Random Sets and Random Fields

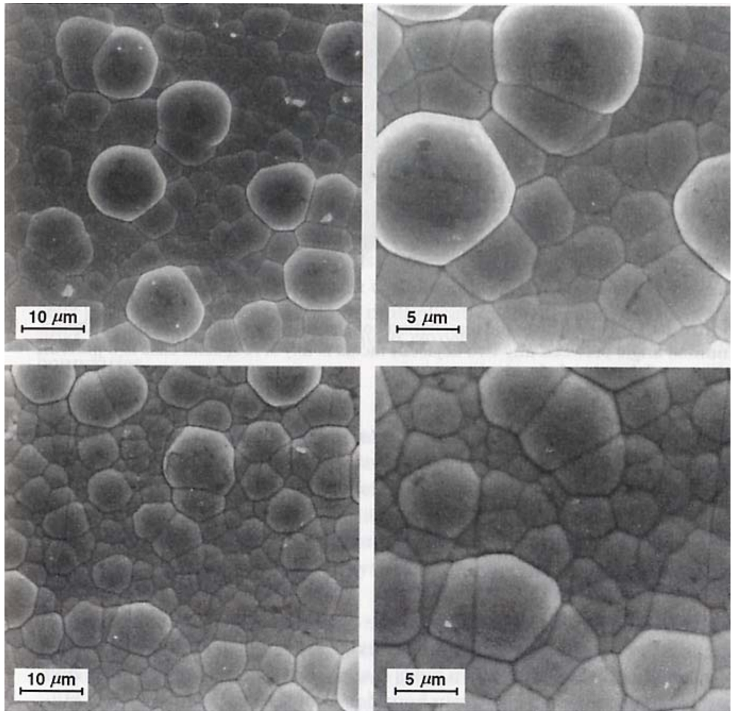
Evgueni Spodarev

Overview

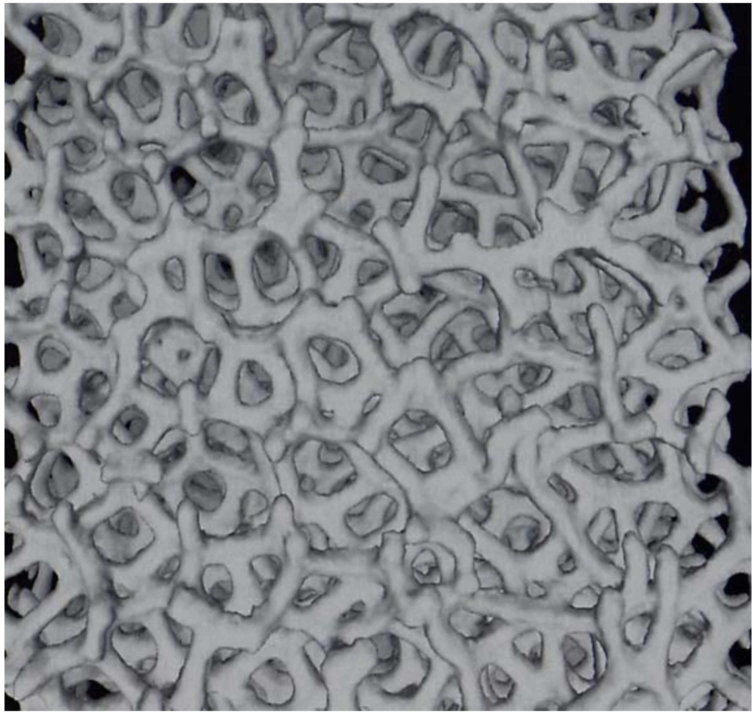
- Motivation
- Basics of mathematical morphology
- Introduction to the theory of random sets
- Comparison of grey scale images
- Random fields
- Examples of random fields
- Extrapolation of spatial data sets
- Applications
- Literature

Motivation

- Geometrical random structures in nature

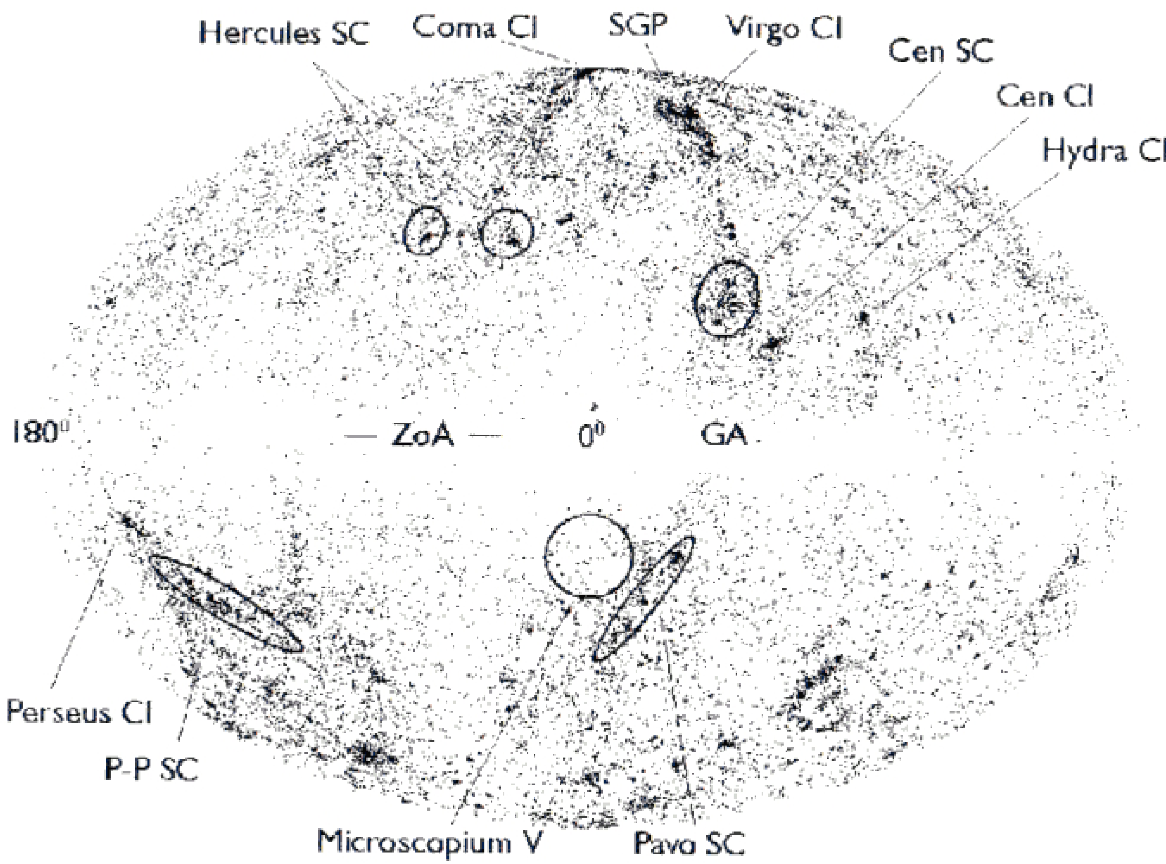


Alcaline zinc–nickel layer on steel



Nickel foam

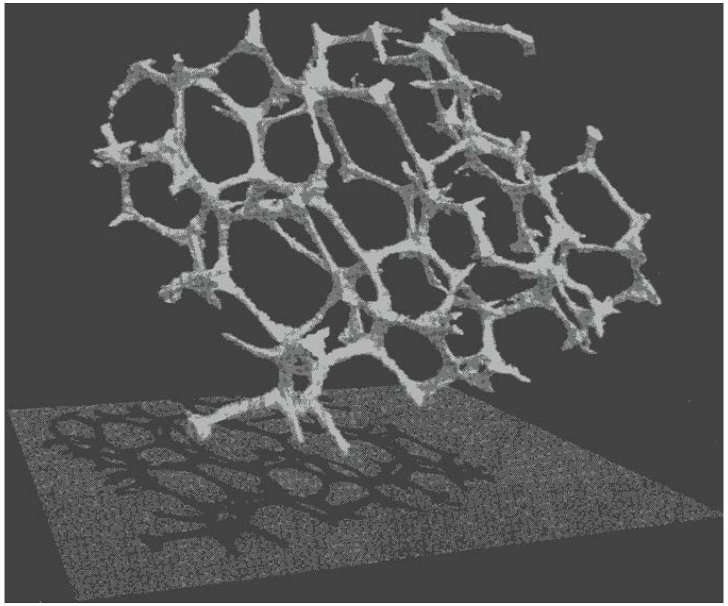
Motivation



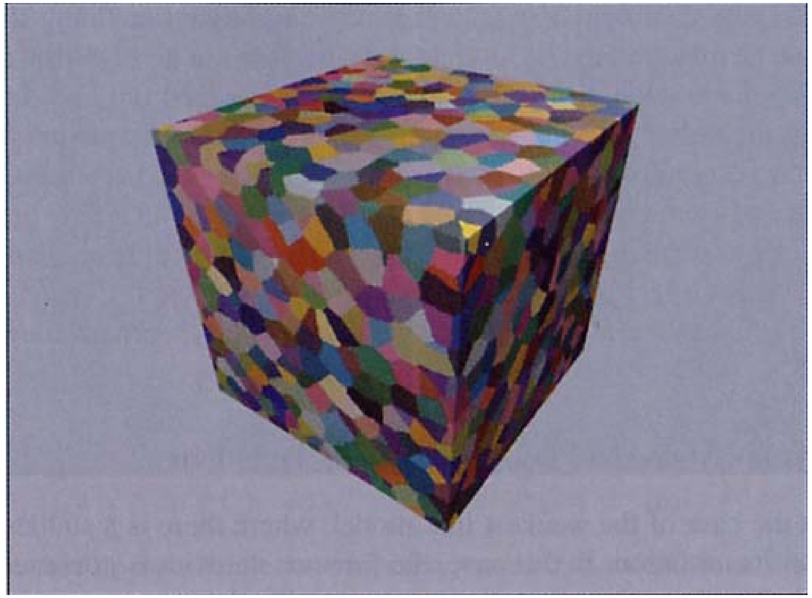
Distribution of galaxies in space (Hammer-Aitoff projection)

Motivation

- Modelling the structure of materials



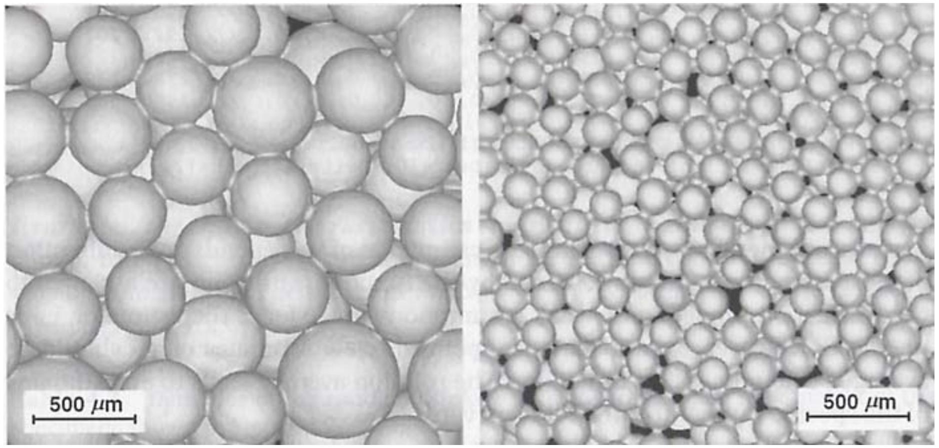
Polyurethan foam



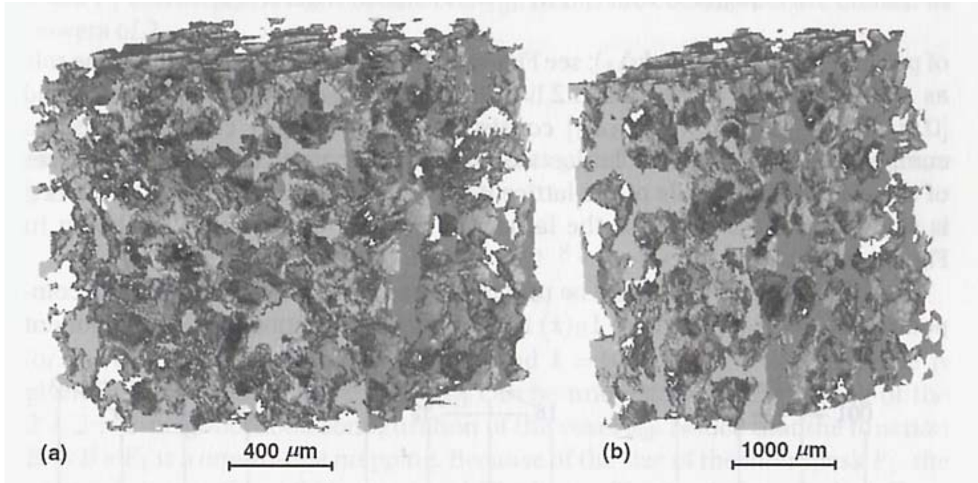
Edges of a 3D Voronoi tesselation

Motivation

- Estimation of the image characteristics



Microscopic structure of Cu powder:
the form and the size of particles



Porous structure of sand stone:
percolation

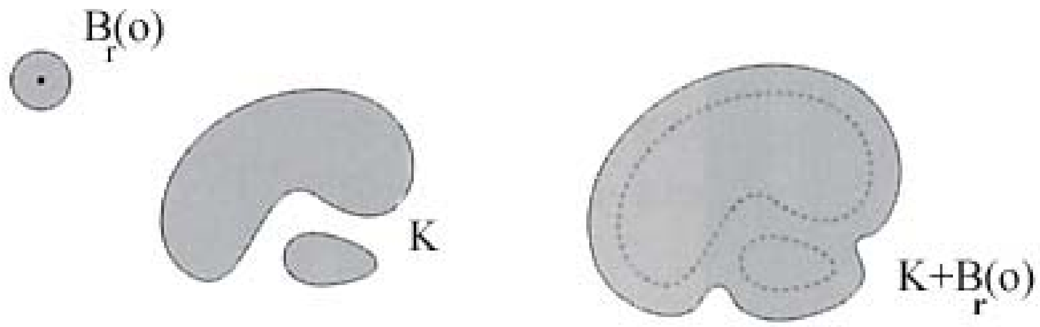
Mathematical morphology

- Basic notation

$\mathcal{K}$	family of all compact convex sets (bodies) in $\mathbb{R}^d$
$\mathcal{R}$	$= \{\bigcup_{i=1}^{n} K_i : K_i \in \mathcal{K},\ i = 1, \dots, n,\ \forall n\}$ convex ring
$\mathcal{S}$	$= \{K : K \cap W \in \mathcal{R},\ \forall W \in \mathcal{K}\}$ extended convex ring
$B_r(a)$	ball with center in a and radius r
κ_j	volume of $B_1(o)$ in $\mathbb{R}^j$, $j = 0, \dots, d$
$K_1 \oplus K_2$	$= \bigcup_{x \in K_2} (K_1 + x)$ Minkowski addition
$K_1 \ominus K_2$	$= \bigcap_{x \in K_2} (K_1 + x)$ Minkowski subtraction

Mathematical morphology

- Morphological operations



Dilation: parallel set $K \oplus B_r(o)$ of K

- Dilation $K \mapsto K \oplus (-B)$
- Erosion $K \mapsto K \ominus (-B)$
- Threshold filter $f(x) \mapsto \mathbf{1}(x : f(x) \geqslant a)$, $x \in W$

Notice that $K \ominus (-B) \oplus B \subseteq K \subseteq K \oplus (-B) \ominus B$

Intrinsic volumes

Steiner formula in $\mathbb{R}^2$

For any $K \in \mathcal{K}$ and $r > 0$

$$A(K \oplus B_r(o)) = A(K) + rS(K) + \pi r^2 \chi(K)\,,$$

where

- $A(K) =$ the area of K
- $S(K) =$ the boundary length of K
- $\chi(K) = 1$ the Euler number of K ("porosity")

Intrinsic volumes

Steiner formula in $\mathbb{R}^d$

- There exist functionals $V_j, W_j : \mathcal{K} \to [0, \infty)$, $j = 0, \dots, d$, (Minkowski functionals, quermassintegrals or intrinsic volumes) such that for any $r > 0$ and $K \in \mathcal{K}$ it holds

$$V_d(K \oplus B_r(o)) = \sum_{j=0}^{d} r^{d-j} \kappa_{d-j} V_j(K) = \sum_{j=0}^{d} r^j \binom{d}{j} W_j(K)$$

- where $W_j(K) = \frac{\kappa_j}{\binom{d}{j}} V_{d-j}(K), \quad \forall K \in \mathcal{K}$, and
- the functionals $V_0, \dots, V_d$ are additive, motion invariant, monotone with respect to inclusion, and locally bounded

Intrinsic volumes

In $\mathbb{R}^3$: For any $K \in \mathcal{K}$, $\partial K \in C^2$, it holds

- $V_3(K) = |K|$ (volume of K)
- $2V_2(K) = S(K)$ (surface area of K)
- $\pi V_1(K) = (1/2) \int_{\partial K} (1/R_1 + 1/R_2)\, d\sigma$ (integral of mean curvature of ∂K or, equivalently, $2\pi\times$ mean breadth of K)
- $4\pi V_0(K) = 4\pi = \int_{\partial K} (1/R_1 \cdot 1/R_2)\, d\sigma$ ($4\pi\times$ Euler number = integral of Gaussian curvature of ∂K),

where R_1 and R_2 are the principal radii of curvature of ∂K.

Intrinsic volumes

Theorem 1 (Hadwiger (1957))
Let $F : \mathcal{K} \to \mathbb{R}$ be any additive, motion invariant and continuous functional. Then, F can be represented in the form

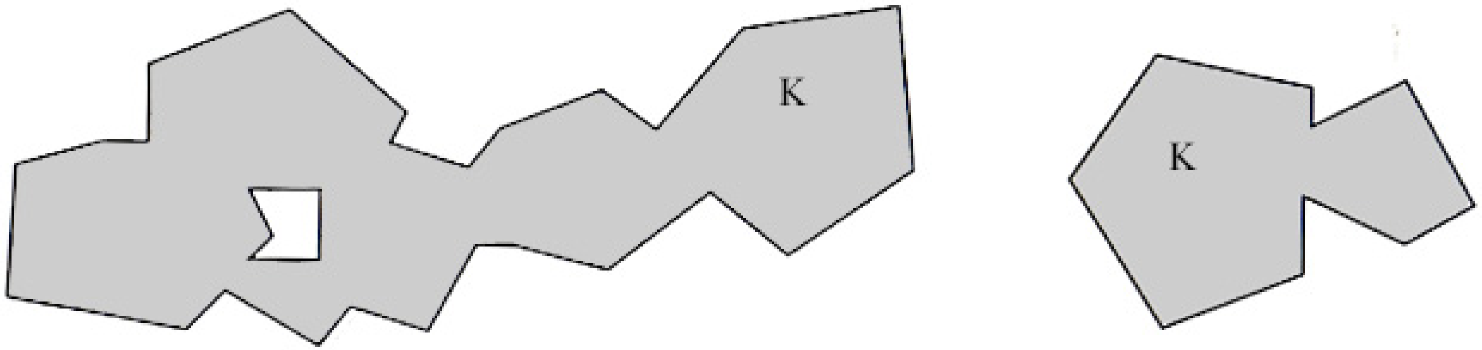
$$F = \sum_{j=0}^{d} a_j V_j$$

for some constants $a_0, \ldots, a_d \in \mathbb{R}$.

Thus, the intrinsic volumes $V_0, \ldots, V_d$ form a basis!

Intrinsic volumes

Additive extension to the convex ring $\mathcal{R}$



For each $j = 0, \ldots, d$, there exists a unique additive extension of $V_j : \mathcal{K} \to [0, \infty)$ to $\mathcal{R}$ given by the inclusion–exclusion formula:

$$V_j(K_1 \cup \ldots \cup K_n) = \sum_{i=1}^{n} (-1)^{i-1} \sum_{j_1 < \ldots < j_i} V_j(K_{j_1} \cap \ldots \cap K_{j_i}), \quad K_1, \ldots, K_n \in \mathcal{K}$$

Intrinsic volumes

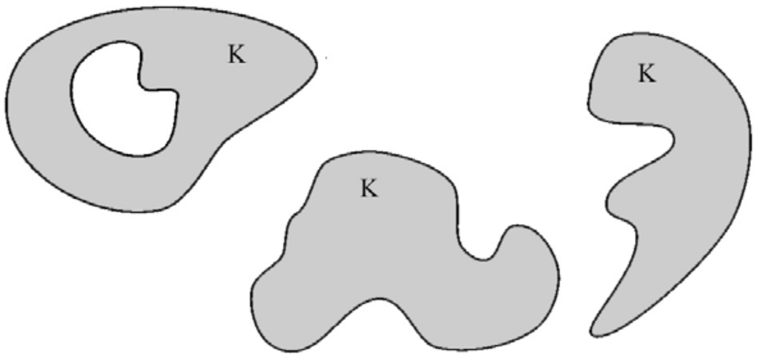
Geometrical interpretation: For any $K \in \mathcal{R}$ with $K \neq \emptyset$,

$V_d(K) = |K|$ (volume)

$2V_{d-1}(K) = S(K)$ (surface area)

$V_0(K) = \chi(K)$ (Euler number)

In $\mathbb{R}^2$: $\chi(K) = \#\{\text{clumps}\} - \#\{\text{holes}\}$



$\chi(K) = 3 - 1 = 2$

Intrinsic volumes

Steiner formula on $\mathcal{R}$ (Schneider (1980))

- Let the functional $\rho_r : \mathcal{R} \to \mathbb{R}$ be given by

$$\rho_r(K) = \sum_{j=0}^{d-1} r^{d-j} \kappa_{d-j} V_j(K), \qquad K \in \mathcal{R}$$

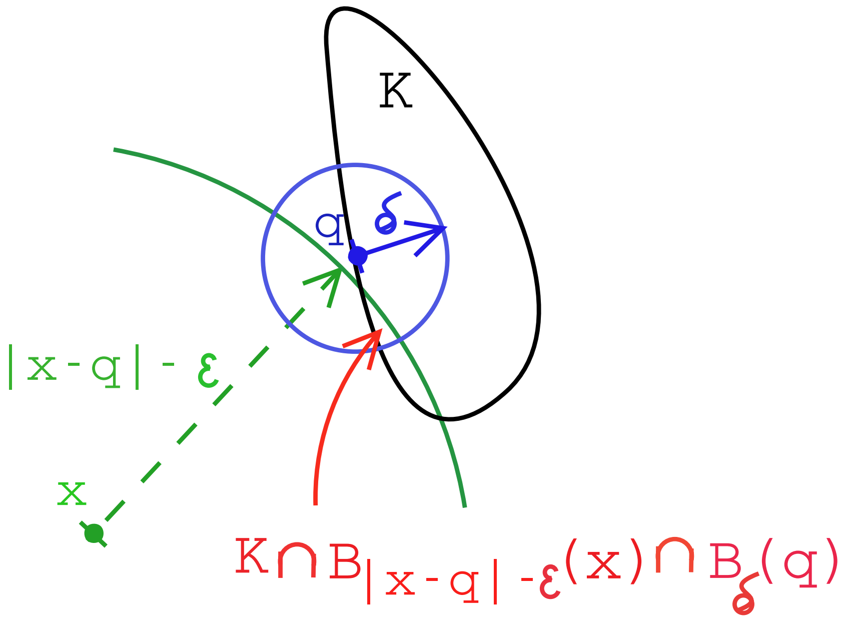
- For $K \in \mathcal{K}$, it holds $\rho_r(K) = V_d\big((K \oplus B_r(o)) \setminus K\big)$
- Geometrical interpretation of $\rho_r(K)$ for arbitrary $K \in \mathcal{R}$?

Intrinsic volumes

- Index function J

  For any $q, x \in \mathbb{R}^d$

  - put $J(\emptyset, q, x) = 0$
  - For $K \in \mathcal{R}$, $K \neq \emptyset$, let

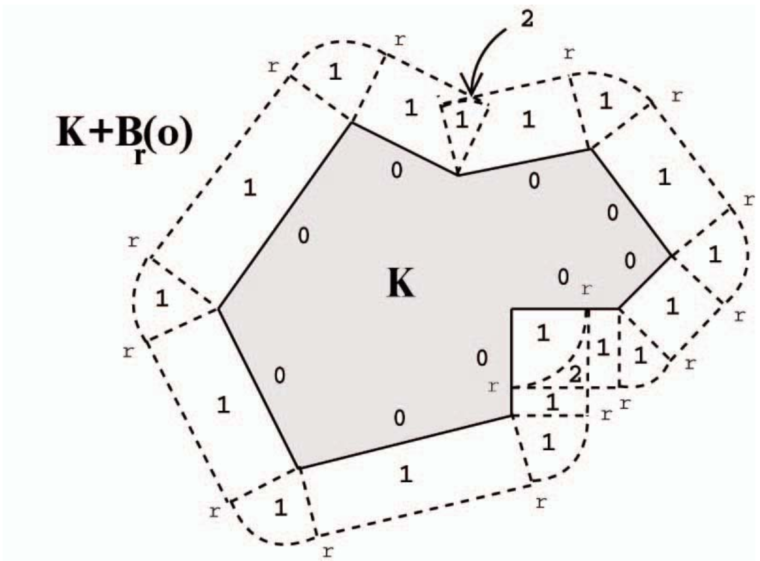


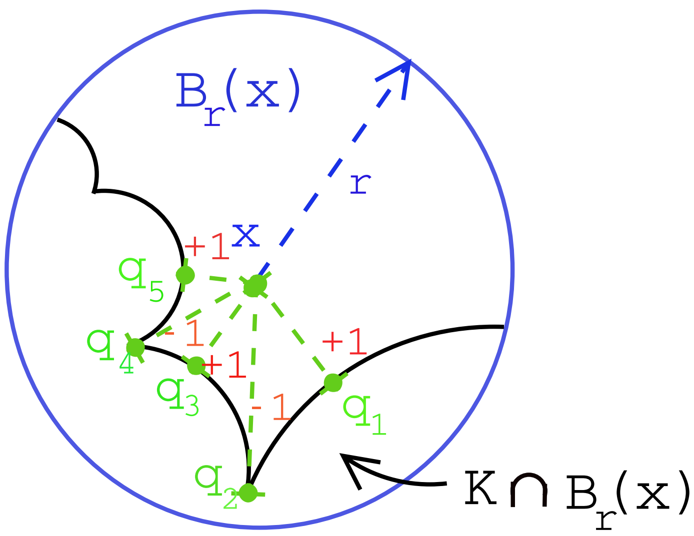
$$J(K, q, x) = \begin{cases} 1 - \lim\limits_{\delta \to +0} \lim\limits_{\varepsilon \to +0} V_0 \left(K \cap B_{|x-q|-\varepsilon}(x) \cap B_\delta(q) \right), & q \in K, \\ 0, & q \notin K. \end{cases}$$

Intrinsic volumes

- For any $r > 0$ and $K \in \mathcal{R}$, it holds

$$\rho_r(K) = \int\limits_{\mathbb{R}^d} I_r(K,x)\,dx, \quad \text{where } I_r(K,x) = \sum_{q \neq x} J\left(K \cap B_r(x), q, x\right)$$





$\rho_r(K) =$ volume of $\left(K \oplus B_r(o)\right) \setminus K$ weighted with multiplicities

Computation of intrinsic volumes

- Simultaneous computation of all $V_0(K), \ldots, V_d(K)$
  - Let $F_i : \mathcal{R} \to \mathbb{R}$, $i = 0, \ldots, d$ be additive, motion invariant and continuous functionals. Then, by Hadwiger's theorem,

$$F_i(K) = \sum_{j=0}^{d} a_{ij} V_j(K), \qquad \forall K \in \mathcal{R}$$

  - If $F = (F_0(K), \ldots, F_d(K))^\top$ can be easily computed and the matrix $A = (a_{ij})_{i,j=0}^{d}$ is regular,
  - then $V = (V_0(K), \ldots, V_d(K))^\top$ can be computed as the (uniquely determined) solution $V = A^{-1}F$ of the system of linear equations $F = AV$

Computation of intrinsic volumes

- Example: Steiner's formula on $\mathcal{R}$
  - $F_i(K) = \rho_{r_i}(K)$, $r_i > 0$, $r_i \neq r_j$ for $i = 0, \ldots, d-1$
  - $F_d(K) = V_d(K)$
  - $A = A_{r_0 \ldots r_{d-1}} =$

$$\begin{pmatrix} r_0^d \kappa_d & r_0^{d-1} \kappa_{d-1} & \ldots & r_0^2 \kappa_2 & r_0 \kappa_1 & 0 \\ r_1^d \kappa_d & r_1^{d-1} \kappa_{d-1} & \ldots & r_1^2 \kappa_2 & r_1 \kappa_1 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{d-1}^d \kappa_d & r_{d-1}^{d-1} \kappa_{d-1} & \ldots & r_{d-1}^2 \kappa_2 & r_{d-1} \kappa_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Computation of intrinsic volumes

- In $\mathbb{R}^2$
  - For any $r_0, r_1 > 0$ with $r_0 \neq r_1$, it holds

$$\begin{pmatrix} V_0(K) \\ V_1(K) \end{pmatrix} = \begin{pmatrix} \pi r_0^2 & 2r_0 \\ \pi r_1^2 & 2r_1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{r_0}(K) \\ \rho_{r_1}(K) \end{pmatrix}$$

$$= \frac{1}{2\pi r_0 r_1 (r_0 - r_1)} \begin{pmatrix} 2r_1 & -2r_0 \\ -\pi r_1^2 & \pi r_0^2 \end{pmatrix} \begin{pmatrix} \rho_{r_0}(K) \\ \rho_{r_1}(K) \end{pmatrix}$$

  - $V_2(K) = A(K)$ has to be computed separately

Random sets

- Let $(\Omega, \mathcal{F}, \mathcal{P})$ be an arbitrary probability space
- $\mathfrak{C}$ = family of all compact sets in $\mathbb{R}^d$
- $\mathfrak{F}$ = family of all closed sets in $\mathbb{R}^d$
- $\sigma(\mathfrak{F}) = \sigma$–algebra in $\mathfrak{F}$, generated by the sets $F_C = \{F \in \mathfrak{F} : F \cap C \neq \emptyset\}$ for any $C \in \mathfrak{C}$

An $(\mathcal{F}, \sigma(\mathfrak{F}))$–measurable mapping $\Xi : \Omega \to \mathfrak{F}$ is called a random closed set (RACS). Its distribution is uniquely determined by the capacity functional $T_\Xi(C) = P(\Xi \cap C \neq \emptyset)$, $C \in \mathfrak{C}$

Random sets

Characteristics of the capacity functional:

1. $0 \leqslant T_\Xi \leqslant 1, \quad T_\Xi(\emptyset) = 0$.
2. From $C_n \downarrow C$ follows $T_\Xi(C_n) \downarrow T_\Xi(C)$.
3. $S_n(C_0; C_1, \ldots, C_n) \geqslant 0$ for all $C_0, \ldots, C_n \in \mathfrak{C}$ and $n \in \mathbb{N}$, where

   $S_0(C_0) = 1 - T_\Xi(C_0)$,

   $S_1(C_0; C_1) = T_\Xi(C_0 \cup C_1) - T_\Xi(C_0)$,

   $\ldots$

   $S_n(C_0; C_1, \ldots, C_n) =$

   $S_{n-1}(C_0; C_1, \ldots, C_{n-1}) - S_{n-1}(C_0 \cup C_n; C_1, \ldots, C_{n-1})$.

Theorem 2 (Uniqueness)

- *If* Ξ_1 *and* Ξ_2 *are two random sets with* $T_{\Xi_1} = T_{\Xi_2}$ *then* $\Xi_1 \stackrel{\mathrm{d}}{=} \Xi_2$.
- *If* T *is a functional on* $\mathfrak{C}$ *satisfying the properties* 1–3 *then there exists a random set* Ξ *with* $T_\Xi = T$.

Random sets

Stationarity and isotropy

A RACS Ξ is called stationary if $\Xi \stackrel{\mathrm{d}}{=} \Xi + x$, $\forall\, x \in \mathbb{R}^d$, and isotropic if $\Xi \stackrel{\mathrm{d}}{=} g\Xi$, $\forall\, g \in SO(d)$

Theorem 3 (Matheron (1975))

- *The RACS Ξ is stationary (isotropic)* $\Longleftrightarrow T_\Xi(C + x) = T_\Xi(C)$ $\forall\, x \in \mathbb{R}^d$ *and* $T_\Xi(gC) = T_\Xi(C)\ \forall\, g \in SO(d)$, *respectively*
- *Each stationary RACS $\Xi \neq \emptyset$ is a.s. unbounded*
- *For any stationary convex RACS Ξ, it holds $\Xi \in \{\emptyset, \mathbb{R}^d\}$ a.s.*

Characteristics of random sets

- Volume fraction: $p_\Xi = P(x \in \Xi)$, $\forall\, x \in \mathbb{R}^d$.
  It holds $p_\Xi = T_\Xi(\{o\}) = E|\Xi \cap W|/|W|$ where $|W|$ is the volume of the observation window W.
- Covariance function: $C_\Xi(x) = P(\{o, x\} \in \Xi)$, $x \in \mathbb{R}^d$.
  It holds $C_\Xi(x) = 2p_\Xi - T_\Xi(\{o, x\}) = E|\Xi \cap (\Xi - x) \cap W|/|W|$.
- Centered covariance function: $Cov_\Xi(x) = C_\Xi(x) - p_\Xi^2$. It holds $Cov_\Xi(x) = E\left[\mathbf{1}(o \in \Xi)\mathbf{1}(x \in \Xi)\right] - E\,\mathbf{1}(o \in \Xi) E\,\mathbf{1}(x \in \Xi)$.
- Contact distribution function: $H_K(r) = P(d_K(o, \Xi) \leqslant r | o \notin \Xi)$, $r \geqslant 0$, where K is a convex body with $o \in K$ and $d_K(x, F) = \min\{r \geqslant 0 : (x + rK) \cap F \neq \emptyset\}$, $F \in \mathfrak{F}$, $x \in \mathbb{R}^d$.

Examples

- Germ–grain models: $\Xi = \bigcup_{i=1}^{\infty} (\Xi_i + X_i)$, where
  - $\{X_1, X_2, \dots\}$ = point process (of germs) and
  - $\{\Xi_1, \Xi_2, \dots\}$ = sequence of nonempty compact RACS (random grains)

Theorem 4

Each RACS Ξ can be represented as a germ–grain model

$$\Xi = \bigcup_{i=1}^{\infty} (\Xi_i + X_i)$$

Examples

Stationary germ–grain models in $\mathbb{R}^2$

Realizations of germ–grain models: Boolean model with spherical and polygonal grains, respectively; cluster process of segments

Examples

Boolean model

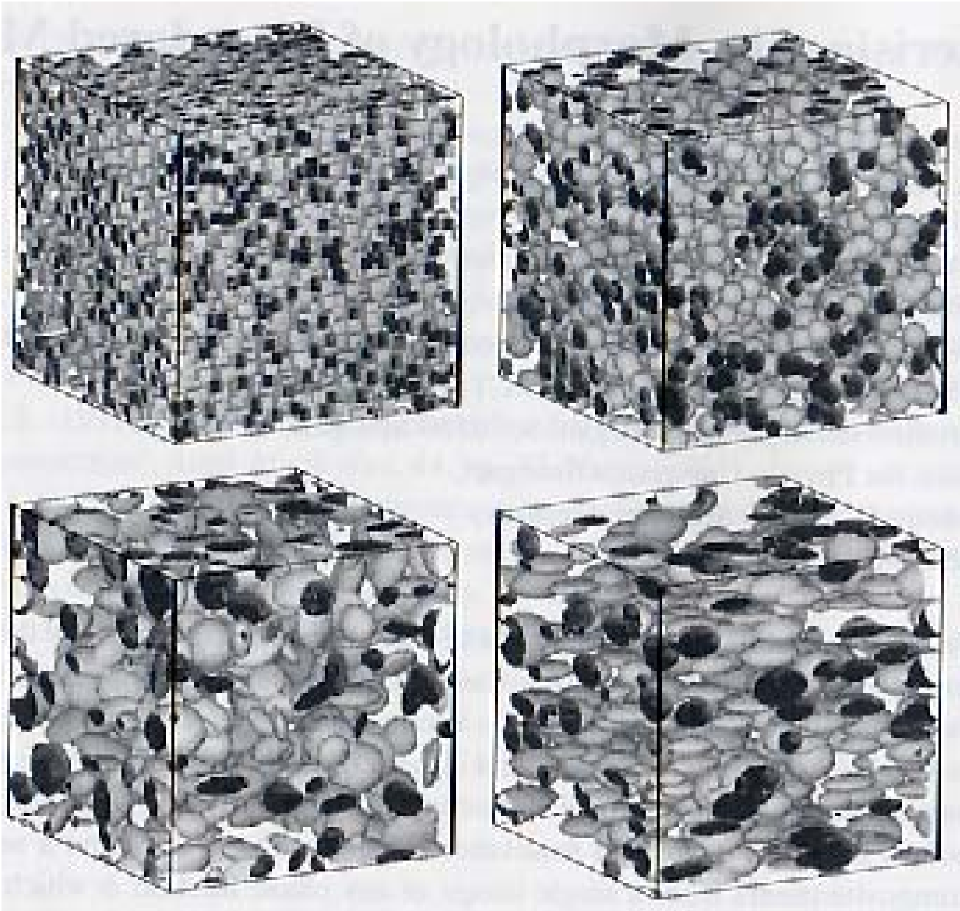
The germ–grain model $\Xi = \bigcup_{i=1}^{\infty} (\Xi_i + X_i)$ is called a Boolean model if

- the point process of germs $\{X_1, X_2, \ldots\}$ is a stationary Poisson process in $\mathbb{R}^d$ (with intensity λ)
- the grains $\Xi_1, \Xi_2, \ldots$ are i.i.d. and independent of $\{X_1, X_2, \ldots\}$; $\Xi_i \stackrel{\mathrm{d}}{=} \Xi_0$
- $E \left|\Xi_0 \oplus K\right| < \infty, \quad \forall K \in \mathcal{K}$.

Capacity functional: $T_\Xi(C) = 1 - e^{-\lambda E\,|(-\Xi_0)\oplus C|}, \quad \forall C \in \mathfrak{C}$

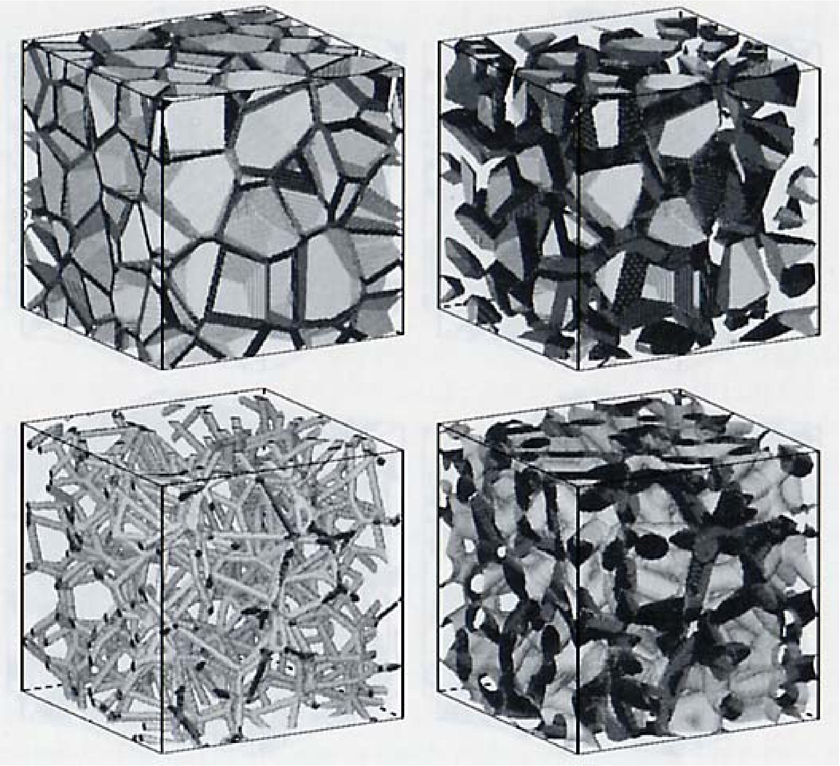
Examples

Boolean models in $\mathbb{R}^3$



Examples

Random sets made from 3D Voronoi tesselations



Specific intrinsic volumes

- Model assumptions
  - Let Ξ be stationary, $\Xi \in \mathcal{S}$ a.s.
  - $E\, 2^{N(\Xi \cap [0,1]^d)} < \infty$, where $N(\emptyset) = 0$ and
    $N(K) = \min\{m \in \mathbb{N} : K = \bigcup_{i=1}^{m} K_i,\ K_i \in \mathcal{K}\}$ for $K \in \mathcal{R} \setminus \{\emptyset\}$

- Specific intrinsic volumes: Let $\overline{V}_j(\Xi) = \lim_{n\to\infty} \frac{E\, V_j(\Xi \cap W_n)}{|W_n|}$
  for $j = 0, \ldots, d$, where $\{W_n\}$ = sequence of monotonously increasing sampling windows $W_n = nW$ with $W \in \mathcal{K}$ and $|W| > 0$

  In particular, $\overline{V}_d(\Xi) = P(o \in \Xi) = E|\Xi \cap W|/|W|$

Estimation of $(\overline{V}_0(\Xi), \ldots \overline{V}_d(\Xi))^\top$

Problem: Estimate $\overline{V}(\Xi) = (\overline{V}_0(\Xi), \ldots \overline{V}_d(\Xi))^\top$ on the basis of a single sample from $\Xi \cap W$

Solution: For each $i = 0, \ldots, d$, consider a random field $Y_i = \{Y_i(x), x \in \mathbb{R}^d\}$ such that

- Y_i is stationary of second order, i.e. $E\,Y_i(x) = \mu_i$ and $Cov(Y_i(x), Y_i(x+h)) = Cov_{Y_i}(h) \quad \forall\, x, h \in \mathbb{R}^d$
- $\mu_i = E\,Y_i(o) = \sum_{j=0}^{d} a_{ij} \overline{V}_j(\Xi)$, where the matrix $A = (a_{ij})_{i,j=0}^d$ is regular

Then, it holds $\overline{V}(\Xi) = A^{-1}\mu$, where $\mu = (\mu_0, \ldots, \mu_d)^\top$

Estimation of $(\overline{V}_0(\Xi), \ldots \overline{V}_d(\Xi))^\top$

- Steiner formula:

$$\lim_{n\to\infty} \frac{E\,\rho_r(\Xi \cap W_n)}{V_d(W_n)} = \sum_{j=0}^{d-1} r^{d-j} k_{d-j} \overline{V}_j(\Xi), \qquad r > 0.$$

- Writing the above formula for radii $r_0, \ldots, r_{d-1}$, $r_i \neq r_j$, $i \neq j$ together with $\frac{E\,V_d(\Xi \cap W_n)}{V_d(W_n)} = \overline{V}_d(\Xi)$, we get a system of $d+1$ linear equations with variables $\overline{V}_j(\Xi)$, $j = 0, \ldots, d$.

Estimation of $(\overline{V}_0(\Xi), \ldots \overline{V}_d(\Xi))^\top$

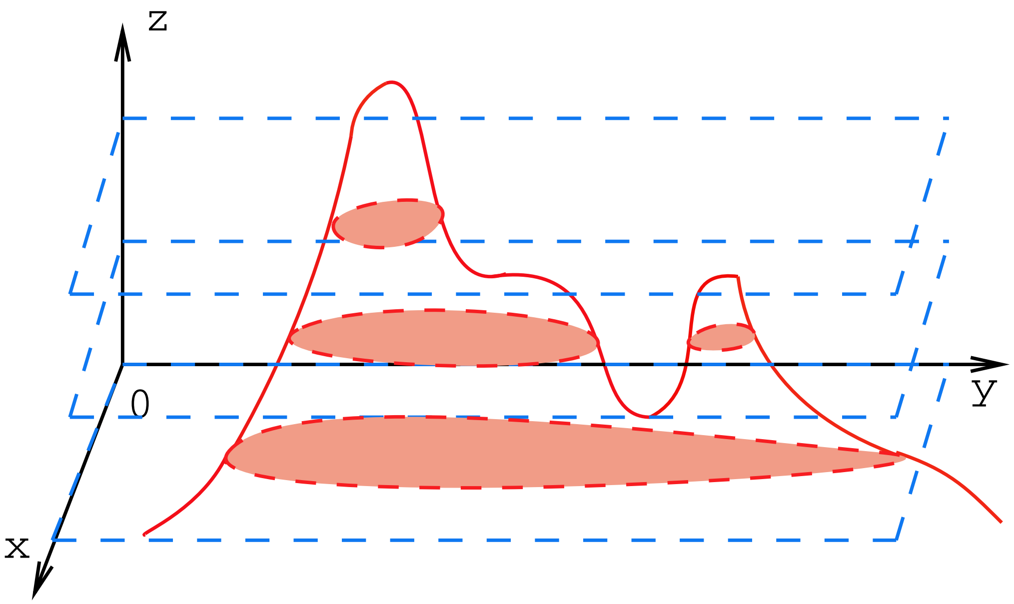
In matrix form: $A_{r_0 \ldots r_{d-1}} \, v = c$, where $v = \left(\overline{V}_0(\Xi), \ldots, \overline{V}_d(\Xi)\right)^\top$,

$$c = \left(\lim_{n \to \infty} \frac{E \, \rho_{r_0}(\Xi \cap W_n)}{V_d(W_n)}, \ldots, \lim_{n \to \infty} \frac{E \, \rho_{r_{d-1}}(\Xi \cap W_n)}{V_d(W_n)}, \frac{E \, V_d(\Xi \cap W_n)}{V_d(W_n)} \right)^\top$$

$$A_{r_0 \ldots r_{d-1}} = \begin{pmatrix} r_0^d k_d & r_0^{d-1} k_{d-1} & \ldots & r_0^2 k_2 & r_0 k_1 & 0 \\ r_1^d k_d & r_1^{d-1} k_{d-1} & \ldots & r_1^2 k_2 & r_1 k_1 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{d-1}^d k_d & r_{d-1}^{d-1} k_{d-1} & \ldots & r_{d-1}^2 k_2 & r_{d-1} k_1 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 \end{pmatrix}.$$

Comparison of grey scale images

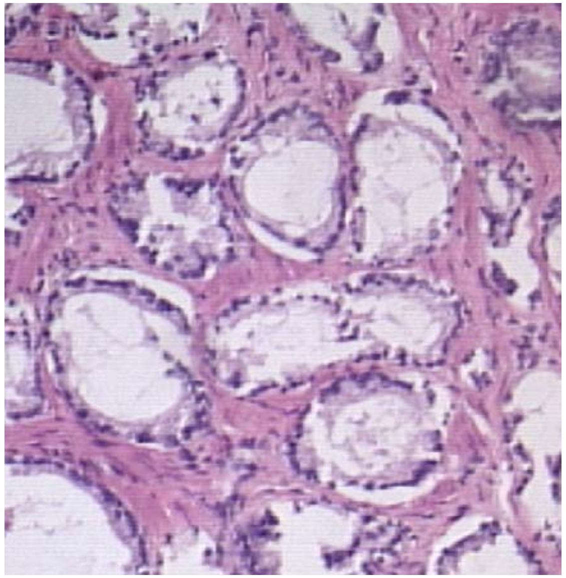
- Grey scale image $\mapsto$ family of binary images

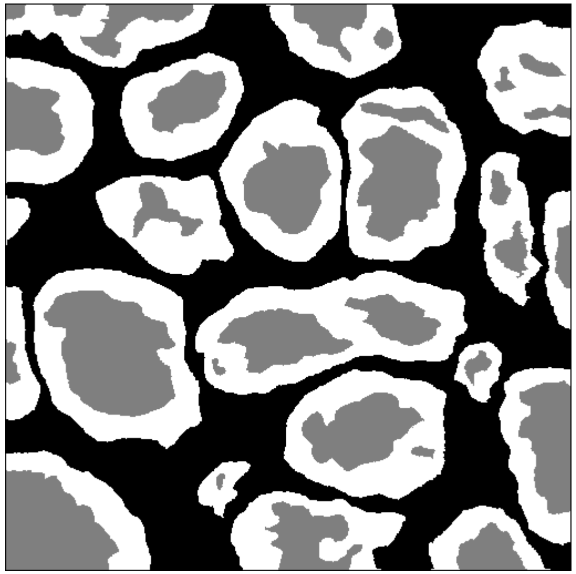


Individual grey scales can be represented by binary images that are analyzed in the sequel.

Comparison of grey scale images

- Comparison of particular grey scales

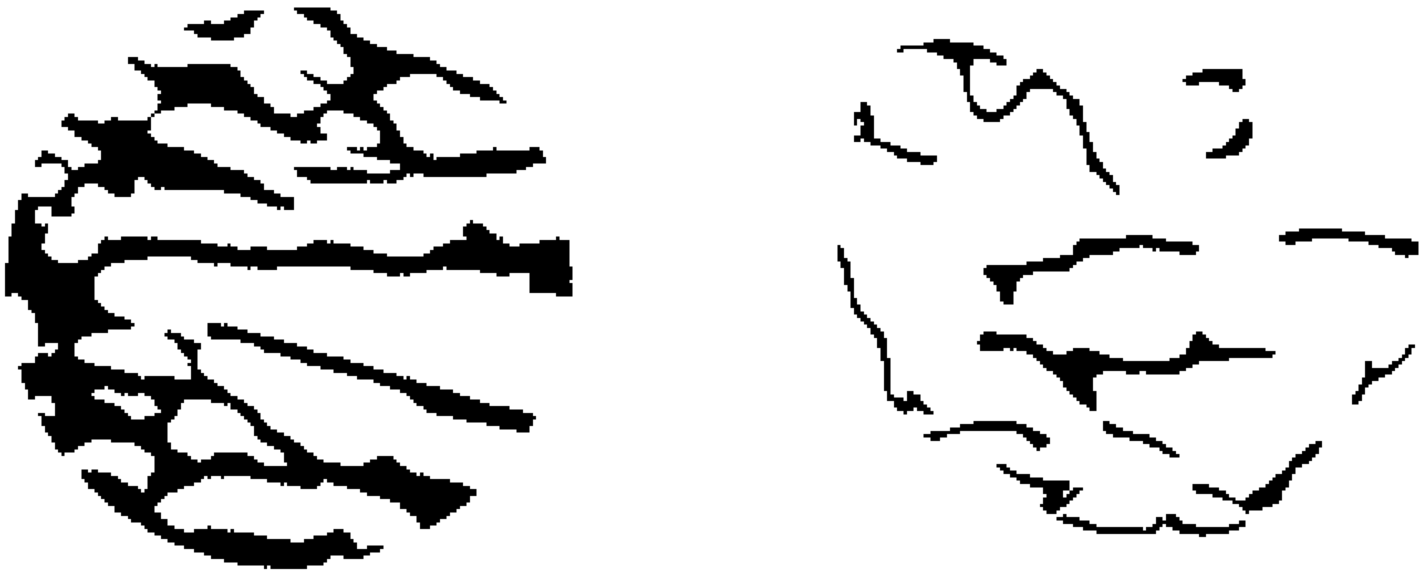




Histological section of prostate tissue: cancer diagnostics

Comparison of grey scale images

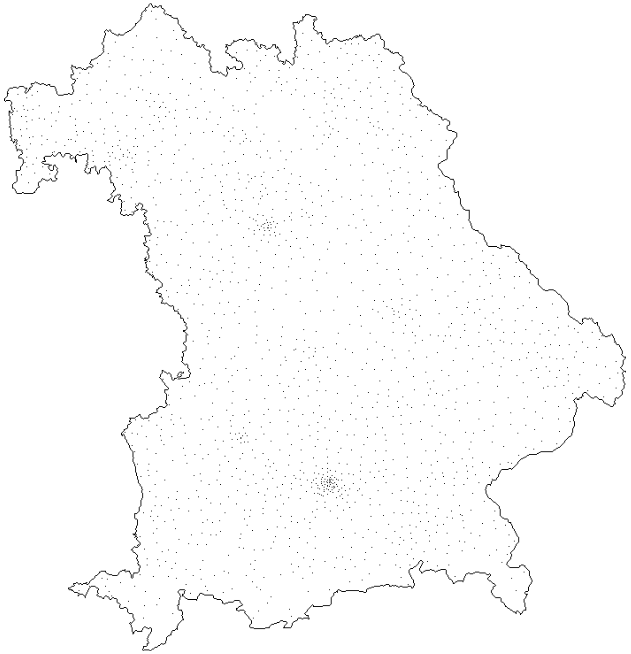
- Binary images: comparison of their estimated (specific) intrinsic volumes



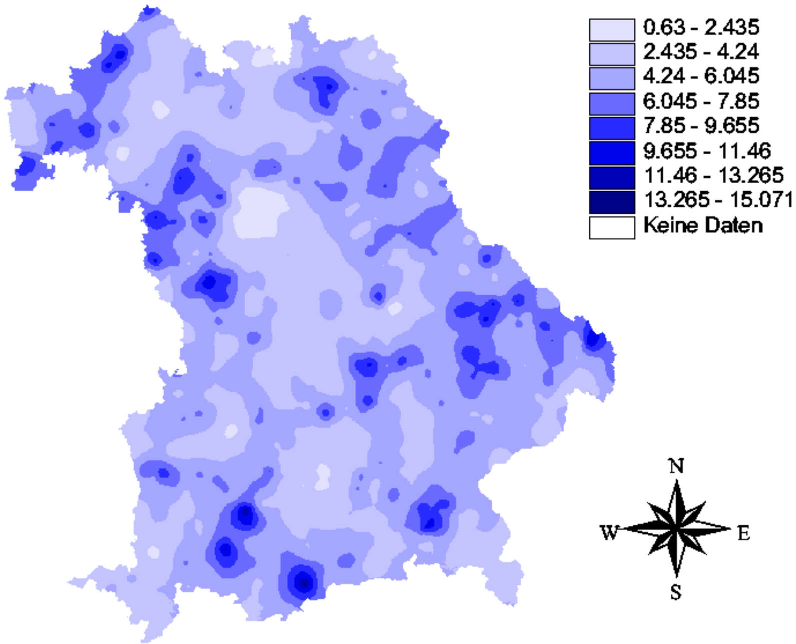
Bone structure: calcium phase of the healthy and deceased bone (osteoporosis)

Random fields: motivation

- Motor car insurance (Bavaria): significant changes of the number of cancellations of insurance policies



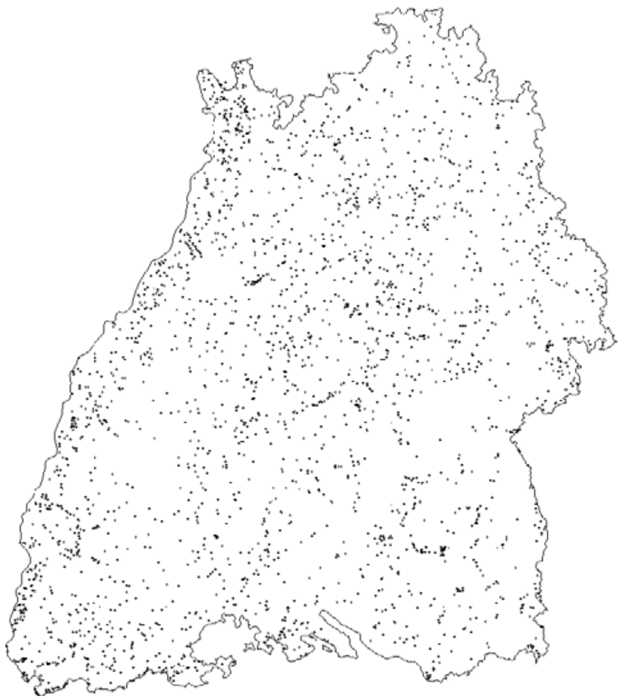
Centers of regions with the same postal code



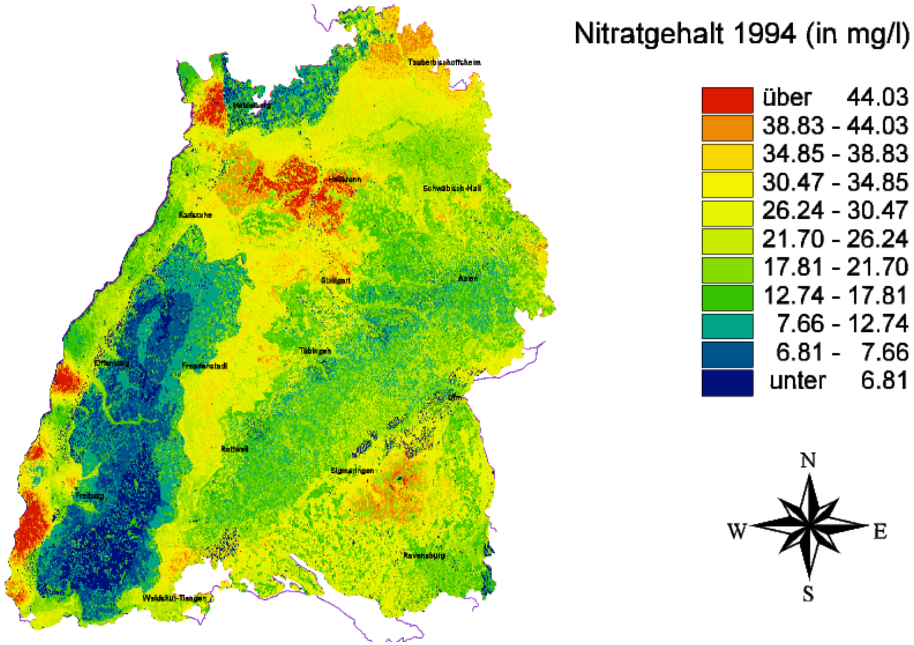
Extrapolated numbers of cancellations in 1998

Random fields: motivation

- Quality of ground water in Baden–Württemberg



Bores

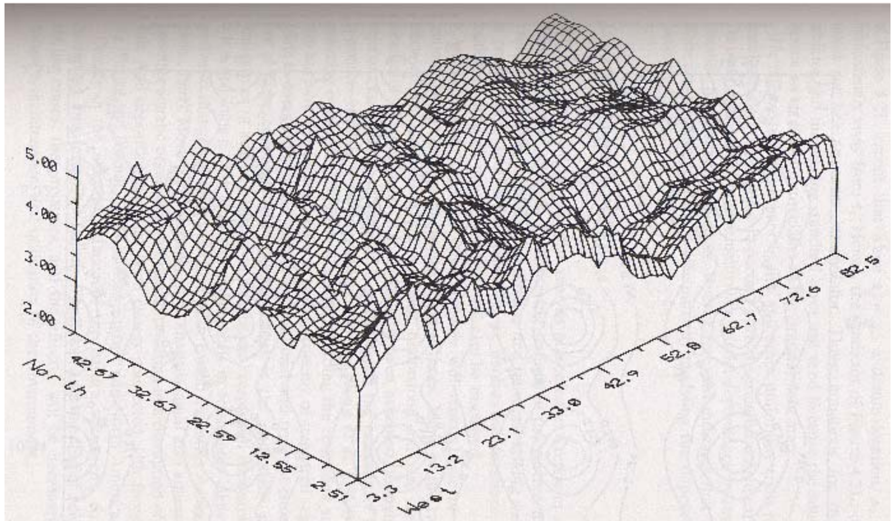


Concentration of nitrate in the water (1994)

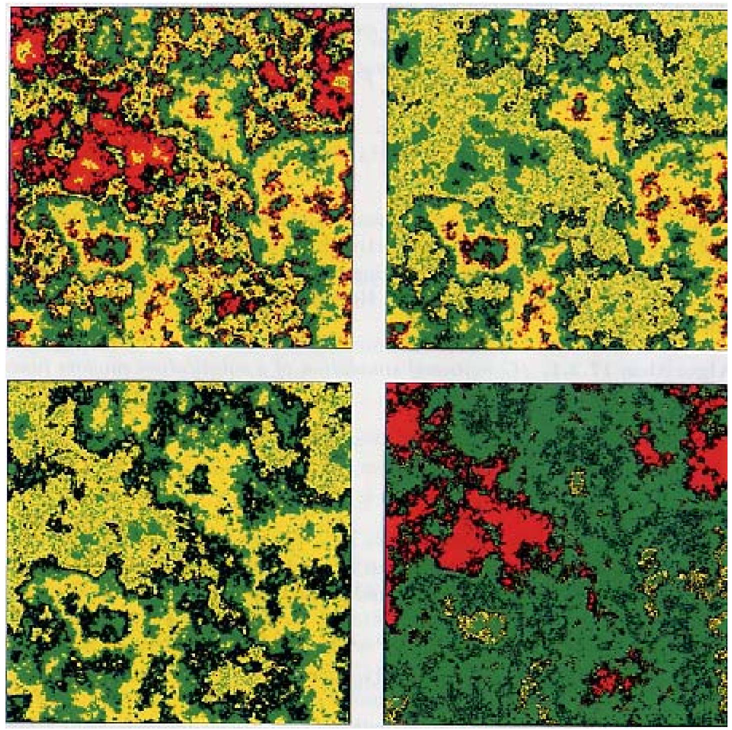
Random fields

- Random field $\{Z(x,\omega) : x \in \mathbb{R}^d, \omega \in \Omega\}$ is a family of random variables:
  - $Z(x,\cdot)$ is a random variable (briefly: $Z(x)$)
  - $Z(\cdot,\omega)$ is a realization of the random field Z (briefly: $z(x)$)
- $Z(x) = m(x) + Y(x)$ where $m(x) = E\,Z(x)$ is the mean (drift) and $Y(x) = Z(x) - m(x)$ is the deviation from the mean (residual). It holds $E\,Y(x) = 0\;\forall\,x$.
- Let $W \subset \mathbb{R}^d$ be an observation window (normally a rectangle)
- Let $x_1, \ldots, x_n \in W$ be measurement locations placed arbitrarily within W. Let $z(x_1), ..., z(x_n)$ be the measured values at these locations.

Random fields



A realization $\{z(x)\}_{x\in\mathbb{R}^2}$ of the random field: $z(x) =$ wheat harvest in pounds at x



Four simulated realizations of a random field

Stationary random fields

- A random field Z is called (strictly) stationary if all its finite dimensional distributions are translation invariant: for all $h \in \mathbb{R}^d$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in \mathbb{R}^d$ holds

$$(Z(x_1 + h), \ldots, Z(x_n + h)) \stackrel{\mathrm{d}}{=} (Z(x_1), \ldots, Z(x_n)).$$

- A random field Z is called stationary of second order if
  - $E\,(Z(x)) = m(x) = z = \text{const} \quad \forall x$
  - $\text{cov}(Z(x), Z(x+h)) = E\,[Z(x)Z(x+h)] - z^2 = C(h)$, the covariance function C exists and depends only on the difference h.

Stationary random fields

- Strict stationarity $\not\Longleftarrow \not\Longrightarrow$ stationarity of second order
- A second order stationary random field is called isotropic if $C(h) = C(|h|)$, $h \in \mathbb{R}^d$.
- A random field Z is mean square continuous (m.s.c.) if $E\left(Z(x) - Z(x_0)\right)^2 \to 0$, $x \to x_0$ for all $x_0 \in \mathbb{R}^d$.
- A second order stationary random field is m.s.c. $\Longleftrightarrow C(h)$ is continuous at $h = 0$.

Stationary random fields of second order

Covariance function

- C is positive definite: $\forall n \in \mathbb{N}$, $w_i \in \mathbb{R}$, $x_i \in \mathbb{R}^d$

$$\sum_{i,j=1}^{n} w_i w_j C(x_i - x_j) = \mathsf{Var}\left(\sum_{i=1}^{n} w_i Z(x_i)\right) \geqslant 0$$

- $|C(h)| \leqslant C(0) = \mathsf{Var} Z$

Stationary random fields of second order

Examples of covariance functions

- Nugget effect (white noise): $C(h) = c > 0$ for $|h| = 0$ and $C(h) = 0$, $|h| > 0$.
- Exponential model: $C(h) = be^{-|h|/a}$, where $b > 0$ is the sill and $a > 0$ is the range.
- Spherical model: for positive a and b

$$C(h) = \begin{cases} b\left(1 - 3/2|h|/a + 1/2|h|^3/a^3\right), & 0 \leqslant |h| \leqslant a, \\ 0, & |h| > a. \end{cases}$$

Stationary random fields of second order

Variogram

- $\gamma(h) \stackrel{def}{=} \frac{1}{2} E\left(Z(x+h) - Z(x)\right)^2$
- It holds $\gamma(h) = C(0) - C(h) = E\left[Z(x)^2\right] - E\left[Z(x)Z(x+h)\right]$, $\gamma(0) = 0$.
- γ is conditionally negative definite: for $n \in \mathbb{N}$, $w_i \in \mathbb{R}$ with $\sum\limits_{i=1}^{n} w_i = 0$ and $x_i \in \mathbb{R}^d$ it holds $\sum\limits_{i,j=1}^{n} w_i w_j \gamma(x_i - x_j) \leqslant 0$.
- γ is a variogram $\Longleftrightarrow \forall\, \lambda\; e^{-\lambda\gamma}$ is a covariance function.
- If $\gamma(h) \leqslant \gamma(\infty) < \infty$ for all h then $C(h) = \gamma(\infty) - \gamma(h)$ is a valid covariance function.

Stationary random fields of second order

Variogram

- If γ_1 and γ_2 are variograms then $\gamma = \gamma_1 + \gamma_2$ is a variogram as well.
- If Z is stationary and isotropic then $\gamma(h) = \gamma(|h|)$, $h \in \mathbb{R}^d$.
- Many istropic variogram models can be constructed using models for covariance functions. But not all of them: $\gamma(h) = b|h|^\alpha$, $b > 0$, $0 < \alpha < 2$.
- Anisotropic variogram models? e.g., geometrically anisotropic...

Stationary random fields of second order

Exponential geometrically anisotropic variogram

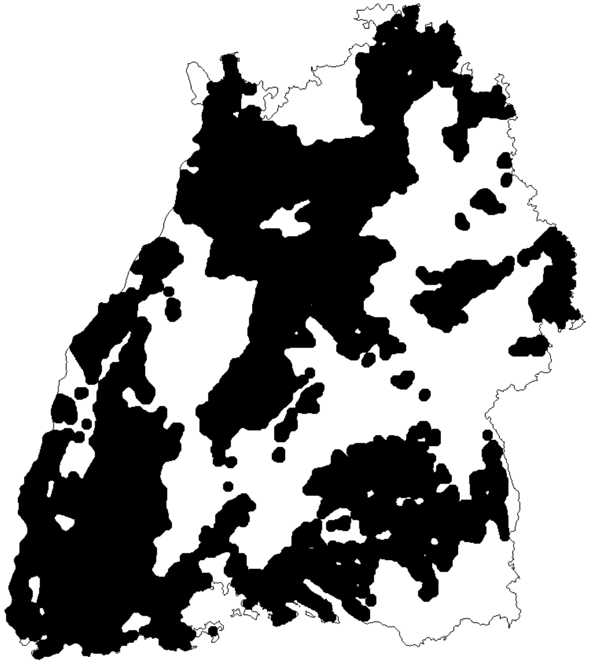
$$\gamma(h) = \begin{cases} 0, & h = 0, \\ a + b(1 - e^{-\sqrt{h^\top K h}/c}), & h \neq 0, \end{cases}$$

- Nugget effect a: discontinuity of the data at the microscopic scale
- Sill b: variability of the data at large distances h
- Range c: the correlation range of random variables $Z(x)$ and $Z(x+h)$
- K is the matrix of the composition of a rotation and a scaling.

Examples of random fields

- Random sets as random fields

Binary image: $Z(x) = I\{x \in \Xi\}$, $x \in \mathbb{R}^d$ for a random set Ξ.

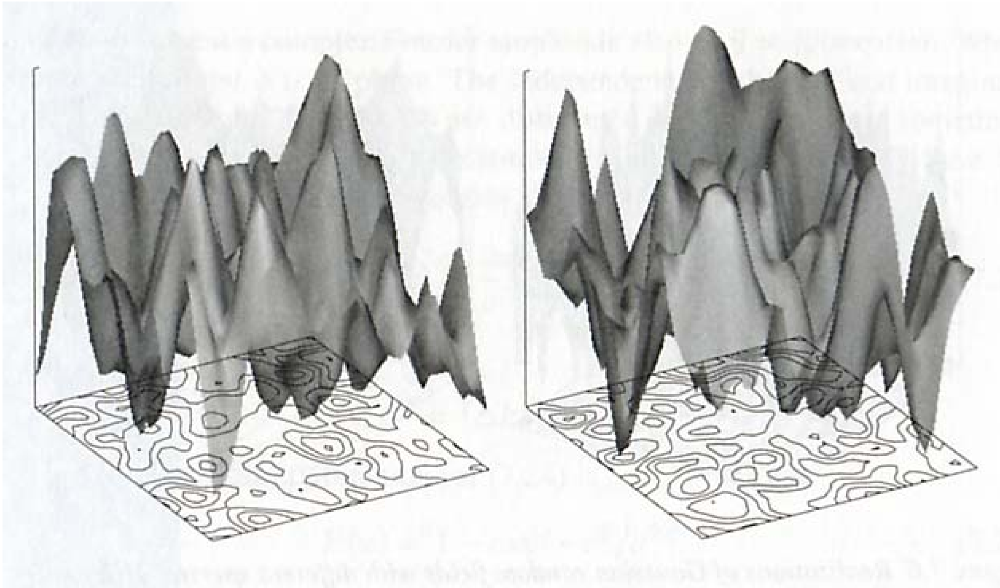


Difference map of nitrate concentrations in the ground water of Baden–Württemberg, 1993–1994

Examples of random fields

- Random fields induced by random sets
  - $Z(x) = V_j(\Xi \cap (W + x))$, $x \in \mathbb{R}^d$, $j = 0, \ldots, d$ where V_j are the intrinsic volumes
  - $Z_r(x) = \sum_{q \in \partial\Xi \cap B_r(x) \setminus \{x\}} J\big(\Xi \cap B_r(x), q, x\big)$, $x \in \mathbb{R}^d$
- Gaussian random fields: Z is a Gaussian random field if all its finite dimensional distributions are normal: for all $n \in \mathbb{N}$, $x_1, \ldots, x_n \in \mathbb{R}^d$ and $w_1, \ldots, w_n \in \mathbb{R}$ holds $w_1 Z(x_1) + \ldots + w_n Z(x_n) \sim N(\cdot, \cdot)$.

Examples of random fields

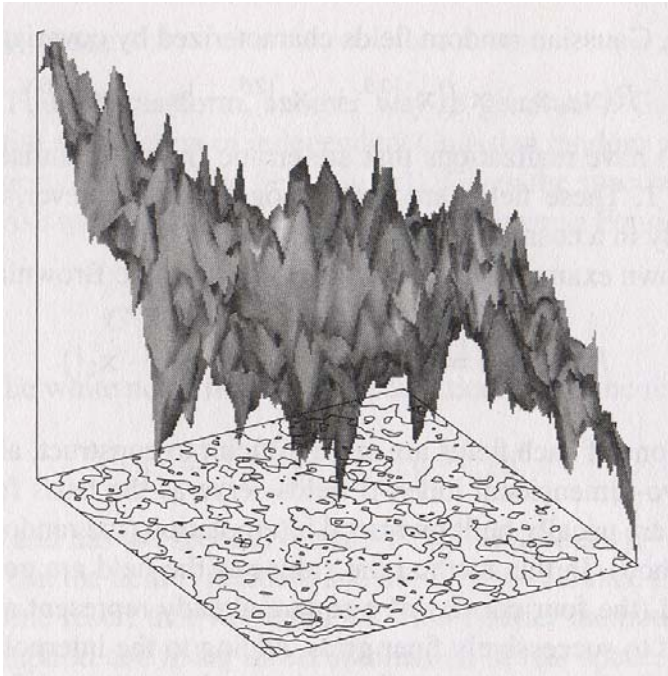


Two realizations of a Gaussian random field

Gaussian random fields are characterized uniquely by their drift $m(x) = E\,Z(x)$ and their covariance function C. Hence: strict stationarity $\Longleftrightarrow$ stationarity of second order.

Examples of random fields

Brownian field: $m(x) = 0$, $C(x, y) = 1/2(|x| + |y| - |x - y|)$, $\gamma(h) = 1/2|h|$



A realization of the Brownian field

Extrapolation of spatial data

Let Z be a stationary random field of second order with mean z, covariance function C and variogram γ.

Goal: Extrapolation of the field Z from the measured data $z(x_1), ..., z(x_n)$.

Estimation of $Z(x)$ $\forall x \Rightarrow$ ordinary kriging

Ordinary kriging

Construct the best linear estimator $\widehat{Z}(x) = \sum_{i=1}^{n} \lambda_i Z(x_i),\ x \in W$

where the weights λ_i satisfy the following requirements:

- $\widehat{Z}(x)$ is unbiased: $E\,\widehat{Z}(x) = z \Longrightarrow \sum_{i=1}^{n} \lambda_i = 1$
- The variance of the estimation error is minimal:
  $E[(\widehat{Z}(x) - Z(x))^2] \longrightarrow \min$

The minimization problem $\Longrightarrow$ solve the Lagrange equations:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \gamma(x_j - x_i) + \mu = \gamma(x - x_i), & i = 1, \ldots, n, \\ \sum_{j=1}^{n} \lambda_j = 1. \end{cases}$$

Ordinary kriging

To find the weights λ_i from the system of linear equations, the variogram $\gamma(h)$ has to be known or estimated from the data $z(x_1), \ldots, z(x_n)$.

- Matheron's estimator:
$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i,j: x_i - x_j \approx h} \left(Z(x_i) - Z(x_j)\right)^2,$$
$N(h)$ is the number of pairs $(x_i, x_j) : x_i - x_j \approx h$.
Computations are made for h on a grid in $\mathbb{R}^d$.
- $\hat{\gamma}(h)$ not conditionally negative definite $\Rightarrow$ a valid variogram model has to be fitted to $\hat{\gamma}(h)$

Extrapolation of spatial data

What if Z is not stationary? $Z(x) = m(x) + Y(x)$

$\Longrightarrow$ Estimation of the drift m and the residual Y

- Estimation of m: e.g. the moving average $\widehat{m}(x) = \frac{1}{N_x} \sum_{x_i \in R(x)} Z(x_i)$ where $R(x)$ is the neighborhood of x and $N_x = \#\{i : x_i \in R(x)\}$.
- Estimated residual $Y^*(x_i) = Z(x_i) - \widehat{m}(x_i)$
- Extrapolation of Y from the data $Y^*(x_1), \ldots, Y^*(x_n)$, e.g. by ordinary kriging provided that Y is stationary of second order.

Applications

- Synthetic data: disturbed Boolean models

  Let Ξ be a stationary Boolean model with intensity λ and deterministic rectangular primary grain $\Xi_0 = [a,b]^2$. Let $\xi = B_r(o)$ be a deterministic disturbance.

  - $m(x) = I(x \in \xi)$
  - $Y(x) = I(x \in \Xi) - p_\Xi$ where

$$p_\Xi = E\, I\{o \in \Xi\} = P(o \in \Xi) = 1 - e^{-\lambda|\Xi_0|} = 1 - e^{-\lambda ab}$$

  is the area fraction of Ξ.

Applications

Y is a stationary random field of second order with the covariance function

$$C(h) = 2p_\Xi - 1 + (1 - p_\Xi)^2 e^{\lambda|\Xi_0 \cap (\Xi_0 - h)|}$$

and the anisotropic variogram

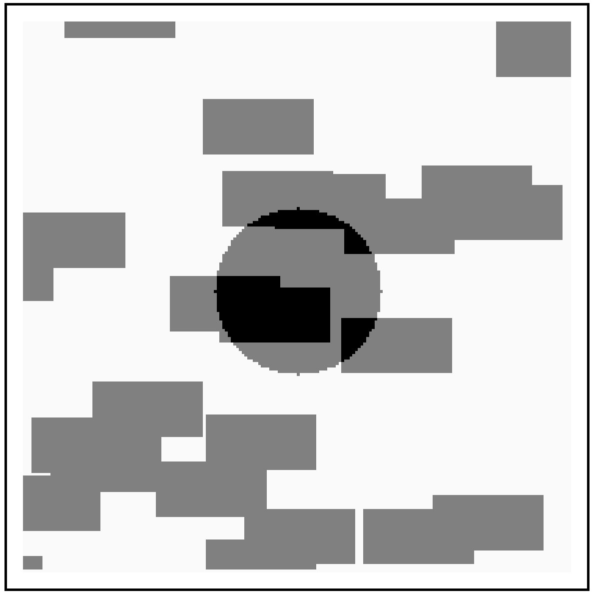
$$\gamma(h) = C(0) - C(h) = 1 - p_\Xi - (1 - p_\Xi)^2 e^{\lambda|\Xi_0 \cap (\Xi_0 - h)|}\,.$$

In the special case of $\Xi_0 = [a,b]^2$ it holds

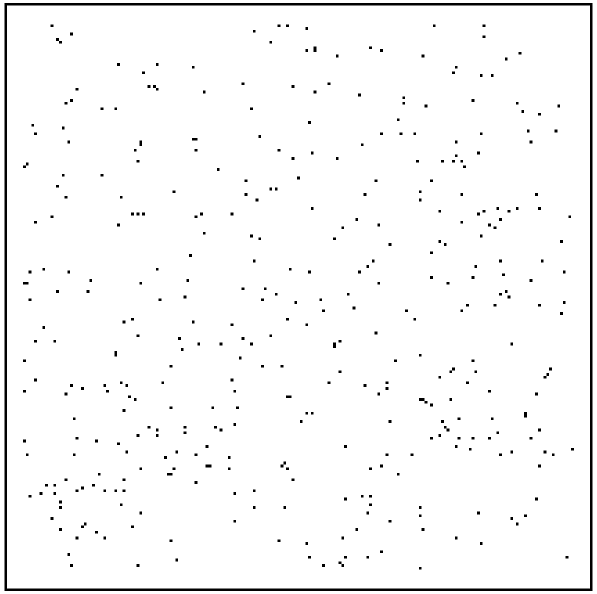
$$\gamma(h) = e^{-\lambda ab}\left(1 - e^{-\lambda(ab - |\Xi_0 \cap (\Xi_0 - h)|)}\right).$$

Applications

Disturbed Boolean models



Realisation of Ξ and ξ



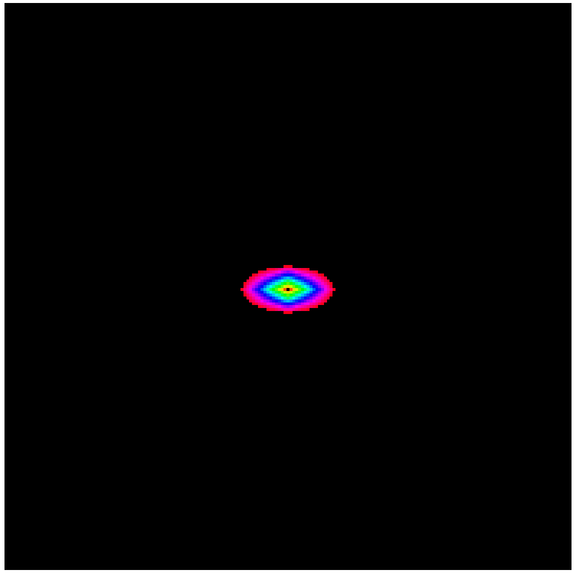
Measurement points

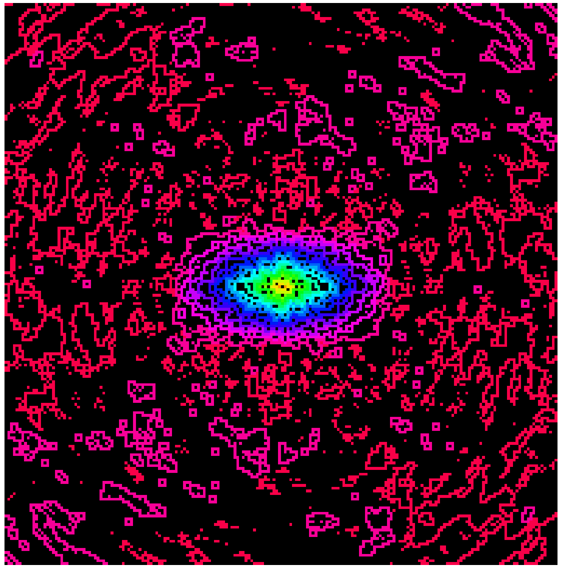
$$x_1 \dots x_n$$

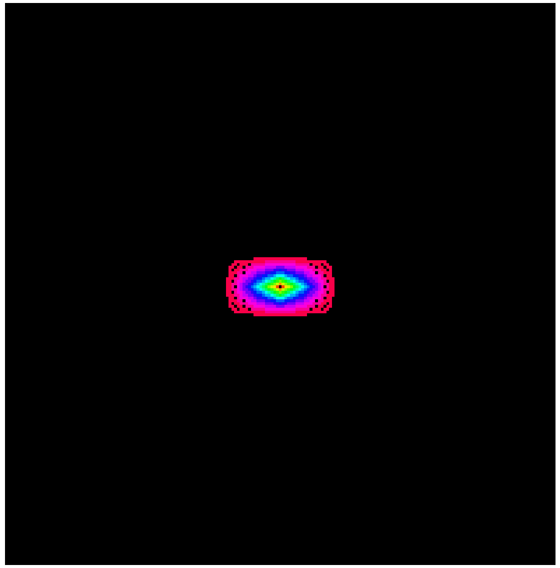
$$Z(x) = m(x) + Y(x) =$$
$$I(x \in \xi) + I(x \in \Xi) - p_\Xi$$

Applications

Disturbed Boolean models

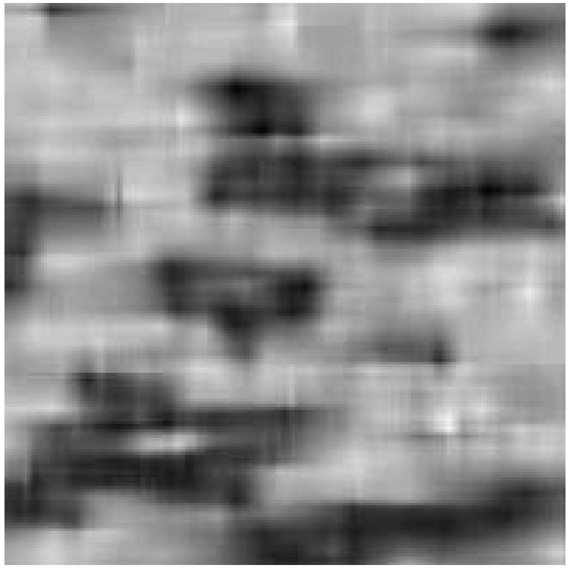

Theoretical variogram γ


Estimated variogram $\hat{\gamma}^*$

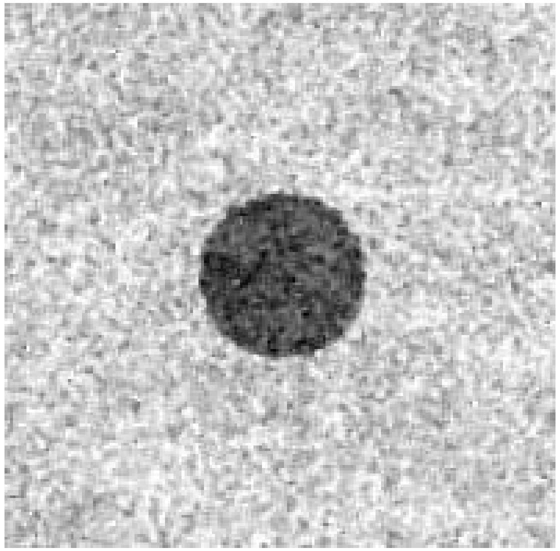

Fitted variogram γ^*

Applications

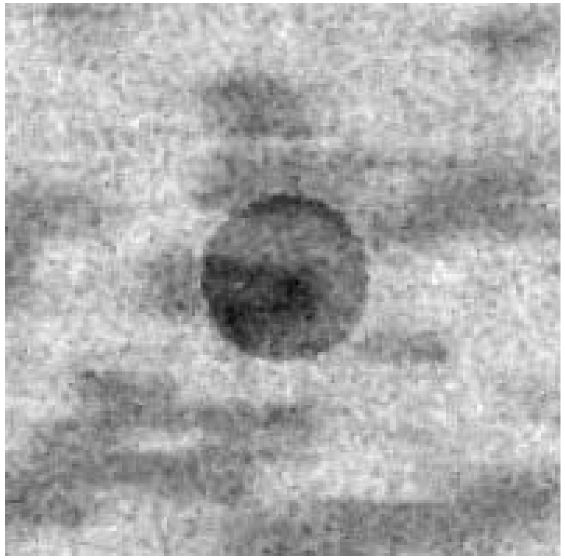
Disturbed Boolean models



+



=



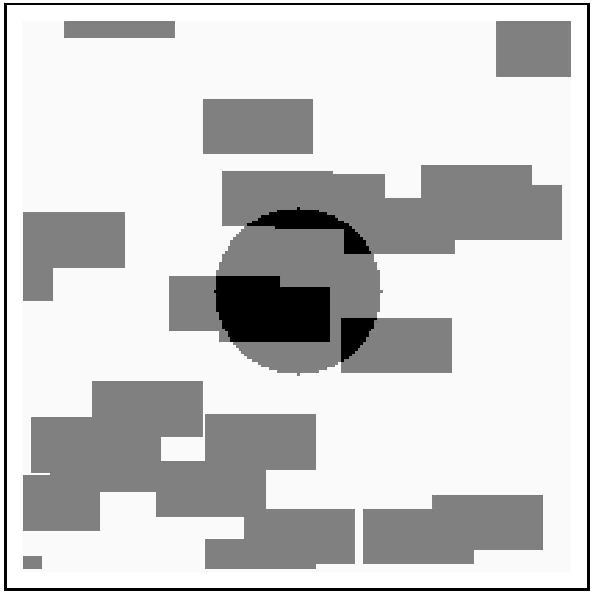
Estimated residual
$\widehat{Y}^*(x)$

Estimated drift
$\widehat{m}(x)$

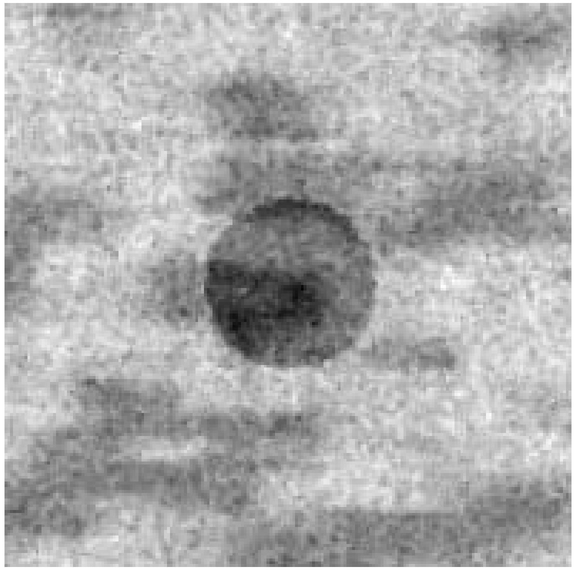
Extrapolated field
$\widehat{Z}(x) = \widehat{m}(x) + \widehat{Y}^*(x)$

Applications

Disturbed Boolean models



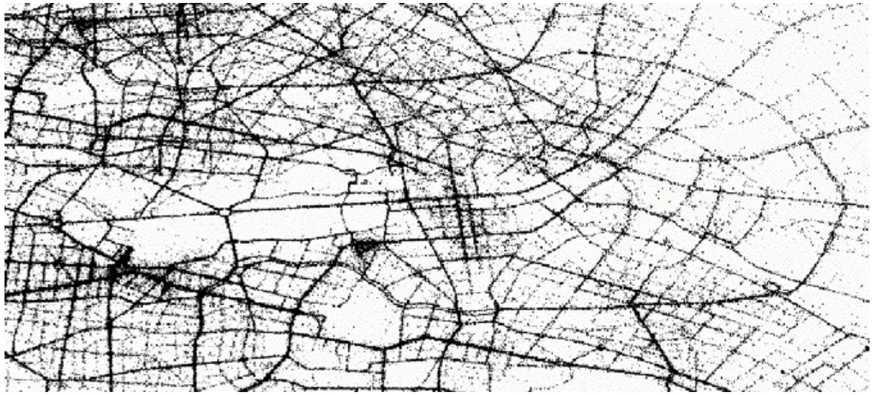
Realisation of
$Z(x) = m(x) + Y(x)$



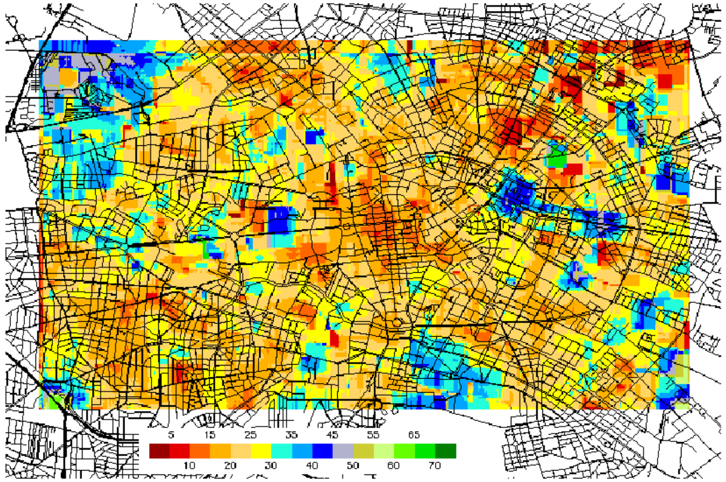
Extrapolated field
$\widehat{Z}(x) = \widehat{m}(x) + \widehat{Y}^*(x)$

Applications

- Making road traffic maps



Taxi positions in Berlin



Traffic velocities: 13.02.2002, 18.00–18.30

Literature

- Martinez, V. J., Saar, E. (2002) Statistics of the galaxy distribution. Chapman& Hall, London.
- Materon, G. (1975) Random sets and integral geometry. J. Wiley & Sons, New York.
- Ohser, J. and Mücklich, F. (2000) Statistical analysis of microstructures in materials science. J. Wiley & Sons, Chichester.
- Santalo, L. A. (1976) Integral geometry and geometric probability. Addison–Wesley, London.
- Schneider, R., Weil, W. (2000) Stochastische Geometrie. Teubner, Stuttgart.
- Stoyan, D., Kendall. W. S., Mecke, J. (1995) Stochastic geometry and its applications. J. Wiley & Sons, Chichester.

Literature

- Adler, R. J. (1981) The geometry of random fields. J. Wiley & Sons. Chichester.
- Chiles, J.-P., Delfiner. P. (1999) Geostatistics. Modelling spatial uncertainty. J. Wiley & Sons.
- Cressie, N. A. C. (1993) Statistics for spatial data. J. Wiley & Sons.
- Ivanov, A. V., Leonenko N. N. (1989) Statistical analysis of random fields. Kluwer.
- Wackernagel, H. (1998) Multivariate geostatistics. Springer.
- Yaglom, A. M. (1987) Correlation theory of stationary and related random functions. Springer.