Geophysical Research Abstracts
Vol. 14, EGU2012-3025, 2012
EGU General Assembly 2012

Robust geostatistical analysis of spatial data

A. Papritz (1), H.R. Künsch (2), C. Schwierz (1,2), and W.A. Stahel (2)

(1) ETH Zurich, Institute of Terrestrial Ecosystems, Zurich, Switzerland, (2) Seminar for Statistics, ETH Zurich, Zurich, Switzerland

Most of the geostatistical software tools rely on non-robust algorithms. This is unfortunate, because outlying observations are rather the rule than the exception, in particular in environmental data sets. Outlying observations may results from errors (e.g. in data transcription) or from local perturbations in the processes that are responsible for a given pattern of spatial variation. As an example, the spatial distribution of some trace metal in the soils of a region may be distorted by emissions of local anthropogenic sources. Outliers affect the modelling of the large-scale spatial variation, the so-called external drift or trend, the estimation of the spatial dependence of the residual variation and the predictions by kriging. Identifying outliers manually is cumbersome and requires expertise because one needs parameter estimates to decide which observation is a potential outlier. Moreover, inference after the rejection of some observations is problematic. A better approach is to use robust algorithms that prevent automatically that outlying observations have undue influence.

Former studies on robust geostatistics focused on robust estimation of the sample variogram and ordinary kriging without external drift. Furthermore, Richardson and Welsh (1995) [2] proposed a robustified version of (restricted) maximum likelihood ([RE]ML) estimation for the variance components of a linear mixed model, which was later used by Marchant and Lark (2007) [1] for robust REML estimation of the variogram. We propose here a novel method for robust REML estimation of the variogram of a Gaussian random field that is possibly contaminated by independent errors from a long-tailed distribution. It is based on robustification of estimating equations for the Gaussian REML estimation. Besides robust estimates of the parameters of the external drift and of the variogram, the method also provides standard errors for the estimated parameters, robustified kriging predictions at both sampled and unsampled locations and kriging variances. The method has been implemented in an R package.

Apart from presenting our modelling framework, we shall present selected simulation results by which we explored the properties of the new method. This will be complemented by an analysis of the Tarrawarra soil moisture data set [3].

References

[1] Marchant, B. P. and Lark, R. M. (2007). Robust estimation of the variogram by residual maximum likelihood. *Geoderma*, **140**, 62–72.

[2] Richardson, A. M. and Welsh, A. H. (1995). Robust restricted maximum likelihood in mixed linear models. *Biometrics*, **51**, 1429–1439.

[3] Western, A. W. and Grayson, R. B. (1998). The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements. *Water Resources Research*, **34**(10), 2765–2768.