23

# Geostatistics

R Webster

## Introduction

Geostatistics began as a body of statistical techniques for estimating the amounts of metals in ore-bearing rocks. Late in the 1970s, pedologists realised that the same technology could be adapted to soil survey; they realised that they could make maps of individual soil properties without first having to classify the soil and getting embroiled in all the doubts and controversy that that entailed. The maps are derived by 'contouring' from dense grids of values estimated from more or less sparse sample data. The procedure for estimation is known as 'kriging', after DG Krige who developed it empirically in the South African gold fields (see Krige 1966). Kriged estimates are unbiased and of minimum variance, and in this sense they are optimal. Theory and practice over the last 25 years has provided earth scientists with a sound technology that can be readily applied for estimating and mapping land resources.

This chapter describes briefly the underlying theory and assumptions of geostatistics. It then tackles the practical steps a surveyor must take to apply the technology. Geostatistics is best suited to intensive studies of small regions where spatially dense sampling is feasible, with sites being located within the range of spatial dependence for each attribute. Larger regions will inevitably include landscapes with diverse histories and the contrasting patterns of soil variation will demand the determination of several sample variograms.

## Theory

The practice of geostatistics is based on the theory of random functions, or stochastic processes. In this theory a variable at any point $\mathbf{x}$ on the land surface possesses not a single value but many values. Most environmental variables, such as the phosphorus content of soil or the salinity of groundwater, are continuous and so the distribution at $\mathbf{x}$ has an infinite number of values. In the real world, there is only one value at $\mathbf{x}$, and the theory treats this value, the value observed there, as just one drawn at random from some probability distribution according to some law.

### Some technical terms and notation

A consistent notation is needed to pursue the theory and for the equations later in the chapter. The location of a place on the Earth's surface is denoted by the vector $\mathbf{x}$, comprising two Cartesian coordinates, $x_1$ for (say) eastings, and $x_2$ for northings. Thus, $\mathbf{x} \equiv \{x_1, x_2\}$. The random variable at $\mathbf{x}$ is denoted $Z(\mathbf{x})$ (notice the capital), and its realisation is denoted $z(\mathbf{x})$ (now with a lower case $z$). The lower case form $z(\mathbf{x})$ is also used for an observation.