Geophysical Research Abstracts
Vol. 20, EGU2018-17808-1, 2018
EGU General Assembly 2018

Learning patterns from data in an unsupervised manner: A Bayesian approach for spatial and statistical pattern extraction of subsoil heterogeneity using satellite derived NDVI time series and electromagnetic induction measurements

Hui Wang (1), Florian Wellmann (1), Maximilian Kanig (2), Christian von Hebel (2), and Jan van der Kruk (2)
(1) The Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Aachen, Germany., (2) Institute of Bio- and Geoscience - Agrosphere (IBG-3), Forschungszentrum Jülich GmbH, Jülich, Germany

Near-surface hydrogeologic properties have natural variation in space as a result of the complex transportation and deposition processes that form the subsurface. In recent years, spatial subsurface heterogeneity has drawn much attention as it does, combined with climatic conditions, affect the crop yield and hence becomes one of the key factors in precision agriculture.

We address the question of linking remotely sensed surface vegetation indices with heterogeneous subsoil physical properties by means of a Bayesian unsupervised machine learning approach used for spatial and statistical pattern extraction. This question has considerable relevance and practical meaning as visible spatial differences in crop development and yield can be related to horizontal and vertical variations in soil texture caused by complex deposition/erosion processes. Besides, active and relict geomorphological settings, such as floodplain, buried paleo-channels, and glacial terrace can cast significant complexity into surface hydrology and crop modeling, this also requires a better approach to detect, quantify, and simulate subsoil heterogeneity. In this work, we introduce a novel spatial and statistical pattern recognition and uncertainty quantification framework based on hidden Markov random field (HMRF) and finite mixture model (FMM), which is validated using synthetic data.

Satellite-derived NDVI time series are combined with proximal soil sensor data acquired by a fixed-beam multi-receiver electromagnetic induction (EMI) survey system, and the multi-dimensional information are clustered. In this way, the similarity in geometry structures between amplitude images from harmonic analysis of NDVI time series and soil apparent electrical conductivity (ECa) for nine depths of investigation (DOI) can be extracted and the associated uncertainty can be quantified in a statistical manner. In addition, we also detect, estimate, and visualize the statistical similarity intra-cluster and the inter-cluster separability. The correlation coefficients between different NDVI amplitude terms and soil apparent electrical conductivity at multiple DOI's are calculated and the spatial and statistical patterns are interpreted by incorporating additional knowledge from geological maps and soil samples. The preliminary results are promising and future investigations focus on the statistical patterns to obtain an improved understanding of the surface vegetation response on the soil heterogeneity.