Geophysical Research Abstracts
Vol. 17, EGU2015-7780, 2015
EGU General Assembly 2015

spsann - optimization of sample patterns using spatial simulated annealing

Alessandro Samuel-Rosa (1), Gerard Heuvelink (), Gustavo Vasques (), and Lúcia Anjos ()

(1) Federal Rural University of Rio de Janeiro, Institude of Agronomy, Graduate Course in Agronomy-Soil Science, Seropédica, Brazil (alessandrosamuelrosa@gmail.com), (2) Soil Geography and Landscape Group, Wageningen University, and ISRIC-World Soil Information, Wageningen, the Netherlands, (3) National Center for Soil Research, Brazilian Agricultural Research Corporation, Rio de Janeiro, Brazil, (4) Department of Soil Science, Federal Rural University of Rio de Janeiro, Seropédica, Brazil

There are many algorithms and computer programs to optimize sample patterns, some private and others publicly available. A few have only been presented in scientific articles and text books. This dispersion and somewhat poor availability is holds back to their wider adoption and further development.

We introduce **spsann**, a new R-package for the optimization of sample patterns using spatial simulated annealing. R is the most popular environment for data processing and analysis. Spatial simulated annealing is a well known method with widespread use to solve optimization problems in the soil and geo-sciences. This is mainly due to its robustness against local optima and easiness of implementation.

spsann offers many optimizing criteria for sampling for variogram estimation (number of points or point-pairs per lag distance class - PPL), trend estimation (association/correlation and marginal distribution of the covariates - ACDC), and spatial interpolation (mean squared shortest distance - MSSD). **spsann** also includes the mean or maximum universal kriging variance (MUKV) as an optimizing criterion, which is used when the model of spatial variation is known. PPL, ACDC and MSSD were combined (PAN) for sampling when we are ignorant about the model of spatial variation. **spsann** solves this multi-objective optimization problem scaling the objective function values using their maximum absolute value or the mean value computed over 1000 random samples. Scaled values are aggregated using the weighted sum method.

A graphical display allows to follow how the sample pattern is being perturbed during the optimization, as well as the evolution of its energy state. It is possible to start perturbing many points and exponentially reduce the number of perturbed points. The maximum perturbation distance reduces linearly with the number of iterations. The acceptance probability also reduces exponentially with the number of iterations.

R is memory hungry and spatial simulated annealing is a computationally intensive method. As such, many strategies were used to reduce the computation time and memory usage: a) bottlenecks were implemented in C++, b) a finite set of candidate locations is used for perturbing the sample points, and c) data matrices are computed only once and then updated at each iteration instead of being recomputed.

spsann is available at GitHub under a licence GLP Version 2.0 and will be further developed to: a) allow the use of a cost surface, b) implement other sensitive parts of the source code in C++, c) implement other optimizing criteria, d) allow to add or delete points to/from an existing point pattern.