Geophysical Research Abstracts
Vol. 18, EGU2016-7276, 2016
EGU General Assembly 2016

Local Partial-Least-Squares Regression Approaches For The Prediction Of Soil Properties From Diffuse Reflectance Spectrosocopy Data

Abhinav Gupta (1), Bhabani Sankar Das (1), Hitesh B. Vasava (1), and Sarathjith M. Chandran (2)
(1) Indian Institute of Technology Kharagpur, (2) International Crop Reserach Institute for Semi Arid Tropics, Mali

The chemometric model building is the most important part of diffuse reflectance spectroscopy (DRS) technology. Several regression methods have been applied for the prediction of soil properties from spectra. Partial-least-squares regression (PLSR) is the most popularly used algorithm in DRS. In PLSR, a single model is calibrated for a spectral library, which is used for the prediction of corresponding soil property for new soil samples. This approach of modeling is called global approach. Recently, local models have gained popularity in DRS due to their ability to model non-linearity. In local modeling, an independent model is calibrated for each target sample based on its similarity with the other samples in the spectral library. K-nearest neighbor PLSR (KNN-PLSR) and locally weighted partial-least-squares regression (LW-PLSR) are the most popularly employed local modeling approaches. However, local models have been applied for large spectral libraries only and LW-PLSR has not been tested comprehensively in DRS. Therefore, we tested the suitability of local models for small spectral libraries. We compared PLSR, KNN-PLSR and different LW-PLSR for the prediction of organic carbon (OC) and sand content in a small spectral library. Different LW-PLSR approaches based on different similarity measures such as Euclidean distance (edLW-PLSR), covariance based (cbLW-PLSR), correlation based (corrLW-PLSR), surface spectrum difference based (sdsLW-PLSR), optimized principal component based (oPC-M) distance were used in this study. Our results showed that LW-PLSR models perform better than PLSR and KNN-PLSR. For small spectral libraries, LW-PLSR improves the prediction accuracy of soil properties and KNN-PLSR didn't have any significant effect on model performance. Among different LW-PLSR approaches, corrLW-PLSR was the best approach, which improved the prediction accuracy of OC from R2 = 0.593 to 0.699 and of sand content from 0.662 to 0.724 followed by sdsLW-PLSR. One drawback of locally weighted modeling is heavy computational load but for small spectral libraries application of LW-PLSR is feasible.