Random Closed Sets

I. Molchanov

Department of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland

Summary. This survey highlights major contributions of George Matheron to developments of random sets theory. It outlines the state of the art of this theory before Matheron, other approaches parallel to Matheron's work and mentions some of more recent developments.

1 Early years of random sets

Concepts and results involving random sets appeared in probabilistic and statistical literature long time ago. The origin of the modern concept of a random set goes as far back as the seminal book by A.N. Kolmogorov [22] (first published in 1933) where he laid out the foundations of probability theory. He wrote [22, p. 46]

> Let G be a measurable region of the plane whose shape depends on chance; in other words, let us assign to every elementary event ξ of a field of probability a definite measurable plane region G.

In modern terminology, G is said to be a random set, which is not necessarily closed, see [37, Sec. 2.5]. It should be noted also that even before 1933 statisticians worked with confidence regions that can be naturally described as random sets.

The next major contribution was due to H.E. Robbins [43, 44] who discovered the formula

$$\mathbf{E}\mu(X) = \int \mathbf{P}\{x \in X\}\mu(\mathrm{d}x)\,, \tag{1}$$

which relates expectation of a σ-finite measure of a random set with the integral of its coverage function. Despite being a simple application of Fubini's theorem, Robbins' formula marked the first rigorous result concerning random sets. Actually, the special case of this formula appears already in [23] (published in 1933 in a physical journal).

For a while results concerning random sets remained scattered in the literature. A rapid development and growing interest in geometric probabilities in the late sixties called for formalisation of the concept of a random set. Developments in mathematical theories of cones and capacities by G. Choquet and the growing literature on set-valued functions greatly facilitated this task. On the other hand, advances in microscopy and image analysis stimulated appearance of new models for random sets and subsequent developments of statistical tools suitable for their formal analysis.

2 The definition of a random set

The crucial breakthrough done by G. Matheron was to concentrate on random sets with closed values and formally define them as random elements whose values belong to the family $\mathcal{F}$ of closed subsets of a given space E. The formal definition required endowing $\mathcal{F}$ with a σ-algebra generated by the topology on $\mathcal{F}$ that is now commonly known under the name of the Fell topology, see [4] and [9] for discussion of this and many other topologies on the space of closed sets. The idea behind the definition of a random closed set is that a random closed set is accessible through the knowledge of the fact whether or not it hits any given compact set. Matheron's definition of a random closed set in [30] is formulated as follows.

Definition 1. *A map $X : \Omega \to \mathcal{F}$ from a probability space $(\Omega, \mathfrak{F}, \mathbf{P})$ to the family $\mathcal{F}$ of closed subsets of a locally compact separable Hausdorff space E is called a* random closed set *if $\{X \cap K \neq \emptyset\} \in \mathfrak{F}$ for every K from the family $\mathcal{K}$ of compact subsets of E.*

The main idea behind Definition 1 is that it is observable if X hits or misses given deterministic sets. Because of this reason the underlying topology on $\mathcal{F}$ is sometimes called the hit-or-miss topology. Note that the book [30] rests on several previous publications of G. Matheron. In particular, in [26] the measurability issues for random closed sets are considered together with the first characterisation theorem for their distributions using the avoidance functional. The report [27] contains the concept of the hit-or-miss topology and the characterisation results for distributions of random closed sets using the capacity functionals.

From the contents of the book [30] it appears that G. Matheron was unaware of the parallel work in set-valued analysis, where E.G. Effros [7] studied the Borel σ-algebra on $\mathcal{F}$ for the case when the carrier space E is a complete separable metric space that is not necessarily locally compact. Since then it is typical to speak about Effros-measurable set-valued functions (also known as multivalued functions or correspondences). Another serious breakthrough was due to C. Himmelberg [19] who proved the following Fundamental Measurability theorem that established equivalence of various definitions of measurability for set-valued functions in Polish spaces.

Theorem 1. *Let E be a separable metric space and let X be a function on $(\Omega, \mathfrak{F}, \mathbf{P})$ with values in the family of closed subsets of E. Consider the following statements.*
(1) $\{\omega : X \cap B \neq \emptyset\} \in \mathfrak{F}$ for every Borel set $B \subset E$.
(2) $\{\omega : X \cap F \neq \emptyset\} \in \mathfrak{F}$ for every $F \in \mathcal{F}$.
(3) $\{\omega : X \cap G \neq \emptyset\} \in \mathfrak{F}$ for every open $G \subset E$ (in this case X is said to be Effros-measurable).
(4) $\varrho(y, X) = \inf\{\varrho(y, x) : x \in X\}$ is a random variable for each $y \in E$.
(5) There exists a sequence $\{\xi_n\}$ of E-valued random elements (measurable selections of X) such that X almost surely coincides with the closure of $\{\xi_n, n \geq 1\}$.
(6) The graph of X, i.e. the set $\{(\omega, x) \in \Omega \times E : x \in X(\omega)\}$, belongs to the product σ-algebra of $\mathfrak{F}$ and the Borel σ-algebra on E.
Then the following results hold.
(i) (1)⇒(2)⇒(3) ⇔ (4)⇒(6)
(ii) If E is a Polish space (i.e. E is also complete) then (3)⇔(5).
(iii) If E is a Polish space and the probability space $(\Omega, \mathfrak{F}, \mathbf{P})$ is complete, then (1)–(6) are equivalent.

Although it is possible to deduce numerous measurability results concerning operations with random sets from Theorem 1, it was G. Matheron who first realised the importance of semicontinuity concept for random closed sets. Many operations with sets are not continuous but only semicontinuous, so that measurability can be deduced by establishing semicontinuity of the corresponding maps. The semicontinuity concept also relates random closed sets to problems in stochastic optimisation [45, 46], where random closed sets naturally appear as epigraphs of lower semicontinuous functions.

3 Distributions of random sets

The next issue dealt by G. Matheron after formally defining a random closed set was to describe its distribution in an "economical" way. This is a highly important question since the Borel σ-algebra on $\mathcal{F}$ is so rich that it is infeasible to explicitly allocate probabilities to every event that belongs to it. This was a typically probabilistic question which is not usually dealt with in the set-valued analysis literature.

G. Matheron followed the traditional approach of constructing a probability measure by extending its values from a semi-algebra of sets to the corresponding σ-algebra. This semi-algebra consists of finite unions of the events $\{X \cap K \neq \emptyset\}$ for all $K \in \mathcal{K}$. Note that the families $\{F \in \mathcal{F} : F \cap K \neq \emptyset\}$ also generate the Borel σ-algebra on $\mathcal{F}$. In other words, G. Matheron found necessary and sufficient condition for a functional

$$T_X(K) = \mathbf{P}\{X \cap K \neq \emptyset\}, \quad K \in \mathcal{K},$$

to be extendible to a probability measure on $\mathcal{F}$. The functional T_X is called a *capacity functional* of X. It is sometimes called the hitting (or trapping) functional or plausibility functional. An immediate observation is that T_X is sub-additive, but typically non-additive (unless X is a random singleton), i.e. $T_X(K_1 \cup K_2)$ is less but not necessarily equal to $T_X(K_1) + T_X(K_2)$ for disjoint K_1 and K_2.

The capacity functional has several basic properties
(i) $T_X(\emptyset) = 0$ and $0 \le T_X(K) \le 1$ for every $K \in \mathcal{K}$;
(ii) T_X is upper semicontinuous on $\mathcal{K}$, i.e. $T_X(K_n) \downarrow T_X(K)$ if $K_n \downarrow K$;
(iii) T_X is completely alternating (also called alternating of infinite order), i.e. the following recurrently defined functionals

$$\Delta_{K_1} T_X(K) = T_X(K) - T_X(K \cup K_1)$$
$$\cdots$$
$$\Delta_{K_n} \cdots \Delta_{K_1} T_X(K) = \Delta_{K_{n-1}} \cdots \Delta_{K_1} T_X(K) - \Delta_{K_{n-1}} \cdots \Delta_{K_1} T_X(K \cup K_n)$$

are non-positive for every $n \ge 1$ and $K, K_1, \dots, K_n \in \mathcal{K}$.

It is easy to see that

$$\Delta_{K_n} \cdots \Delta_{K_1} T_X(K) = -\mathbf{P}\{X \cap K = \emptyset,\ X \cap K_1 \neq \emptyset, \dots, X \cap K_n \neq \emptyset\}\,,$$

so that condition (iii) generalises the monotonicity concept for multivariate cumulative distribution functions. Note that the above notation for the successive differences is taken from the harmonic analysis literature [5] and so differs from the notation used in [30].

A function ϕ on the family of all subsets of E with values in the extended real line is called a capacity if it is monotone, $M_n \uparrow M$ implies $\phi(M_n) \uparrow \phi(M)$ for arbitrary sets $M, M_n \subset E$, and $\phi(K_n) \downarrow \phi(K)$ if $K_n \downarrow K$ are compact sets. The above properties single out those capacities (obtained by extending T onto the family of all subsets of E) that correspond to distributions of random closed sets. The key result in random sets theory says that every functional T satisfying (i)-(iii) above corresponds to the distribution of a unique random closed set.

Theorem 2 (Choquet-Kendall-Matheron theorem). *Let $T : \mathcal{K} \mapsto [0,1]$. There exists a unique random closed set X with capacity functional T such that $\mathbf{P}\{X \cap K \neq \emptyset\} = T(K)$ for every $K \in \mathcal{K}$ if and only if T satisfies conditions (i)-(iii).*

G. Matheron [30] attributed this theorem to G. Choquet [6], where it indeed appears but not in such an explicit form. It was observed by D.G. Kendall [21] that the trapping probabilities define the distribution of a random set with not necessarily closed values provided the family $\mathcal{K}$ is replaced by another appropriately chosen family of trapping sets. G. Matheron gave a clear formulation of this theorem given above and provided an independent proof based exclusively upon the first principles of extending a probability measure from a semi-algebra to a σ-algebra. Later on with advances of harmonic analysis on semigroups [5] and the theory of lattices [11, 40] new proofs of this result appeared. Indeed, the family of closed sets is a semigroup and a lattice with the main operation being union. Within both of these frameworks, defining a measure on a semigroup or lattice is one of the key issues. The original Choquet's proof of Theorem 2 is based on a representation of positive definite functions on cones and is similar to the harmonic analysis proof outlined below.

Proof. The family $\mathcal{K}$ of compact sets is an Abelian semigroup with respect to the union operation. Let $\mathcal{I}$ be the set of all sub-semigroups I of $(\mathcal{K}, \cup)$, which satisfy

$$K, L \in I \;\Rightarrow\; K \cup L \in I \quad \text{and} \quad K \subseteq L,\; L \in I \;\Rightarrow\; K \in I\,.$$

Define $\tilde{K} = \{I \in \mathcal{I} :\; K \in I\}$ and equip $\mathcal{I}$ with the coarsest topology in which the sets $\tilde{K}$ and $\mathcal{I} \setminus \tilde{K}$ are open for all $K \in \mathcal{K}$. It is possible to prove that the map

$$c(I) = E \setminus \bigcup_{K \in I} \mathrm{Int} K$$

is continuous on $\mathcal{I}$ ($\mathrm{Int}K$ is the interior of K), and

$$c^{-1}(\mathcal{F}^K) = \bigcup_{L \in \mathcal{K},\; K \subset \mathrm{Int} L} \tilde{L}\,,$$

where $\mathcal{F}^K = \{F \in \mathcal{F} :\; F \cap K = \emptyset\}$. Indeed, $c(I) \cap K = \emptyset$ if and only if there exists $L \in I$ such that $K \subset \mathrm{Int}L$. It follows from (ii) that

$$Q(K) = \sup\{Q(L) :\; L \in \mathcal{K},\; K \subset \mathrm{Int}L\}\,,$$

where $Q(K) = 1 - T(K)$.

Note that $\mathcal{I}$ (with identical involution) is isomorphic to the set of semi-characters on $(\mathcal{K}, \cup)$, i.e. complex-valued maps on $\mathcal{K}$ satisfying $\chi(\emptyset) = 1$ and $\chi(K \cup L) = \chi(K)\chi(L)$. Property (iii) implies that T is a completely alternating function on $(\mathcal{K}, \cup)$. It is possible to prove that the corresponding function Q is negative definite on $\mathcal{K}$, i.e.

$$\sum_{i,j=1}^{n} a_j \bar{a}_k Q(K_j \cup K_k) \le 0$$

for any complex numbers $a_1, \ldots, a_n$, $n \ge 1$. By [5, Prop. 4.17], there exists a measure ν on $\mathcal{I}$ such that $Q(K) = \nu(\tilde{K})$. Now the continuity property of (Radon) measures ($\sup_\alpha \mu(G_\alpha) = \mu(\cup_\alpha G_\alpha)$ for upward filtering family of open sets G_α) yields

$$\nu(\cup_{L \in \mathcal{K},\; K \subset \mathrm{Int}L} \tilde{L}) = \sup\{\nu(\tilde{L}) :\; L \in \mathcal{K},\; K \subset \mathrm{Int}L\} = \nu(c^{-1}(\mathcal{F}^K))\,,$$

so that $Q(K) = \mu(\mathcal{F}^K)$, where μ is the image measure of ν under the continuous mapping $c : \mathcal{I} \mapsto \mathcal{K}$.

The uniqueness part follows from the fact that the families $\{F \in \mathcal{F} : F \cap K = \emptyset, F \cap K_1 \neq \emptyset, \dots, F \cap K_n \neq \emptyset\}$ generate the Borel σ-algebra on $\mathcal{F}$.

The lattice-theoretic proof [40] is even more powerful, since it is applicable for a non-Hausdorff space E. However, it is generally unknown how to characterise distributions of random closed sets in a space E that is either not locally compact or not separable, see [38].

4 Further developments

When random closed sets have been properly defined, their distributions characterised and measurability properties of some operations established, the theory of random sets was brought to a stage when it was desirable to obtain results parallel to those well-known in probability theory for random variables and stochastic processes. This was not easy however as the space of all closed (or compact) sets is not a linear space, while most of conventional techniques in probability theory are adapted to studies of random elements in linear spaces.

4.1 Special random sets and models

The capacity functional $T_X(K)$ defined for all compact sets K is a complicated object. In simple cases it is possible to define it using direct probabilistic arguments. For example, if $X = \{\xi\}$ is a random singleton, then $T_X(K) = \mathbf{P}\{\xi \in K\}$ is a probability measure that can be efficiently defined. More complicated examples of random sets appear from stochastic processes as excursion sets, e.g. $X = \{x \in E : \ \xi(x) \geq t\}$, where t is a real number and ξ is a real-valued random process indexed by E with upper semicontinuous paths (in order to ensure that X is closed). However, it remains an important task to develop new models for random sets, provide manageable expressions for their capacity functionals and relate properties of capacity functionals to those of the corresponding random closed sets.

For instance, it is possible to characterise random closed sets with almost surely convex realisations in terms of their capacity functionals. A random closed set X is almost surely convex if and only if

$$T_X(K) + T_X(K \cup K_1 \cup K_2) = T_X(K \cup K_1) + T_X(K \cup K_2) \qquad (2)$$

for every convex compact sets K, K_1 and K_2 such that K_1 and K_2 are separated by K in a sense that the segment joining any two points of K_1 and K_2 hits K, see [30].

G. Matheron introduced one extremely important model for random sets called the Boolean model. The basic idea follows the concept of a point process, which is a collection of points in a carrier space. The principal new feature is that the carrier space becomes the family of compact sets, so that one works with a collection of sets instead of collection of points, see Chap. 1.2 of this volume for an in-depth survey of this important model.

A Poisson point process $\{F_1, F_2, \ldots\}$ on $\mathcal{F}$ is determined by a measure on $\mathcal{F}$, that is not necessarily finite. Similarly to the Choquet theorem, it is possible to show that every such measure ν uniquely corresponds to a function $\Psi(K) = \nu(\{F \in \mathcal{F} : \ F \cap K \neq \emptyset\})$, which is upper semicontinuous, completely alternating, satisfies $\Psi(\emptyset) = 0$, but not necessarily is bounded by 1 from above and may even be infinite. Then

$$T_X(K) = 1 - \exp\{-\Psi(K)\} \tag{3}$$

is the capacity functional of a random closed set X that is the union of the sets $F - 1, F - 2, \ldots$ that form the underlying Poisson process on $\mathcal{F}$. If the functional Ψ satisfies (2), then X is called the semi-Markov random closed set. It was shown by G. Matheron [28] that these sets include Boolean models with convex grains and also unions of Poisson flats.

Many other important random closed sets are related to paths of a Brownian motion W_t, $t \geq 0$, or other stochastic processes with values in $\mathbb{R}^d$. Assume that W_t starts at x from a compact set D and is killed whenever it leaves D. Denote by X the path of W_t (the set of points visited at least once). This is an example of a random *fractal* set [8, Chap. 15,16]. The corresponding capacity functional is related to hitting probabilities of the Wiener process. For instance, if the initial position x is distributed according to the equilibrium probability measure on D, then $T_X(K)$ is the ratio $C(K)/C(D)$, where $C(\cdot)$ stands for the Newton capacity of the corresponding set. If $X_t = \{W_s : \ s \leq t\}$ is the part of the path up to time t, then its r-neighbourhood $X_t^r = \{x : \ \varrho(x, X_t) \leq r\}$ is called the Wiener sausage. Results on volumes of the Wiener sausage are closely related to the probability that the Wiener process hits obstacles that form a Boolean model, see [50].

Further recent results concern such concepts like capacity equivalence for random sets. Two random closed sets X and Y are called capacity equivalent if there are positive constants c_1 and c_2 such that

$$c_1 T_X(K) \leq T_Y(K) \leq c_2 T_X(K)$$

for every $K \in \mathcal{K}$. It is shown in [41] that the path of the Wiener process is capacity equivalent to a sequence of sets related to some branching processes on a tree generated by successive partitions of the unit square. This concept if closely related to Radon-Nikodym derivatives of capacities considered in [15].

4.2 Expectation

The concept of averaging for random sets was not mentioned at all in [30], although the relevant ideas in set-valued analysis (concerning integration of set-valued functions) appeared well before 1975 in R.J. Aumann's pioneering work [2]. Aumann defined the integral of a set-valued function F as the set of integrals of all measurable functions f such that $f(t) \in F(t)$ for all parameter values t. In application to random sets the idea of the corresponding expectation was first explicitly spelt out in [1]. The crucial step was to consider all random singletons ξ that almost surely belong to a random closed set X. Such ξ is called a selection of X. It is well-known that an almost surely non-empty random closed set possesses a selection under rather mild conditions on the carrier space E.

A random closed set X in a Banach space E is called integrable if it has at least one integrable selection. The selection expectation (also called Aumann or set-valued expectation) of an integrable random closed set X is defined as the closure of the set of expectations of all integrable selections of X

$$\mathbf{E}X = \mathrm{cl}\{\mathbf{E}\xi : \ \xi \in X \text{ a.s.}, \ \xi \text{ integrable}\} .$$

Taking closure in the right-hand side is essential as the family of all $\mathbf{E}\xi$ is not necessarily closed if the carrier space E is infinite dimensional.

Numerous results concerning the selection expectation include dominated convergence theorem and the Fatou lemma. It is possible to define the conditional expectation that leads to the concept of set-valued martingales [17]. However, the selection expectation has a serious drawback that reduces the range of its practical applications for averaging of sets. On a non-atomic probability space it always returns convex results, i.e. $\mathbf{E}X$ coincides with the expectation of the convex hull of X. Furthermore, if X is bounded, then $\mathbf{E}h(X,u) = h(\mathbf{E}X,u)$, where

$$h(K,u) = \sup\{u(x) : \ x \in K\}$$

is the support function of K, and u is a linear continuous functional on E.

Alternative definitions of expectation [34, 36] make it possible to work with non-convex set, although these expectations do not have so nice mathematical properties as the selection expectation.

4.3 Minkowski sums

Despite Minkowski addition and related morphological operations with sets were described and the corresponding measurability results established in [30], the corresponding limit theorems remained beyond the scope of G. Matheron's attention. These limit theorem were derived first for random compact sets in Euclidean spaces and then generalised for random closed sets in Banach spaces without the compactness and boundedness assumptions, see e.g. [14, 18]. For simplicity we consider here only the case of random compact sets in the Euclidean space $E = \mathbb{R}^d$. Recall that $K \oplus L = \{x + y : \ x \in K, \ y \in L\}$ denotes the Minkowski sum of K and L.

Let $X, X_1, X_2, \ldots$ be a sequence of independent identically distributed random compact sets in $\mathbb{R}^d$. Assume that X is integrably bounded, i.e. $\|X\| = \sup\{\|x\| :\ x \in X\}$ is an integrable random variable. In this case all selections of X are integrable. The strong law of large numbers for random compact sets [1] establishes

$$n^{-1}(X_1 \oplus \cdots \oplus X_n) \to \mathbf{E}X\,, \tag{4}$$

where the convergence is understood with respect to the Hausdorff metric ϱ_H. Recall that the Hausdorff distance between two sets K and L is the smallest positive r such that K is contained in the r-neighbourhood of L and L is contained in the r-neighbourhood of K.

The central limit theorem concerns the speed of convergence in (4). Since it is not possible to subtract $\mathbf{E}X$ from the Minkowski average in the left-hand side of (4), one usually formulates the central limit theorem for square integrable (i.e. $\mathbf{E}\|X\|^2 < \infty$) random sets as the following convergence in distribution property

$$\sqrt{n}\varrho_H(n^{-1}(X_1 \oplus \cdots \oplus X_n), \mathbf{E}X) \xrightarrow{\mathrm{d}} \sup_{\|u\|=1} |\zeta(u)|\,, \tag{5}$$

where ζ is a Gaussian random field on the unit sphere with the covariance explicitly determined by the distribution of the support function $h(X, \cdot)$ of X, see [51]. Related works include characterisation of stable and infinite divisible for Minkowski addition random sets [12, 13]. It should be noted that the limiting random field (5) does not have an explicit geometrical meaning, and it is an open problem to provide a sensible geometric interpretation of ζ.

4.4 Weak convergence

Weak convergence of random closed sets is a special case of weak convergence of probability measures. Along the same line with the Choquet-Kendall-Matheron theorem, it is possible to show that a sequence of random closed set $\{X_n,\ n \geq 1\}$ converges weakly to a random closed set X if and only if $T_{X_n}(K) = \mathbf{P}\{X_n \cap K \neq \emptyset\}$ converges to $T_X(K) = \mathbf{P}\{X \cap K \neq \emptyset\}$ as $n \to \infty$ for each $K \in \mathcal{K}$ satisfying $T_X(K) = T_X(\mathrm{Int}K)$. These sets K correspond to the families of closed sets $\{F \in \mathcal{F} :\ F \cap K \neq \emptyset\}$ that are continuous with respect to the probability measure on $\mathcal{F}$ that describes the distribution of X, see [39]. In an internal report [29] G. Matheron suggested an equivalent definition that relies on the convergence of the capacity functionals $\limsup T_{X_n}(K) \leq T_X(K)$ for all compact sets K and $\liminf T_{X_n}(G) \geq T_X(G)$ for all open sets G.

It is well-known that the weak convergence of random variables is metrised by the Lévy metric. The weak convergence of random compact sets can be also described using the Lévy distance between their distributions. For random closed sets X and Y define

$$\mathfrak{L}(X, Y) = \inf\{r > 0 :\ T_X(K) \leq T_Y(K^r) + r,\ T_Y(K) \leq T_X(K^r) + r,\ K \in \mathcal{K}\}\,.$$

It is shown in [33] that X_n weakly converges to a random compact set X if and only if $\mathfrak{L}(X_n, X) \to 0$ as $n \to \infty$.

4.5 Unions

While the Minkowski addition of random sets generalises conventional sums of random vectors, unions of random sets have their parallel interpretation in the studies of extremes of random variables. Consider a random half-line $X = (-\infty, \xi]$ and independent realisations $X_n = (-\infty, \xi_n]$, $n \geq 1$, of the random closed set X. Then $X_1 \cup \cdots \cup X_n$ is the half-line bounded by $\max(\xi_1, \ldots, \xi_n)$, while $X_1 \oplus \cdots \oplus X_n$ is the half-line bounded by $\xi_1 + \cdots + \xi_n$.

A random closed set X is infinite divisible for unions if, for every n, X can be represented as a union of n independent identically distributed random closed sets. Infinite divisible random closed sets were characterised by G. Matheron in [30] as having the capacity functional of the form (3), where Ψ is a capacity that satisfies the same conditions (i)-(iii) as T with the only exception of the range of values that can now be $[0, \infty]$. The infinite values of Ψ appear due to the fixed points x that belong to X almost surely. The modern proof of (3) for infinitely divisible sets rests on the theory of lattices, see [40].

If, for every $n \geq 1$, there is a number $a_n > 0$ such that $a_n X$ has the same distribution as $X_1 \cup \cdots \cup X_n$ for $X_1, \ldots, X_n$ being independent identically distributed realisations of X, the random set X is called union-stable. Union-stable sets are infinite divisible, and it is possible to show that (3) holds with an additional requirement that $\Psi(sK) = s^\alpha \Psi(K)$ for some $\alpha \neq 0$ and every $s > 0$, $K \in \mathcal{K}$. The case of X without fixed points [30] is much easier to handle than the general case. This is due to the fact that non-trivial random sets with fixed points (e.g. the set of zeroes for the Wiener process or a randomly rotated cone) may satisfy $T_X(sK) = T_X(K)$ for every $s > 0$ and so it is difficult to turn the union-stability condition into a functional equation for T_X, see [33].

The characterisation of union-stable sets is naturally accompanied by a spectrum of limit theorems where union-stable random closed sets appear as weak limits, see [33]. These limit theorems are typically formulated in terms of capacity functionals of random sets, e.g. using the function $f(x) = T_X(xK)$ that should be regularly varying at infinity (or zero) for a sufficiently large family of compact sets K.

4.6 Functionals of random sets

A measurable functional of a random closed set automatically becomes a random variable. The earliest result concerning functionals of random sets is Robbins' formula (1) that is applicable to relate the expectation of $\mu(X)$ for a general σ-finite measure μ to the covering probabilities $\mathbf{P}\{x \in X\}$ of X. While the assumption of σ-finiteness of μ is absolutely essential, it was apparently overlooked in [30]. However, many interesting functionals of X can be represented as values $\mu(X)$ for a not necessarily σ-finite μ. The most well-known examples of such functionals are the surface area and the cardinality of X. As the capacity functional is the ultimate characteristics of a random closed set, it is quite natural to conjecture that the expected value of $\mu(X)$ for a general measure μ can be expressed using the capacity functional of X. Although this problem remains open, some preliminary results can be found in [3].

Another family of functionals of random sets is closely related to problems that appear in stochastic optimisation. For a real-valued random function $\xi(x)$, $x \in E$, with almost surely lower semicontinuous realisations, its epigraph

$$\operatorname{epi}\xi = \{(x,t) :\ \xi(x) \le t\}$$

becomes a random closed set in $E \times \mathbb{R}$. The crucial point is to notice that the infimum of $\xi(x)$ for x from a compact set K is a random variable and the points, where this infimum is achieved, form another random closed set. This observation sparked a considerable activity in stochastic optimisation literature, see [45, 46]. For instance, to ensure weak convergence of infima, it suffices to prove that the sequence of the corresponding epigraphs converges weakly as random closed sets.

Further results on functionals of random sets rely on assuming particular models for random sets. Examples of these results are integral geometrical formulae for Boolean models of random sets [49], Boolean random functions [20, 48] and results for convex hulls of random points [47].

4.7 Statistics

Despite the fact that G. Matheron's book [30] does not deal explicitly with any statistical issue concerning random sets, the developed probabilistic tools formed a platform for further developments of statistical techniques. While statistical issues for point processes had been in the focus of attention of statisticians for quite a while before 1975, the first statistical paper on random sets [42] appeared later in 1977. It concerned estimation of a domain accessed though Poisson points inside it. The natural estimator is the convex hull of these Poisson points. This estimator is however biased and has to be rescaled to eliminate the bias. Modern developments in this problem are surveyed in [24].

Realisations of random sets are available through values of some functionals or numerical measurements. Determining of sets using values of functionals or measurements was initiated in [25] and further put into the framework of mathematical morphology in [16].

For general random sets, it is possible to build the empirical capacity functional in the same manner as empirical measures are defined. Let X_n, $n \ge 1$, be i.i.d. realisations of a random closed set X. Define the empirical capacity functional as

$$T_n^*(K) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \cap K \neq \emptyset}\,, \quad K \in \mathcal{K}\,.$$

The strong law of large numbers immediately implies that $T_n^*(K)$ converges almost surely to $T(K)$ for any given K. However, the uniform convergence may fail even over a simple family of sets K. For instance, let X be a random closed subset of $\mathbb{R}$ defined as $\xi + M$, where ξ is normally distributed (say) and M is a nowhere dense set of a positive Lebesgue measure. Then it is easy to see that $|T_n^*(\{x\}) - T(\{x\})|$ does not converge uniformly to zero over $x \in [0,1]$. The uniform convergence properties have been explored in [31], where it was shown that the empirical capacity functional converges uniformly over the family of all compact sets if X coincides almost surely with the closure of its interior, $\mathrm{Int}X$, and $\mathbf{P}\{\mathrm{Int}X \cap K \neq \emptyset\} = T(K)$ for each $K \in \mathcal{K}$. The corresponding central limit theorem and applications to estimation of quantiles of random sets were discussed in [32]. Further results on weak convergence of families of probability measures dominated by empirical capacity functionals can be found in [10].

While statistical techniques for general random closed sets are still quite scarce, more is known for particular models of random sets. This concerns, in particular, the Boolean model, where a range of statistical tools exists [35, 49], and union-stable random sets [33]. Statistical techniques for random sets commonly rely on minimisation of minimum contrast or method of moments. This means that parameters are estimated by matching moments of some functionals of the sample with the moments calculated (theoretically, numerically or by simulations) for the underlying model. Approaches based on likelihood are understandably quite complicated to work out, since the complete likelihood function is very difficult to write even for models based on the Poisson assumption.

5 Final remarks

The range of citation of Matheron's random sets book [30] is extremely wide and stretches far beyond the literature specifically concerned random sets. Apart from a tremendous impact on mathematical morphology and image analysis, its random sets chapters have been cited by many authors who wrote on harmonic analysis on semigroups, sample paths properties of stochastic processes, set-indexed processes, set-valued analysis, stochastic optimisation and integral geometry. The up-to-date state of the random sets theory is presented in [37].

Matheron's book on random sets left enough open ends in random sets theory to ensure its fruitful development for nearly thirty years. I am pleased to note that this book was translated into Russian and published in 1978, very soon after its English edition appeared in 1975. The translator, V.P. Nosko, and the editor, V.M. Maksimov, did a great job going through the uneasy text and supplying their comments to occasional unclear or difficult places. When it came to writing my second year undergraduate project in 1980, I discovered the Russian translation of Matheron's book in a book shop and was impressed by the way topology, convex geometry and probability theory merge together. And I am still fascinated by it.

References

1. Artstein, Z. and Vitale, R. A. (1975). A strong law of large numbers for random compact sets. *Ann. Probab.* **3**, 879–882.
2. Aumann, R. J. (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1–12.
3. Baddeley, A. J. and Molchanov, I. S. (1997). On the expected measure of a random set. In D. Jeulin, editor, *Advances in Theory and Applications of Random Sets*, 3–20, Singapore. Proceedings of the International Symposium held in Fontainebleau, France (9-11 October 1996), World Scientific.
4. Beer, G. (1993). *Topologies on Closed and Closed Convex Sets*. Kluwer, Dordrecht.
5. Berg, C., Christensen, J. P. R. and Ressel, P. (1984). *Harmonic Analysis on Semigroups*. Springer, Berlin.
6. Choquet, G. (1953/54). Theory of capacities. *Ann. Inst. Fourier* **5**, 131–295.
7. Effros, E. G. (1965). Convergence of closed subsets in a topological space. *Proc. Amer. Math. Soc.* **16**, 929–931.
8. Falconer, K. J. (1990). *Fractal Geometry*. Wiley, Chichester.
9. Fell, J. M. G. (1962). A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space. *Proc. Amer. Math. Soc.* **13**, 472–476.
10. Feng, D.-J. and Feng, D. (2004). On a statistical framework for estimation from random set observations. *J. Theoret. Probab.* **17**, 85–110.
11. Gierz, G., Hofmann, K. H., Keimel, K., Lawson, J. D., Mislove, M. and Scott, D. S. (1980). *A Compendium of Continuous Lattices*. Springer, Berlin.
12. Giné, E. and Hahn, M. G. (1985). Characterization and domains of attraction of p-stable compact sets. *Ann. Probab.* **13**, 447–468.
13. Giné, E. and Hahn, M. G. (1985). M-infinitely divisible random sets. *Lect. Notes Math.* **1153**, 226–248.
14. Giné, E., Hahn, M. G. and Zinn, J. (1983). Limit theorems for random sets: application of probability in Banach space results. In A. Beck and K. Jacobs, editors, *Probability in Banach spaces, IV (Oberwolfach, 1982)*, volume 990 of *Lect. Notes Math.*, 112–135. Springer, Berlin.
15. Graf, S. (1980). A Radon-Nikodym theorem for capacities. *J. Reine Angew. Math.* **320**, 192–214.
16. Heijmans, H. J. A. M. and Molchanov, I. S. (1998). Morphology on convolution lattices with applications to the slope transform and random set theory. *J. Math. Imaging and Vision* **8**, 199–214.
17. Hess, C. (1999). Conditional expectation and martingales of random sets. *Pattern Recognition* **32**, 1543–1567.
18. Hess, C. (1999). The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive Banach spaces. *J. Convex Analysis* **6**, 163–182.

19. Himmelberg, C. (1974). Measurable relations. *Fund. Math.* **87**, 53–72.
20. Jeulin, D. (1997). *Morphological Models of Random Structures*. CRC Press, Boca Raton, Florida.
21. Kendall, D. G. (1974). Foundations of a theory of random sets. In E. F. Harding and D. G. Kendall, editors, *Stochastic Geometry*, 322–376. Wiley, New York.
22. Kolmogorov, A. N. (1950). *Foundations of the Theory of Probability*. Chelsea, New York.
23. Kolmogorov, A. N. and Leontovitch, M. A. (1992). On computing the mean Brownian area. In A. N. Shiryaev, editor, *Selected works of A. N. Kolmogorov, Volume II: Probability and mathematical statistics*, volume 26 of *Mathematics and its applications (Soviet series)*, 128–138. Kluwer, Dordrecht, Boston, London.
24. Korostelev, A. P. and Tsybakov, A. B. (1993). *Minimax Theory of Image Restoration*. Springer, New York.
25. Lyashenko, N. N. (1983). Statistics of random compacts in the Euclidean space. *J. Soviet Math.* **21**, 76–92.
26. Matheron, G. (1967). *Eléments pour une Théorie des Milieux Poreux*. Masson, Paris.
27. Matheron, G. (1969). Théorie des ensembles aléatoires. Technical report, Les Cahiers du Centre de Morphologie Mathematique, Fascicule 4, Paris School of Mines.
28. Matheron, G. (1972). Ensembles fermés aléatoires, ensembles semi-Markoviens et polyèdres poissoniens. *Adv. Appl. Probab.* **4**, 508–541.
29. Matheron, G. (1975). La convergence en loi des fermés aléatoires. Technical Report Internal Report N-409, Paris School of Mines, Fontenebleau.
30. Matheron, G. (1975). *Random Sets and Integral Geometry*. Wiley, New York.
31. Molchanov, I. S. (1987). Uniform laws of large numbers for empirical associated functionals of random closed sets. *Theory Probab. Appl.* **32**, 556–559.
32. Molchanov, I. S. (1990). Empirical estimation of distribution quantiles of random closed sets. *Theory Probab. Appl.* **35**, 594–600.
33. Molchanov, I. S. (1993). *Limit Theorems for Unions of Random Closed Sets*, volume 1561 of *Lect. Notes Math.*. Springer, Berlin.
34. Molchanov, I. S. (1997). Statistical problems for random sets. In J. Goutsias, R. Mahler and H. T. Nguyen, editors, *Applications and Theory of Random Sets*, volume 97 of *The IMA Volumes in Mathematics and its Applications*, 27–45, Berlin. Springer.
35. Molchanov, I. S. (1997). *Statistics of the Boolean Model for Practitioners and Mathematicians*. Wiley, Chichester.
36. Molchanov, I. S. (1998). Random sets in view of image filtering applications. In E. R. Dougherty and J. Astola, editors, *Nonlinear Filters for Image Processing*, chapter 10. SPIE.
37. Molchanov, I. (2005). *Theory of Random Sets*. Springer, New York.
38. Nguyen, H. T. and Nguyen, N. T. (1998). A negative version of Choquet theorem for Polish spaces. *East-West J. Math.* **1**, 61–71.
39. Norberg, T. (1984). Convergence and existence of random set distributions. *Ann. Probab.* **12**, 726–732.
40. Norberg, T. (1989). Existence theorems for measures on continuous posets, with applications to random set theory. *Math. Scand.* **64**, 15–51.
41. Peres, Y. (1996). Intersection-equivalence of Brownian paths and certain branching processes. *Comm. Math. Phys.* **177**, 417–434.

42. Ripley, B. D. and Rasson, J.-P. (1977). Finding the edge of a Poisson forest. *J. Appl. Probab.* **14**, 483–491.
43. Robbins, H. E. (1944). On the measure of a random set. I. *Ann. Math. Statist.* **15**, 70–74.
44. Robbins, H. E. (1945). On the measure of a random set. II. *Ann. Math. Statist.* **16**, 342–347.
45. Rockafellar, R. T. and Wets, R. J.-B. (1998). *Variational Analysis*. Springer, Berlin.
46. Salinetti, G. and Wets, R. J.-B. (1986). On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima. *Math. Oper. Res.* **11**, 385–419.
47. Schneider, R. (1988). Random approximations of convex sets. *J. Microscopy* **151**, 211–227.
48. Serra, J. (1989). Boolean random functions. *J. Microscopy* **156**, 41–63.
49. Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and its Applications*. Wiley, Chichester, second edition.
50. Sznitzman, A.-S. (1998). *Brownian Motion, Obstacles and Random Media*. Springer, Berlin.
51. Weil, W. (1982). An application of the central limit theorem for Banach-space-valued random variables to the theory of random sets. *Z. Wahrsch. verw. Gebiete* **60**, 203–208.

Index