



Extracting heterogeneity of subsoil from geophysical measurements using unsupervised learning algorithms

Christoph Herbert (1,2), Florian Wellmann (2), Hui Wang (3), and Christian von Hebel (4)

(1) Politechnic University of Catalonia, Department of Signal Theory and Communications, Barcelona, Spain (christoph.josef.herbert@upc.edu), (2) RWTH Aachen University, Department of Computational Geoscience and Reservoir Engineering, Aachen, Germany (christoph.herbert@rwth-aachen.de), (3) University of Dayton, Department of Civil and Environmental Engineering and Engineering Mechanics, Ohio, USA, (4) Research Centre Jülich, Agrosphere, Institute of Bio- and Geoscience (IBG-3), Jülich, Germany

Heterogeneities in subsurface soil structures have a potentially important influence on crop performance in agricultural fields. However, determining these subsurface structures can be difficult and time-intensive. The relation between surface patterns link crop performance and spatial variation of soil-physical properties is investigated on the basis of a Bayesian unsupervised learning algorithm. Remote sensing techniques are assumed to be indirectly sensitive to soil water content. Finding a correlation between those techniques and depth-sensitive on-site geophysical measurements may enable us to extract subsurface heterogeneity. A deduced general correlation could then be used to harness increasingly available satellite data sets for prediction of subsurface heterogeneity.

We investigate here the relation between surface patterns link crop performance and spatial variation of soil-physical properties using a Bayesian unsupervised learning algorithm. The algorithm is applied to field data combining multi-configuration electromagnetic induction (EMI) measurements and spatio-temporal normalized difference vegetation index (NDVI) images. It is validated with geological subsurface structure estimation based on soil development and deposition processes. The quantification and visualization of intra-cluster cohesion and inter-cluster separability shows clear correlations for certain sub-regions. Furthermore, subsurface heterogeneity can be extracted and conditionally verified with prior estimates of the number of soil horizons. A specifically developed anomaly analysis extension of the algorithm enables outlier detection based on spatial and statistical criteria and has the capability to reveal hidden sub-structures for cases where noisy input features bias segmentation. Our developments provide a path to a more robust estimation of subsurface soil heterogeneity from remote sensing information, although significant challenges remain due to variability between different plant types and seasonal effects.