Geophysical Research Abstracts
Vol. 17, EGU2015-6408, 2015
EGU General Assembly 2015

A stratified two-stage sampling design for digital soil mapping in a Mediterranean basin

Michael Blaschek and Rainer Duttmann
Kiel University, Department of Geography, Ludewig-Meyn-Str. 14, D-24118 Kiel

The quality of environmental modelling results often depends on reliable soil information. In order to obtain soil data in an efficient manner, several sampling strategies are at hand depending on the level of prior knowledge and the overall objective of the planned survey. This study focuses on the collection of soil samples considering available continuous secondary information in an undulating, 16 km^2-sized river catchment near Ussana in southern Sardinia (Italy). A design-based, stratified, two-stage sampling design has been applied aiming at the spatial prediction of soil property values at individual locations.
The stratification based on quantiles from density functions of two land-surface parameters – topographic wetness index and potential incoming solar radiation – derived from a digital elevation model. Combined with four main geological units, the applied procedure led to 30 different classes in the given test site. Up to six polygons of each available class were selected randomly excluding those areas smaller than 1ha to avoid incorrect location of the points in the field. Further exclusion rules were applied before polygon selection masking out roads and buildings using a 20m buffer. The selection procedure was repeated ten times and the set of polygons with the best geographical spread were chosen. Finally, exact point locations were selected randomly from inside the chosen polygon features. A second selection based on the same stratification and following the same methodology (selecting one polygon instead of six) was made in order to create an appropriate validation set. Supplementary samples were obtained during a second survey focusing on polygons that have either not been considered during the first phase at all or were not adequately represented with respect to feature size. In total, both field campaigns produced an interpolation set of 156 samples and a validation set of 41 points. The selection of sample point locations has been done using ESRI software (ArcGIS) extended by Hawth's Tools and later on its replacement the Geospatial Modelling Environment (GME). 88% of all desired points could actually be reached in the field and have been successfully sampled.
Our results indicate that the sampled calibration and validation sets are representative for each other and could be successfully used as interpolation data for spatial prediction purposes. With respect to soil textural fractions, for instance, equal multivariate means and variance homogeneity were found for the two datasets as evidenced by significant ($P > 0.05$) Hotelling T^2-test (2.3 with df1 = 3, df2 = 193) and Bartlett's test statistics (6.4 with df = 6). The multivariate prediction of clay, silt and sand content using a neural network residual cokriging approach reached an explained variance level of 56%, 47% and 63%.
Thus, the presented case study is a successful example of considering readily available continuous information on soil forming factors such as geology and relief as stratifying variables for designing sampling schemes in digital soil mapping projects.