PREFACE

This book is exceptional on more than one account. First, it was never published, although it is the source of hundreds of papers, theses, and books on geostatistics. But, curiously, until now, seminal text never existed, formally speaking. Another exceptional feature is also worth mentioning. When Georges Matheron wrote this work, he considered the theory of linear geostatistics complete and thought that this book would be the final one on the subject. Indeed, the branch of geostatistics he developed later on, about changes of support, is based on non-linear tools. However, What Matheron considered to be an end point turns out to have been the starting point for most of the readers. Some of them have had direct access to the book, but most of them discovered its content indirectly, via the many borrowings one can find in several books on geostatistics.

In such a situation, it is always instructive to come back to the original document, where the author's motivations, his physical intuitions, and his thoughts on the meaning of what he does, are detailed in most pages. We are motivated to renew readership of the book, to extend it to a larger audience, and to provide a more accessible version of the original. The result is this first formal edition of the book.

Before commenting on the text itself, we would like to briefly recall the professional course which led to the book. Georges Matheron landed in Algeria, with his wife and child, as an engineer at BRGA[1] in 1954. He was twenty-three. He quickly took over the scientific management of the BRGA (in 1956), and then the general management (1958/1959) of the BRMA[2]. It takes imagination to realise how important this was for a young French engineer. The huge Algerian territory has an area of 2, 400, 000 km^2, that is, half the area of the European Union, and stretches as far as the Saharan South. It abounds in ore bodies of all kinds. This is one of the reasons why the International Geological Congress was held in Algiers in 1952, just two years before G. Matheron arrived. It was also in the early fifties that papers written by three South African authors, Krige, Sichel, and de Wijs, laid the statistical foundations on which G. Matheron would base his revolutionary theory of geostatistics.

In the early sixties, French Algeria collapsed, France recalled its executives back to the home country and reorganised its mining research by creating the BRGM (Geological and Mining Survey) in Paris. G. Matheron was assigned a 'geostatistical department' with one single member, himself. At that time, geostatistics was not well accepted. The BRGM did not believe in it, and the first mining partners were the CEA (Atomic Energy Commission), and the iron mines of Lorraine (east of France). This solitude was, in fact, a blessing which enabled G. Matheron to devote himself to the final development of what would be later

[1] BRGA: *Bureau de Recherches Géologiques d'Algérie,* i.e. Geological Survey of Algeria.

[2] BRMA: *Bureau de Recherches Minières d'Algérie,* i.e. Mining Survey of Algeria.

Georges Matheron in 1958, at the time he invented geostatistics.

called linear geostatistics. This isolation ended in 1968, when he could create the Centre of Mathematical Morphology (including geostatistics) at the École des Mines de Paris.

The starting point of geostatistics is the de Wijsian formula, which gives the variance of a sample of size v moving in a field of size V by the relation

$$\sigma(v|V) = \alpha(\log V - \log v).$$

It was discovered in the gold ore body of the Witwatersrand, in South Africa. The mining of this huge placer deposit provided the grade in gold in elementary squares of $10 \times 10\,\mathrm{m}^2$ over kilometres, so that the experimental basis of the relation, established on hundreds of thousands of measurements, is indisputable. It teaches us first that the variance does not depend on the number of samples, as in classical statistics. If you continuously move a small square around in a larger one, and you measure the grades at an infinity of locations, their variance does not tend towards zero but towards the logarithmic ratio of the two areas of the squares, up to a constant. Second, this variance increases infinitely with the size V of the field. No bound appears in the experimental domain. Third, the formula reduces to a difference involving a logarithm. Is that particular to the Witwatersrand deposit, or is this circumstance general?

Matheron's answer to this last question launched geostatistics. He introduced the extension variance (equation 2.14), the de Wijsian formula being a particular case, and he separated, in the extension variance, the roles of the sampling pattern from that which is intrinsic to the mineralisation. The sampling pattern appears as domains of integration, and the intrinsic function, or variogram $\gamma(h)$, as the quantity to integrate. The logarithm, in de Wijsian formula, comes from the choice of the specific variogram $\gamma(h) = \alpha \log h$, although equation 2.14 is a general expression.

Matheron invented both extension variance and the variogram in Algeria during the fifties. Both are based on increments only (if you add a constant to the grade, they do not change), and both get around the trouble with possible infinite variances in infinite fields. They summarise the core notions of the theory, and lead to kriging, by minimising extension variances. Matheron had preserved from Algeria a huge record of experimental results, measurements, and estimations, which he used in the first volume of his *Traité de Géostatistique Appliqué* (*Treatise on Applied Geostatistics*) in 1962. One year later, the second volume appeared. It was devoted to kriging, a theory of cartography initially written to solve problems of local estimations raised by the uranium deposits of the CEA.

However, he was sceptical about the underlying probabilistic framework. Every deposit is a unique phenomenon that occurred once only in geological timescales. This phenomenon by itself does not offer any more hold on probabilities than if one wanted to know the proportion of hearts in a deck of cards by drawing only one card once. However, the stochastic approach works. Take, for instance, the rule of one-to-one correspondence given in Section 1.3. It associates each term of the limited expansion of the theoretical variogram (whose estimation is empirically accessible), with a term of the limited expansion of the estimation variance (which is sought for). Now, this rule can be equivalently formulated in a deterministic framework or in probabilistic terms. This surprising similarity led him to a third book, which was entitled *La Théorie des Variables Régionalisées et Leur Estimation*, which became his PhD thesis in 1965. The deterministic and random parts were developed successively, with much rigour, and he showed that randomness arrives surreptitiously, in the interpolation between the origin and the first experimental point of the variogram or the transitive covariogram (see Section 1.4).

This sequence of events from Matheron's life gives us a better appreciation of his style, pedagogical choices, and developments in the book. As the document is a set of lectures, most of the mathematical technicalities that one can find elsewhere are removed. This includes the proof of the rule of correspondence, or the passage to continuous formalism in universal kriging (e.g. as in equation 5.5).

For the same reason, no practical case study is described. Emphasis is put on the major results, like estimation variances, or kriging systems, which are established by simple proofs. At the same time, the reader is invited to wonder about the physical meaning of the notions Matheron deals with. We already quoted the ambiguous status of randomness; infinity also seems strange. A physical phenomenon is always finite, of course. But when you toss a coin, the plot of the cumulative results is a random function with a linear variogram, which means a lack of covariance and an infinite variance. And, if you restrict yourself to local variances and covariance, that is, to a finite number of games, then you strongly distort the phenomenon (see the superb Exercise 3.11.18 in Chapter 3).

Universal kriging dates from the end of the sixties and is Matheron's response to a question set by the hydrographic service of the French Navy, which wanted to draw bathymetric maps of the seas near the coasts. It starts from the observation that $\gamma(h)$ is defined modulo a constant, and replaces the constant by a linear term, or a more complex one. This permits interpolations and cartography in presence of a drift, as in bathymetry. In the book, Matheron put it together with three other chapters of a mining origin, because they all deal with linear estimators.

The chapters of the text are based on the lectures of a summer course given by G. Matheron in 1970. Initially written in French, they were translated into English by the PhD student Charles Huijbregts and were never published as a book. We have remained faithful to these original notes. But we tried to help the reader by giving a common structure to the chapters and sections, numbering equations sequentially within each chapter, numbering figures sequentially, redrawing most of them, and adding captions. In the exercises, we have separated the statements from the solutions, or suggestions for the solution, given by Matheron. While editing the notes, some ideas on new research lines came up, such as reviewing the principles underlying the theory of regionalised variables, taking into account recent advances in compositional data analysis. After so many years, and so many hours devoted to the field by so many researchers, the notes are still inspiring new ideas. We think this is the highest recognition a theory can achieve.

Vera Pawlowsky-Glahn and Jean Serra
Pals and Fontainebleau, 23 June 2016