{F}_u = \int \mid cos(u,m) \mid R(dm)$. Its variance is:

$$\mathrm{var}(L_2) = \frac{1}{n^2\nu(H_i \cap B)^2\mathcal{F}_u^2}(n\mathrm{var}(N_1) + \sum_{i \neq j} cov(N_i, N_j)) \tag{5}$$

$\mathrm{cov}(N_i, N_j)$ being equal to

$$\lambda\int_r\int_l \nu((B \cap H_i) \cap (B \cap H_j)_{v(a|i-j|,l)})(r\cos(\theta) - a \mid i-j \mid)_+ \rho(l)h(r)drdl \tag{6}$$

where $v(r,l)$ is the vector of H_i of length $\mid r\tan(\theta)\mid$ and angular orientation ϕ and $t_+ = \max(0,t)$.

PROJECTION ESTIMATOR

Suppose a tends to 0, then an intersection measure Φ_u is defined as

$$\Phi_u(C) = \int_{\mathrm{Pr}(C)} N_y dy \tag{7}$$

where C is any borel set (Benes *et al* 1993) and $\mathrm{Pr}(C)$ is the projection of C on the vertical axis. We have $E(\Phi_u(B)) = L\nu(B)\mathcal{F}_u$ and an unbiased estimator of L is

$$L_3 = \frac{\Phi_u(B)}{\nu(B)\mathcal{F}_u} \tag{8}$$

whose variance is

$$\mathrm{var}(L_3) = \frac{L^2}{\nu(B)^2}\int g_B(x)(p_u(x)-1)dx \tag{9}$$

where $p_u(x)$ is the pair correlation function of Φ_u.
Using a similar method than for evaluating $p(x)$, one gets that the mean fibre length in the sector $s(O,r,l)$ under the condition that a fibre with orientation m hits the origin is:

$$E_0\left(s(O,r,l)\mid m\right)\begin{cases} = \mathcal{F}_u LV(s(0,r,l)) & \text{for } m\notin s_l \\ = \mathcal{F}_u LV(s(0,r,l)) + f(r)\mid\cos(u,m)\mid & \text{for } m\in s_l \end{cases} \tag{10}$$

The probability density of the tangent orientation at the typical fibre point being $\frac{|\cos(u,m)|}{\mathcal{F}_u}\rho(m)$, one gets

$$p_u(x) = 1 + \frac{\cos(u,l)^2}{\mathcal{F}_u^2}\frac{2\rho(l)}{\lambda\bar{H}^2 r^2}\int_r^\infty (y-r)h(y)dy \tag{11}$$

so that

$$\mathrm{var}(L_3) = \frac{2\lambda}{\nu(B)^2\mathcal{F}_u^2}\int_{(r,l)} g_B(r,l)\cos(u,l)^2\rho(l)\int_r^\infty (x-r)h(x)dxdrdl \tag{12}$$

The projection estimator is the limit of the serial section estimator when $a \to 0$. Figure 1 illustrates this convergence on the following example:

- the sampling volume is a parallelepiped of square basis of unit area and of height 10.
- the intensity of the Poisson point process of segment centers is equal to 1,
- the length distribution of each segment follows an exponential law with mean length 1,
- the angular distribution is a Bingham-Mardia distribution: $p(l) = c\exp\left\{-\kappa\left(\cos(2\theta) - 2\cos(2\theta_0)\right)^2\right\}$ where c is a normalizing constant. In the considered example, $\theta_0 = 0$ and $\kappa = 10$,

For small values of n, the intersection processes defined on each hyperplane have small correlations. Therefore, the variance of the serial section estimator decreases as $1/n$ for small values of n, flattens out and tends to the variance of the projection estimator when $n \to \infty$.

k-FACES ESTIMATOR

Let U_k, $k = 1..K$ be K faces of the sampling parallelepiped B, u_k a unit vector orthogonal to U_k. Let $N_k = \Phi \cap U_k$ the number of intersecting points. An unbiased estimator of L can be defined as in the serial sampling estimator as

$$L_4 = \frac{1}{K}\sum_{k=1}^{K}\frac{N_k}{\nu(U_k)\mathcal{F}_{u_k}} \tag{13}$$

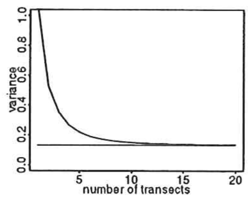

Figure 1. variances of the intersection estimators and the projection estimator (horizontal line)

whose variance is

$$\mathrm{var}(L_4) = \frac{1}{K^2}\left(\sum_{k,k'} \frac{\mathrm{cov}(N_k, N_{k'})}{\nu(U_k)\nu(U_{k'})\mathcal{F}_{u_k}\mathcal{F}_{u_{k'}}}\right) \tag{14}$$

where $\mathrm{cov}(N_k, N_{k'})$ is equal to $\lambda \int_r \int_l \nu((U_k \oplus rl) \cap (U_{k'} \oplus rl))h(r)\rho(l)drdl$ if $U_k \oplus rl$ denotes the dilation of U_k by the vector rl

SECOND MOMENT ESTIMATION

Let us suppose that Φ is known up to a given number of parameters, denoted θ in the following. Using the classical serial section estimator, $L\mathcal{F}_u$ only can be estimated, where u is the perpendicular to the sections. Estimation may then be performed using the covariances $\mathrm{cov}(N_0, N_y)$ or the second moments $E(N_0 N_j)$ between the numbers of intersections of Φ and two parallel planar sets V_1 and V_2 of same shape and distant of y.
Let us suppose $L = G(\theta)$, G continuously derivable, suppose the fibre process Φ to be a mixing field and denote $(N_j)_{1 \le j \le n}$ the number of intersections between Φ and (V_j) a set of parallel equidistant planar sets of same shape of surface area $\nu(V)$ and a the distance between two consecutive sets. Then (Guyon 1985) the vector $Z_n = \frac{1}{n-p}\left(\sum_{j=1}^{n-p} N_j N_{j+k}\right)_{1 \le k \le p}$ tends in probability to the vector $E(Z) = E((N_0 N_k)_{1 \le k \le p})$ and

$$\sqrt{n-p}\,(Z_n - E(Z)) \tag{15}$$

tends to a gaussian law of mean 0 and variance Γ.
Let

$$\theta_n = \mathrm{argmin} \sum_{k=1}^{p} \left\{ \mathrm{cov}_\theta(ak) - Z_{k,n} + \nu(V)^2 G(\theta)^2 \mathcal{F}_{u,\theta}^2 \right\} = \mathrm{argmin} f(\theta, Z_n) \tag{16}$$

Then, under classical regularity conditions on f and unicity of the solution of $f(\theta, E(Z_n)) = 0$, ensuring the existence of a two times derivable implicit function $\theta = h(Z)$, using the convergence property given above, one obtains (Dacunha-Castelle & Duflo 1983):
θ_n tends to the actual parameters θ_0 in probability and $\sqrt{n-p}(\theta_n - \theta_0)$ is asymptotically gaussian with 0 mean and covariance matrix $I = {}^t\left(\frac{\partial F}{\partial z}\right)^t \left(\frac{\partial F}{\partial t}\right)^{-1} \Gamma \left(\frac{\partial F}{\partial t}\right)^{-1} \left(\frac{\partial F}{\partial z}\right)$
where ${}^tX$ denotes the transpose of X, $\left(\frac{\partial F}{\partial t}\right)$ denotes the matrix of coefficients $\frac{\partial^2 f}{\partial t_i \partial t_j}$ and $\left(\frac{\partial F}{\partial z}\right)$ the matrix of coefficients $\frac{\partial^2 f}{\partial t_i \partial z_j}$
so that finally $L_n = G(\theta_n)$ tends to L in probability and $\sqrt{n-p}(L_n - L)$ is asymptotically gaussian with 0 mean and variance ${}^t\left(\frac{\partial G}{\partial t}\right) I \left(\frac{\partial G}{\partial t}\right)$
The variance of L_n depends of the moments of order 4 of the process N_j, for whom explicit formulas are difficult to obtain, even for simple models as the boolean segment process. Variance estimation can then be performed by simulation, either directly or by estimating the covariance matrix Γ if the estimation procedure is time consuming.

CASE STUDY

The full description of a natural earthworm burrow system was performed *in situ* with the following method: a face, from a large pit, was rendered as flat and vertical as possible. A column of soil (1x1x3dm) was described by destroying it little by little and every burrow segment was characterized by the three dimensional coordinates of its extremities and by its diameter. This natural burrow system was observed into a soil under permanent pastures; the soil is a brown soil near Dijon (France) where the earthworm population is mainly due to *Aporrectodea longa* and *A. nocturna* (80 % in biomass) which can be regarded as the main burrowing species. Burrows were modeled as a stationary boolean segment process with independent length and orientation defined as follows :

- the length is exponentially distributed
- let θ the colatitude and ψ the longitude of a segment, the distribution function of (θ, ψ) is modeled as $p(\theta, \psi) = ce^{-\kappa|cos(2\theta)|^{\alpha}}$ where c is a normalizing constant. It is an extension of the Bingham-Mardia distribution with $\theta_0 = 0$. The latter is obtained as soon as $\alpha = 2$.

The intensity of the point process was estimated as $\hat{\lambda} = \frac{N}{\nu B}$ where N is the number of upper segments points observed inside B obtaining $\lambda = 17.9$. The angle distribution was estimated by maximum of likelihood using the fibres whose upper point is in B, the estimated values being $\hat{\alpha} = 0.14$ and $\hat{\kappa} = 6.27$. The mean segment length τ was estimated as $\frac{L}{\lambda}$ were L is the estimated segment process intensity, $\hat{\tau} = 1.09$.

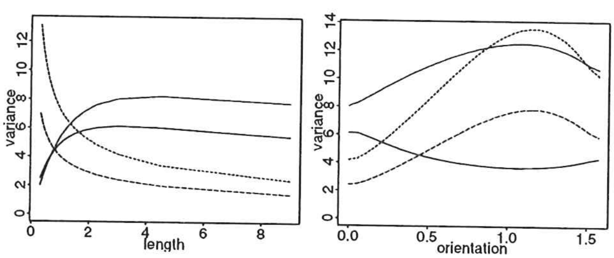

Figure 2. influence of the shape (vertical dilation : left picture) and orientation(on the right) of a parallelepiped of constant volume. plain line : classical estimator, dotted : projection estimator, broken lines: 1 and 2 -faces parallel to the projection.

In figure 2, the left picture presents the variance of L_1, L_3 and of the 1-face and 2-face estimators when the sampling volume is a parallelepiped of volume 3dm^3 whose vertical edge length z vary from 0.1dm to 9dm. The projection vector was the vertical unit vector, the faces were two consecutive vertical faces. var(L_1) and var(L_3) have similar variations, the variance of L_1 being higher to that of L_3. The 1-face and 2-face estimators present also similar variations.

For small values of z, due to the flatness of the parallelepiped and the strong anisotropy of the fibres, quite all fibres in the volume intersected its vertical face. So, the horizontal area explored by L_1 and L_3 was large (it varies as $1/z$) whereas the area explored by the 1-face and 2-face estimators varied as $1/\sqrt{z}$ and the variance of L_1 and L_3 was smaller than that of the 1-face and 2-face estimators.

For great values of z the same volume was explored by the four estimators, the variance of L_1 and L_3 tend to a plateau, before decreasing slowly to 0. The variance of the 1-face and 2-face estimators tend sharply to 0.

The right picture in figure 2 presents the variance evolution of the same estimators when the colatitude of the $1 \times 1 \times 3$dm^3 parallelepiped vary from 0 to $\pi/2$. The face of the 1-face estimator was chosen so that it became horizontal for a colatitude of $\pi/2$. L_1 and the 2-face estimator are less sensitive to the angular variations. The projection estimator and the 1-face estimator, much more sensitive to the angle of the fibres, show greater variance variations. The value at which an estimator is better than another depends in fact on the shape of the parallelepiped. The last method, was applied on the $1 \times 1 \times 3\text{dm}^3$ volume. The distance between two consecutive intersecting planes was 1mm. The estimation was performed using covariances estimated at 1, 2, 3, 4 and 5cm. The estimated variance was 91.

In this method, one needs to estimate all the parameters of the model. So, one can expect a large variance. Moreover, the size of the sampling volume is relatively small when compared with the mean size of the fibres and their orientation so that it leads to a poor estimation of the covariances, and consequently to large estimation variances of the parameters and of the intensity. To get a better estimation should require large samples which will give more precise covariance estimation. This will be balanced by less constraints on the object measurement as local angle measurements, precise location of the intersecting points and independent random sampling.

CONCLUSION

Assuming that the length distribution and the orientation distribution of a boolean segment process were known, we computed the variance of various estimators of the length density.

Approximating the variance of the serial section estimator by that of the associated projection estimator when the distance between successive sections tends to 0, we studied the influence of the shape of the explored sample on the variance of the density estimator.

When the angle and length distributions are known up to a given finite number of parameters, statistics based on the first moments cannot be used to estimate the length density when the process is anisotropic and if isotropic random sampling cannot be performed. Statistics based on the second moments can then provide an asymptotically unbiased estimator, whose variance is much more difficult to obtain.

Such analysis was carried on a soil analysis example, were serial sections along preferential sections is the only practical sampling method.

REFERENCES

Beneš V, Chadœuf J & Ohser J. On some characteristics of anisotropic fibre processes. Math Nachrichten, to appear.

Cruz-Orive LM & Howard CV. Estimating the length of a bounded curve in three dimensions using total vertical projections. J Micros **163**, 1991: 101-113.

Dacunha-Castelle D & Duflo M. Probabilités et statistiques. 2. Problèmes à temps mobile. Masson, Paris.1983: 286p.

Gokhale AM. Unbiased estimation of curve length in 3-D using vertical slices. J Micros 1990: **159**; 133-141.

Gokhale AM. Utility of the horizontal slice for stereological characterization of lineal features. J Micros 1993: **170**; 3-8.

Guyon X. Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas markovien. Actes de la 6^{me} Rencontre Franco-Belge des Statisticiens. Ed. des Fac. Univ. de Saint Louis. Bruxelles 1985.

Hilliard JE. Determination of structural anisotropy. In Stereology H Elias edit., Springer, New-York, 1967: 219-227.

Jensen EB & Gundersen HGJ. Fundamental stereological formulae based on isotropically oriented probes through fixed points-with application to particle analysis. J Micros **153**, 1989: 249-267.

Kanatani K. Stereological determination of structural anisotropy. Int J Engin Sci **22**, 1984: 531-546.

Kieu K & Vedel Jensen EB. Stereology using isotropic slices through fixed points. J Micros 1993: **170**; 45-51.