Geophysical Research Abstracts
Vol. 15, EGU2013-10981, 2013
EGU General Assembly 2013

Predicting local Soil- and Land-units with Random Forest in the Senegalese Sahel

Tobias Grau, Martin Brandt, and Cyrus Samimi
University of Vienna, Institut of Geography, Vienna, Austria (martin.brandt@univie.ac.at)

MODIS (MCD12Q1) or Globcover are often the only available global land-cover products, however ground-truthing in the Sahel of Senegal has shown that most classes do have any agreement with actual land-cover making those products unusable in any local application.

We suggest a methodology, which models local Wolof land- and soil-types in an area in the Senegalese Ferlo around Linguère at different scales. In a first step, interviews with the local population were conducted to ascertain the local denotation of soil units, as well as their agricultural use and woody vegetation mainly growing on them. "Ndjor" are soft sand soils with mainly Combretum glutinosum trees. They are suitable for groundnuts and beans while millet is grown on hard sand soils ("Bardjen") dominated by Balanites aegyptiaca and Acacia tortilis. "Xur" are clayey depressions with a high diversity of tree species. Lateritic pasture sites with dense woody vegetation (mostly Pterocarpus lucens and Guiera senegalensis) have never been used for cropping and are called "All". In a second step, vegetation and soil parameters of 85 plots ($\sim$1 ha) were surveyed in the field. 28 different soil parameters are clustered into 4 classes using the WARD algorithm. Here, 81% agree with the local classification. Then, an ordination (NMDS) with 2 dimensions and a stress-value of 9.13% was calculated using the 28 soil parameters. It shows several significant relationships between the soil classes and the fitted environmental parameters which are derived from field data, a digital elevation model, Landsat and RapidEye imagery as well as TRMM rainfall data. Landsat's band 5 reflectance values (1.55 – 1.75 μm) of mean dry season image (2000-2010) has a R^2 of 0.42 and is the most important of 9 significant variables (5%-level). A random forest classifier is then used to extrapolate the 4 classes to the whole study area based on the 9 significant environmental parameters. At a resolution of 30 m the OBB (out-of-bag) error rate is 6.55%. At this scale, even small depressions and local discrepancies at field-level can be identified. Scaling to 250, 500 and 1000 m still gives a reliable overview of prevalent soil-units in the region. By relating soil to vegetation parameters we prove, that these maps indicate the potential dominant woody vegetation.

Our example demonstrates, that solely the use of native Wolof land-types, which are gathered by interviews, can be used to get a proper scientific classification of (a) the agricultural suitability and (b) the dominant woody vegetation in an area of the Senegalese Sahel region.