

Exploring Data-Rich Sources of Environmental Information using Geostatistics

Margaret A. Oliver

Department of Soil Science, University of Reading, Whiteknights, Reading RG6 6DW, United Kingdom

m.a.oliver@rdg.ac.uk

Abstract.

In general environmental and earth scientists have only sparse data for analysis. Scientists have faced the problems of prediction from such data for a long period before Matheron developed the Theory of Regionalized Variables. There are now many sources of intensive data, however, and these are at increasing fine spatial resolutions. They include digital image data, digital elevation models (DEM), electromagnetic induction information (EMI), yield monitoring data, ground penetrating radar and so on. Since optimal predictions are an important aim of geostatistical analysis there seemed to be no reason to analyse such sources of information geostatistically. However, the surfeit of information together with interference during recording, can make it difficult to identify and interpret any structures present. Identifying these aids the interpretation of images and can provide insight into factors controlling the observed variation. In addition it is likely that more than one scale of spatial variation will be present. If this is so it will be apparent in the variogram. It is then possible to describe the variogram as the sum of separate components or sources of variation.

The components of the variogram can be used to estimate the contributions of the individual sources to the whole picture using a procedure developed by Matheron (1982) known as 'kriging analysis' or factorial kriging. The merit of this method is that it allows the variation at each evident scale to be treated in turn as the signal, and to be separated from variation at all other evident scales, which is treated as noise. What is regarded as noise at one scale may be the signal at another. The analysis of nested spatial variation by factorial kriging will be illustrated with data from a SPOT image, a DEM, an EMI scan and crop yield.

Another feature of intensive data, especially digital elevation and EMI data, is the presence of trend or deterministic variation. This violates the assumptions of geostatistics, which is based on a random function model. The deterministic component of the variation must be removed and further analyses carried out on the residuals. The effect on the variogram before and after removing trend will be illustrated with an analysis of a DEM and an EMI scan.

Matheron, G. 1982. Pour une analyse krigéante de données régionalisées. Note N-732, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau..