Geophysical Research Abstracts
Vol. 17, EGU2015-14979, 2015
EGU General Assembly 2015

Using spectral subspaces to improve infrared spectroscopy prediction of soil properties

Andrew Sila (1,2), Keith D. Shepherd (1), Ganesh P. Pokhariyal (2), Erick Towett (1), Elvis Weullow (1), and Mercy K. Nyambura ()

(1) World Agroforestry Centre (ICRAF), P.O. Box 30677-00100, Nairobi, Kenya (a.sila@cgiar.org), (2) School of Mathematics, University of Nairobi, P.O Box 30196-00100 GPO, Nairobi, Kenya

We propose a method for improving soil property predictions using local calibration models trained on datasets in spectral subspaces rather that in a global space. Previous studies have shown that local calibrations based on a subset of spectra based on spectral similarity can improve model prediction performance where there is large population variance. Searching for relevant subspaces within a spectral collection to construct local models could result in models with high power and small prediction errors, but optimal methods for selection of local samples are not clear. Using a self-organizing map method (SOM) we obtained four mid-infrared subspaces for 1,907 soil sample spectra collected from 19 different countries by the Africa Soil Information Service. Subspace means for four sub-spaces and five selected soil properties were: pH, 6.0, 6.1, 6.0, 5.6; Mehlich-3 Al, 358, 974, 614, 1032 (mg/kg); Mehlich-3 Ca, 363, 1161, 526, 4276 (mg/kg); Total Carbon, 0.4, 1.1, 0.6, 2.3 (% by weight), and Clay (%), 16.8, 46.4, 27.7, 63.3. Spectral subspaces were also obtained using a cosine similarity method to calculate the angle between the entire sample spectra space and spectra of 10 pure soil minerals. We found the sample soil spectra to be similar to four pure minerals distributed as: Halloysite (n1=214), Illite (n2=743), Montmorillonite (n3=914) and Quartz (n4=32). Cross-validated partial least square regression models were developed using two-thirds of samples spectra from each subspace for the five soil properties.We evaluated prediction performance of the models using the root mean square error of prediction (RMSEP) for a one-third-holdout set.
Local models significantly improved prediction performance compared with the global model. The SOM method reduced RMESP for total carbon by 10 % (global RMSEP = 0.41) Mehlich-3 Ca by 17% (global RMESP = 1880), Mehlich-3 Al by 21 % (global RMSEP = 206), and clay content by 6 % (global RMSEP = 13.6), but not for pH. Individual SOM groups gave up to 94% reduction in RMSEP for total carbon, compared with the global model. Overall SOM gave better predictions than the cosine method but the latter is worth researching further as it may provide additional interpretative insight into soil mineralogical differences.