1 Boolean simulations

This exercice, which is mainly visual, is aimed at highlighting the general aspects and construction flexibility of boolean models ; it will also give an idea about the variability (very large, sometimes) from one realization to the other when the parameters of the model are determined. Some of these simulations shall be saved and will be used for the exercice about tests.

1. Simulate boolean sets with square, hexagonal or circular grains, with a fixed or variable size.

[*For the first ones, one only has to dilate Poisson points, for the second ones, to consider the union of several dilations of various sizes. Two procedures-* ***isobool*** *and* ***isobool 2*** *- may also be applied*].

2. The previous simulations are isotropic, at least on their definition grid. Carry out now some anisotropic simulations from segments or ellipses.

[*As previouly, make a union with Poisson points dilations with different size and orientation segments, or with ellipses (procedures* ***eldil*** *and* ***elbool****)*]

3. Simulate boolean sets with primary asymmetric grains, as for instance a mixture of equilateral upwards and downwards triangles.

[*Use the procedure* ***tribool****. The shapes catalogue may also be grown with* ***dropdil****, which dilates according to asymmetric droplets of the eight possible orientations*].

4. As a last example of non convex primary grains, we will use doublets of disjoint disks which centers are d apart (cell division model)

[*Make a union of a first realization of Poisson points with its horizontal translate of distance d, then a second one with its translate at angle 45° and of distance d, etc ... Make a union of all the points you got, and dilate them with a small disk*].

5. Start from a Poisson points simulation with a regionalized density g (cf. exercise 9-5), and pass to the boolean system with small homogranular circles. Explain.

[*We get a* binarization *technique of the function g*].

6. Simulate conic boolean functions (we will use the cones sizes and heights whether to isolate the islets or whether to invade space). Simulate a cones doublets realization.

[*Use procedures* ***conedil 1 and 2,*** *applied to Poisson points, consider the sup for several points density, matching several cones sizes. In case of doublets, act as for question 4 to get the points, and dilate with a cone*].

2 Boolean tests

1. Symbols $a\prime$ and $u\prime$denote the mean area and perimeter of the primary grain respectively, and q and u the porosity and specific perimeter after a booleanization. We know that

$$q = \exp\{-\theta a'\} \quad \text{and} \quad u = \theta u' \exp\{-\theta a'\}$$

Check both relations on several simulations of the previous exercise, and particularly on the disks doublets. What about the boarder effects ?

[*Use the function* ***perim****. Note the generality of the approach, which does not imply that the primary grain is connected. Note also that, only the products* $\theta u'$ *and* $\theta a'$ *are accessible, but neithert* u'*nor* a']

2. When the primary grain is supposed to be convex, the first boolean test is based on linear erosions, which must be negative exponentials. If K is the segment with length L and direction α, we indeed have

$$logQ(K) = logQ_1(L, \alpha) = -\theta(Ld'_\alpha + a') \qquad (1)$$

where d'_α is the mean diameter of the primary grain in direction α. Test relation (1) on one of the convex grains simulations of exercise 1, and apply it successively to grains and pores. Test it afterwards for circles doublets, and for primary grains with regionalized density.

[*Use the procedure* ***pvonl*** *(which corrects boarder effects) and report the measurements on semi-log paper. Note that* θ *is still inaccessible*].

3. Use the images of the previous question, and erode them now by hexagons or squares. Prove that the logarithm of Q(K), for a square with side L and directions α, β, is

$$\log Q(K) = \log Q_2(L, \alpha) = -\theta \left[L^2 + L(d'_\alpha + d'_\beta) + a'\right] \qquad (2)$$

which yields estimator θ such that

$$\theta L^2 = \log Q_1(L, \alpha) + \log Q_1(L, \alpha + \pi/2) - \log Q_2(L, \alpha) - \log q \qquad (3)$$

Discuss the experimental results. Does the formula remain admissible when the primary grain is not convex anymore ?

4. Measure the covariance of the disks doublets realization. Interpret the results.

[*The mean distance between two primary grain disks turns out to appear as a hole effect in the covariance*].

5. Develop the questions (2) and (3) for boolean islets.

[*Start from the simulations of conic boolean functions of the previous exercise. The function is boolean if and only if all sections are boolean sets. Tests arising from relations (1) and (2) are applied to each grey level t. By integrating over the gray levels t, we obtain*

$$-\int_T \log Q_1(t; L, \alpha)dt = \theta\left[L\rho'_\alpha + v'\right] \tag{1'}$$

with ρ'_α and v'integral of the directional gradient and volume of the primary grain respectively (similar extension for relation(2))]

6. Check that the specific number ν(t) of the maxima at altitude t of a random boolean function simulation fulfills the relation :

$$\nu(t) = \theta q(t)\varpi(t)$$

when $\varpi(t)$ represents the probability that the primary grain has a height t.

3 Boolean model and counting

The boolean model allows to calculate explicitly a number of relations between observable characteristics of the primary grain. These relations can be used as tests (cf exercise above), or, on the contrary, can be admitted in a heuristic fashion, without further discussion.

For instance,consider, in R^2, a random primary grain with a mean area a' and a mean perimeter u'. After booleanization, the number θ of grains per unit area is linked to the porosity q and to the specific perimeter s by the two relations

$$\theta a' = -\log q \quad \text{and} \quad \theta u' = s/q \tag{1}$$

1. Simulate a succession of boolean R.A.C.S. with circular grains (for the sake of simplicity) of variable densities, homogranular or not. Once s and q have been measured on the simulations (whose parameter θ is given), check the relevance of the first relation (1) and of the following relation

$$u'/a' = -s/q \log q \tag{2}$$

[*Use procedures **isobool** and **isobool2**. Define, for $\theta^* = n$ fixed, for instance, at 50, the validity scope of relations (1) and (2) with respect to a'*]

2. Simulate hierarchical R.A.C.S. (therefore non boolean), for which all variables taking place in relations (1) and (2) are known. To what extent are these relations still correct ? Explain.

[*You can start from the procedure **rose**, or directly from **isobool**, when a realization is intersected with a Poisson points set, followed with a circular dilation. The first relation (1) implies that we have quite a precise idea about the mean size of the primary grain ; the relation (2) rather implies that its shape is known, and that its size changes very little. Therefore rel. (2) may be better for the low values of q. Formula (1) and (2) are anyhow satisfied as long as the* *object under study can be locally modelled by a boolean model, that allows free covering*]

3. Extend the first relation (1) to boolean random functions, and test it by simulating boolean islets.

[*Dilate Poisson points by cones, by considering the sup of two or three realizations with different densities or sizes of cones. The relation (1) must now be integrated with respect to grey levels :*

$$\theta v' = -\int \log q_t dt$$

Check numerically the ranges of validity of the various parameters]

4 Hierarchical models

The adjective "hierarchical" indicates here the implementation of a second boolean generation (θ_2, X_2'), conditionally to a first one (θ_1, X_1'). The second grains mays lean on the first ones or avoid them. This leads to aggregates in the first case, and to separate sets in the second one. Although these models are rapidly incalculable, they can nevertheless be developed in easy and instructive simulations.

1. Simulate hierarchical models for each of both types, and choose $\theta_1 \neq \theta_2$ and primary grains of different sizes.

[*Procedure* ***rose****(as in desert rose) and* **hard**].

2. Iterate the hierarchies in order to construct a sequence. Boolean models will be considered with respect to smaller and smaller disks.

[*Procedure* ***flake*** *(as in Von Koch "snow flake") and* **disjoint**].

5 Boolean domain of attraction

Choose half a dozen ordinary tesselations. For instance, the starting point may be the watershed lines of images such as *electrop* or *barrier* (after a small preliminary filtering in order to avoid too many classes). Skeletons may also to drawn according to influence zones or even Poisson lines. In each partition, assign the value 1 or 0 to each class with a low probability p for the 1s. Consider the union of the results, and test if it can be considered as boolean. Draw the conclusions.

6 Poisson lines in the plane

1. Use a square grid and simulate Poisson points (density λ) on the axes O_x and O_y.Take the perpendiculars on these points. This leads to a family of Poisson anisotropic lines, that segment the space into rectangles.

Convert the previous drawing into an approximately isotropic process, by incorporating lines with slopes ± 1, ± 2 and $\pm 1/2$. Simulate several realizations for both methods and note the important size variability.

[*The procedure* ***lines*** *creates Poisson rectangles, and* ***diags****, Poisson polygons in the six other directions. In* ***lines****, n is the effective number of points on each axis ; in diags,* $n\sqrt{2}$ *points are implemented on each diagonal, and* $n\sqrt{5/2}$ *points on each axis with* ± 2 *and* $\pm 1/2$ *gradient. These operations can be deconditionalized when extracting beforehand n values in Poisson law with parameters* λL *(L being the unit length of the axes)*].

2. Prove that Poisson lines induce, on each line, a Poisson points process, with a density 2λ in the isotropic case. Compare this result with the previous 8 directions approximation.

[*Isotropic case : the lines with directions* $(\alpha,\ \alpha+d\alpha)$ *induce a small Poisson with a density* $\lambda \sin \alpha d\alpha$*, which implies, through a summation over* π*, a total density of* 2λ*. In digital case, the points induced on the axis* O_x *(for instance) have a density of* $(1+\sqrt{2}+6\sqrt{5})/\pi = 1,62$].

3. Conditional sections. Suppose K and K' be two compact convex sets with respective perimeters u and $u\prime$, and K'$\subseteq$ K. Prove that if one and only one line intersects K, then, the probability it also intersects K' equals the perimeters ratio u'/u.

[*If a line has a direction* $(\alpha, \alpha+d\alpha)$*, such a probability equals the diameters ratio* D'_α/D_α *in the direction* α*. As and varies the searched conditional probability is then given by*

$$\int_0^\pi \frac{D'_\alpha}{D'_\alpha}\frac{D_\alpha}{u} d\alpha = \frac{u'}{u}]$$

4. Mean chord. Knowing that a line intersects K with an area a, prove that the conditional expectation of the intersected length is $\pi a/u$.

[*In direction* α*, this expectation equals* a/D_α *; when deconditionalizing in* α*, we get* $\pi a/u$].

5. Specific number of peaks. Prove that if two lines intersect a disk of radius r, there is one chance out of two they intersect inside the disk. Infer that the number of intersections per surface unit equals $\nu = \pi\lambda^2$.

[*1st point : if the intercepted chord on the first line has a length l, the second one intersects it with the probability* $2l/2\pi r$ *(question 3). But the conditional expectation of l is* $\pi r/2$ *(question 4), which leads to the result.*

2nd point : the a-priori probability that two lines intersect the disk equals $\frac{(2\pi r\lambda)^2}{2}\exp\{-2\pi r\lambda\}$ *. When* $r \rightarrow r_{0,}$ *half the number equals* $\pi\lambda^2$].

6. Determine the relative variance var(A), in number, of the polygons size,by means of their average in measure M(A). It is reminded that

$$var(A) = \frac{\sigma^2(A)}{[E(A)]} = \frac{M(A)}{E(A)} - 1 \tag{1}$$

[Let Y_0 be the polygon including the origin. The probability it also contains the small area r dr dα equals exp $\{-2\lambda r\}$, which leads to :

$$M(Y_0) = \int_0^{2\pi} d\alpha \int_0^{\infty} \exp(-2\lambda r) r dr = \pi/2\lambda$$

On the other hand, since each peak labels one and only one polygon, we have $E(A) = 1/v$, and (1) induces $var(A) = \pi^2/2 - 1 \sim 4$. This particularly high value explains the large disparity of the sizes considered in simulation].

7 Poisson tesselations

The random function which is studied here-below comes from the "turning band method", a method from G. Matheron, and is constantly used in fractals construction. It is instructive for several reasons, not only as an application of Poisson lines, but also as a mean and variance-free example of a model, and that can only be reached through its increments. Consequently, we will also wonder about the meaning of the induced experimental covariances (critics also from G. Matheron, in his theory of regionalized variables).

1. In R^1, a Poisson tesselation is defined as follows. Given a realization of Poisson points $\{x_1, i \in I\}$, we consider the function f, constant betwen two consecutive points, and which jump by s_i at point x_i , where the s_i are independent random variables, with a mean 0 and a variance σ^2.

Simulate f when the jumps value ± 1 with a probability 1/2, and when Poisson density varies from 1 to 10.

[Procedure ***steps****, with $n \leq 40$; when n increases, notice the fluctuations for larger and larger ranges].*

2. Prove that the random function f has no mean nor variance, whereas the increment $|f(x+h) - f(x)|$ is stationary , with a zero mean and with a variance equals to $2\sigma^2 |h|$ so that the variogram of f is $\gamma(h) = \sigma^2 |h|$

[We find $2\sigma^2 |h|$ by poissonizing the number of jumps on $[0,h]$].

3. For $h > 0$, we know that the variogram of order one, $\gamma_1(h)$, is given by

$$\gamma_1(h) = E|f(h) - f(0)| = 2/\pi \int_0^{\infty} \{[1 - \exp \lambda h(\cos u - 1)/u]\} du \qquad (1)$$

By a direct proof, show that

$$\textit{for small } h \quad \gamma_1(h) \sim 2\lambda h - 4\lambda^3 h^3 + \varepsilon(h^3) \qquad (2)$$

$$\textit{for large } h \quad \gamma_1(h) \sim (\lambda h/\pi)^{1/2} \qquad (3)$$

Measure the γ_1's for the above simulations of ex. 7.1. Why does the observed linear behaviour continues on quite long distances ?

[In order to set the limited expansion near the origin (2), assign the probabilities P_n to the upwards jumps and P'_n to the downwards ones, with $P_n = P'_n = e^{-\lambda h}(\lambda h)^n/n!$. For small h, we have

$$\gamma_1(h) = P_0 P_1' + P_1 P_0' + 2\,[P_0 P_2' + P_2 P_0' + P_1 P_1'] + ...$$

which leads to relation (2). The lack of second order terms explains the quasi-linear experimental behaviour. For large values of h, set $\theta = (\lambda h/2)^{1/2}$. The sum of the positive jumps tends towards the variable $\theta^2 + Y\theta$, and the negative jumps one towards $-\theta^2 + Y\theta$, where Y is the reduced normal variable, hence

$$\gamma_1(h) \quad \sim \theta E\,|Y_1 - Y_2| = (\lambda h/\pi)^{1/2} \; .$$

Thus, γ_1, which was initially proportional to γ, is finally proportional to its square root].

4. Build rectangular Poisson tessalations, by summing the simulations of horizontal and vertical bands of the first question (two different densities λ_1 and λ_2 have to be considered). Measure the $\gamma_1(h)$ in the direction $\pi/2$. Interpret. What happens when these directions increase ?

[We come back to the previous case, and note that jumps on a diagonal segment of length h admit a decomposition into two independent families of positive and negative jumps, with the same Poisson parameter $(\lambda_1 + \lambda_2)h/2\sqrt{2}$. Up to factor $\sqrt{2}$, this results in the sum of an horizontal and a vertical component for γ, as well as for γ_1 when h is small and h large. Finally, as these characteristics do not depend upon the number of directions, they remain valid under averaging of Poisson bands in all directions of the space : numerous fractal reliefs are simulated on this principle].

5. Critics of the "finitary" mathematician to the above approach : " What does a infinite random function f mean ? You will ever have finite means to build it, which will lead to finite numbers as well. So, please, keep your subtleties and variograms for yourself, and use a covariance as everyone else does".

Well, let's try for a realization f of vertical Poisson bands, that we intersect by a segment of lenght L, and whose direction is orthogonal to the banding.

(a) Estimate the hypothetical covariance $C(h)$, which in fact does not exist here, from the experimental quantities

$$m^* = (1/L)\int_0^L f(x)dx \quad \text{and} \quad C^*(x,y) = (f(x) - m^*)(f(y) - m^*).$$

By putting

$$C^*(h) = (1/L-h)\int_0^{L-h} C^*(x+h,x)dx \tag{4}$$

Show that for h>0

$$E\left[C^*(x+h,x)\right] = 2L/3 + (x^2 + (x^2+h^2))/L - 2x - 2h \tag{5}$$

and, by integrating in x, derive

$$E\left[C^*(h)\right] = L/3 - 3/4h + 2h^2/3L \tag{6}$$

Calculate the mean of the experimental variogram and covariance on several band simulations. Comment on the results.

[*An apparent variance* $E\left[C^*(0)\right] = L/3$ *is found, depending on the length* L *of the segment considered. It is a pure artefact, since the true variance is infinite. Although there is no covariance (the variogram is linear), the biases introduced by this procedure of estimation result in an apparent confirmation of the existence of a covariance (with a range!). It will be noted also that the stucture of the phenomenon is extremely distorded: not only is the straight line replaced by a parabola, but even the* slope at the origin is changed *(* 3/4 instead of 1) *Thus* $C^*(h)$ *represents almost nothing of the true structure*]

(b) Show that the experimental variogram

$$2\gamma^*(h) = (1/L-h)\int_0^{L-h} \left[f(x+h) - f(x)\right]^2 dx \tag{7}$$

has an expectation

$$E[\gamma^*(h)] = \gamma(h) = |h|$$

and does not run into the same bias problem as $C^*(h)$.Comment

[*We observe that the experimental variogram of relation (7) does not involve any cumbersome mean value (so that* $E(\gamma^*(h) = \gamma(h))$*; it is linked to* $C^*(h)$ *by the expression.*

$$2C^*(h) = -2\gamma^*(h) + (L/L-h)m^* + (1/L-h)\int_h^{L-h} \left[(f(x) - m^*\right]^2 dx$$

which introduces a variance in right member. When the experimental variance of a phenomenon increases with the size of the zone investigated, without tending towards a horizontal asymptote, it is not wise to fit the structure under study with a stochastic model possessing a covariance. However, in such a case, the variogram still exists and its estimation is significant. Consequently it provides a safer *method than the covariance.*]

8 Poisson partitions

The model here-under was initially proposed by F. Conrad (1972) in order to describe aerial photographs of geological cracks as those presented in the course. But the approach by "mixed" random set, on which the model is based, is extremely general and applies to any set or random function in R^n. We shall limit ourselves here to R^2.

1. Let X_1 be a tesselation and X_2 a R.A.C.S. of the euclidian plane. Let us build a realization X_1^* of X : each polygon is kept or rejected, independently of the others, with a probability p. Affect the restriction of a distinct realization in X_2 to each remaining polygon Π. The union of these portions, together with X_1^*, may be considered as the realization of a random closed set A.

Denoting by Q the characteristic functional of A, prove that the following relation (1) is satisfied for any connected compact set K

$$Q(K) = Q_1(K)\,[q + p\; Q_2(K)] \tag{1}$$

where q = 1-p, and where Q_1 and Q_2 stand for the characteristic functionals of X_1 and X_2 respectively.

[*If K misses X_1^*, then it is contained in a single polygon* Π, *therefore we can equivalently say that K misses A, or misses the restriction $X_2^* \cap \Pi$, or again misses X_2^*. This leads to the relation (1)*].

2. Let X_1 and X_2 be two tesselations of isotropic Poisson lines in the plane, with respective densities λ_1 and λ_2. The relation (1) becomes

$$Q(K) = \exp\{-\lambda_1 u(K)\}\,[q + p\exp\{-\lambda_2 u(K)\}] \tag{2}$$

where u is the perimeter. Simulate A in a square grid, and check relation (2). When p = 1, is Q(K) different from the Poisson lines functional ? What is the explanation ?

[*We can use the procedure **lines 2**, which directly builds A according to the square grid isotropy, or start as well from a simulation **diags**, richer in directions, and then simulate A polygon after polygon. Take a λ_2 at least twenty times larger than λ_1, and for K a square or an hexagon with side k. Q(k) is estimated by the procedure **binerotest**. Denote, when setting on semi-logarithmic coordinates, that log Q(k) looks like two successive segments, whose angle is all the larger as p is small*].

3. Similarly, draw a simulation A where X_2 is a boolean R.A.C.S., and another one where X_2 is a random function of Poissons tubes.

[*In the latter case, we will get X_2 by dilating the simulations **diags** by cones*].

9 A few point models

1. Order two analysis. The points models that follow are stationary and generally admit a covariance *measure* C(dx), which can be decomposed into the sum of a Dirac measure and a *function* g(x)

$$C(dx) = m\delta(dx) + g(x)dx \tag{1}$$

The mean number of points Z_B in the borelian set B is

$$m(Z_B) = m\ a(B)$$

and the covariance between the numbers in B and in B' is

$$C(Z_B, Z_{B'}) = m\ a\ (B \cap B') + \int_{R^2} \int_{R^2} 1_B(x) 1_{B'}(y) g(x-y) dx\ dy \qquad (2)$$

In the isotropic case, there is a simpler way to reach g from experiments. Note H(r), the average number of points which are included in the disc of radius r, centered on a point x_i (not counted), which is itself part of the set model. By taking for B' the disc of radius ds centered x_i,and for B the disc (r, x_i) minus B', relation (2) becomes :

$$C(Z_B, Z_{B'}) = 2\ \pi\ dx \int_0^r g(s)\ s\ ds \qquad (3)$$

But, this covariance is nothing else than the centered version of $mH(r)$

$$C(Z_B, Z_{B'}) = E\left[Z_B - m(Z_B)\right]\left[Z_{B'} - m(Z_{B'})\right] = m\left[H(r) - \pi r^2 m\right] dx \qquad (4)$$

where the primitive $\left[H(r) - \pi r^2 m\right]$ of sg(r) is experimentally accessible.

Write a computing procedure for H(r)

[*Procedure* ***pp****, for "point packing"*].

2. Anisotropic Poisson points. Work again with procedure **points** and divide the simulation by setting in a memory the random points (x_i, y_i) and in the other (x_i, $y_i/2$). To which extend does the affinity of ratio 1/2 make the simulation more anisotropic from a visual point of view ? Interpret by calculating the number of points which are included in the flat (sides 2,1) and long (side 1,2) rectangles.

[*Only the density varies, here doubles. Such thing as Poisson points cannot be anisotropic !*]

3. Measure the primitive H(r) for Poisson points simulations, and for the centres in the model of disc doublets. Derive g(s) and explain. Calculate H(r) for the germs centers in the hierarchical sets derived from the boolean model.

[*For Poisson points with a density* λ*, we find, in square grid H(k)* $\sim \lambda(k + r)^2$ *and in hexagonal grid H(k)* $\sim \lambda(3\ k^2\ +\ 3\ k\ +\ 1)$*, which duly corresponds to Poisson law variance, and implies* $g(s) \equiv 0$*. In case of doublets from d apart, H(r) undergoes a thrust, at distance d. In hierarchical cases, clustering situations (****rose, flake****) lead to an additional increase of H(r), and repulsive ones (****hard, disjoint****) to a decrease, followed an asymptote in both cases. The range of function H(r) -* $\pi r^2 m$ *corresponds to the average size of the interactions between points*].

4. Calculate H(r) for the set of points generated by the crossing of isotropic Poisson lines, with an intensity λ. Derive the measure g(r) $= 4\lambda^3/r$. Simulate such a process and calculate its H(r). Give an interpretation.

[*We will simulate a process wich is close to the model, but easier to implement, by considering two realizations of* ***lines*** $\cup$ ***diags****, and by intersecting them. A linear increase of H(r) is noticed, unlike all the other point processes considered until now. In theory, if* x_i *is a point of the set, the disk B(r,* x_i*) includes, on the one hand, lines intersections different from those which define* x_i *(average number* $\pi r^2 \lambda^2$*, see "Poisson lines" exercice), and, on the other hand, the points of the two lines that intersect at point* x_i *(average number* $8\lambda r$*). Consequently, H(r) =* $\pi r^2 \lambda^2 + 8\lambda r$*, and*

$$\int_0^r sg(s)\, ds = 4\lambda^3/r \qquad \text{therefore} \qquad g(r) = 4\lambda^3/r\ ]$$

5. Regionalized density. In this last exercice, we come back to image processing. Let f be a numerical function of the plane. Simulate Poisson points with a variable density f. Apply the simulation by chosing for f a grey tone image. Explain.

[*Procedure* ***regpoints****. This simulation is explained as the binarization of the numerical grey image, whose quality is all the smaller since the involved gradients are steeper and steeper*].