Geophysical Research Abstracts
Vol. 14, EGU2012-2896, 2012
EGU General Assembly 2012

Quality measures for geostatistical prediction of log-normal soil properties.

R. M. Lark
British Geological Survey, Nottingham, United Kingdom (mlark@nerc.ac.uk)

A signature of non-linear processes in the soil is the non-normal distribution of soil properties. A common non-normal distribution is the log-normal, in which the variable Z can be transformed to a variable with a normal distribution by $Y = \log_e(Z)$. Log-normal variables are common in soil geochemistry and hydrology. It is standard practice in geostatistics to use the log-transformation for such variables before spatial modelling and prediction, and there are procedures to back-transform predictions of Y to the original scale of measurement Z. This is important because values on the original scale are commonly required either for scientific purposes or for practical applications such as the assessment of potential contaminant concentrations in soil.

One of the strengths of geostatistics is that geostatistical prediction returns a prediction error variance. Furthermore, this variance can be computed before a survey is undertaken, for a range of possible different sampling networks, since it depends only on the disposition of sample sites, and the variogram model of spatial dependence. This allows the most efficient network to be selected: one which will provide estimates of sufficient precision (where the prediction error variances are within acceptable bounds) without over-sampling. In log-normal kriging the prediction error variance depends not only on the variogram and the sampling array, but also on the conditional mean value of the variable, which is not known until after sampling. This means that the usual pre-survey quality measures which can be computed to guide the planning of geostatistical surveys are not available for log-normal variables. Given that many critical variables, such as contaminant concentrations, are often log-normally distributed, this is a serious gap in the capablity of geostatistics to facilitate rational sampling design for environmental management and monitoring.

In this paper I propose and demonstrate some quality measures that can be computed, pre-survey, for different sampling networks for log-normal variables. I propose measures that are suitable both for point kriging, when predictions are obtained for the same sample support as the observations, and for block kriging when the predictions are upscaled to a region, such as a management unit or a regular block. The quality measures are dimensionless prediction intervals, scaled relative to the unknown median value of the target conditional distribution. In the point kriging case the distribution of interest is the conditional distribution of the variable at the target location. In the block kriging case the distribution is that of the spatial median value over the block.

These quality measures are illustrated in a case study on the spatial variability of some critical metals and metalloids in the soil of eastern England. Target values of the quality measures are defined in terms of the requirements that land managers or regulators may specify for uncertainty in final predictions, and the sampling effort required to meet these is computed.