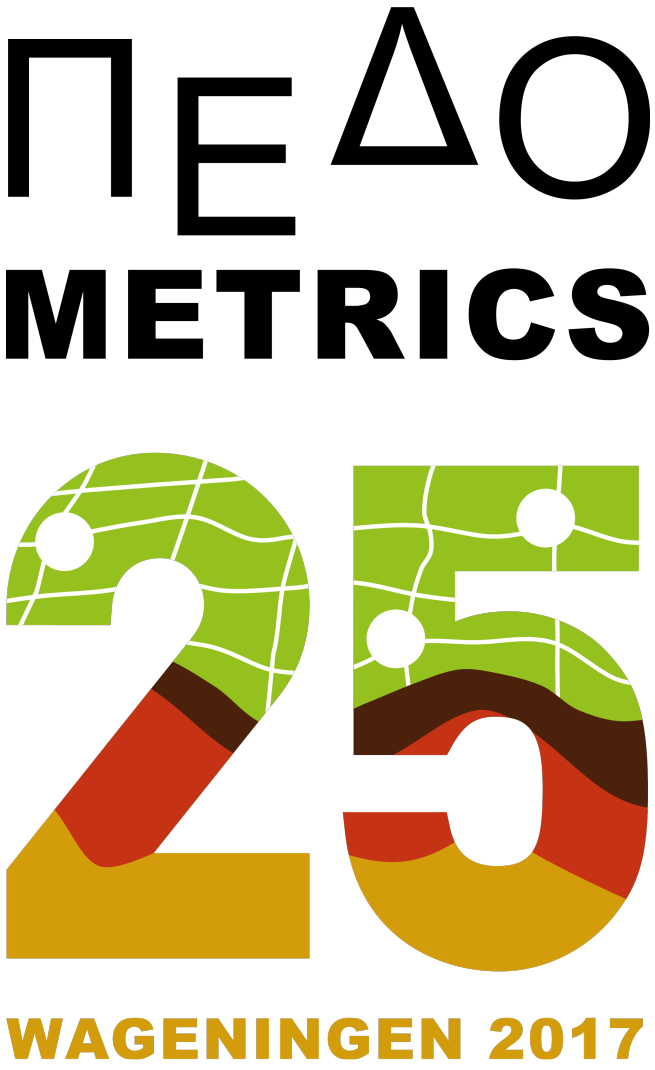
Abstract Book Pedometrics 2017



Wageningen, 26 June – 1 July 2017

Comparison between random forest and partial least square regression of on-line vis-NIR spectroscopy measurements of soil total nitrogen and organic carbon

Said Nawar – Cranfield university, United Kingdom
Abdul M. Mouazen – Ghent University, Belgium

Accurate and detailed spatial soil information about within field variability is essential for variable-rate applications of farm resources. Soil total nitrogen (TN) and organic carbon (OC) are important fertility parameters that can be measured with on-line (mobile) visible and near infrared (vis-NIR) spectroscopy, whose the calibration method may considerably affect the measurement accuracy. This study aims at comparing the performance of local farm scale calibrations with those based on spiking of local samples into an European continental dataset (ECD) for TN and OC estimation using two modelling techniques, namely, random forest (RF) and partial least squares regression (PLSR). An on-line sensor platform equipped with a mobile, fiber type, vis-NIR spectrophotometer (AgroSpec from tec5 Technology for Spectroscopy, Germany), with a measurement range of 305–2200 nm was used to acquire soil spectra in diffuse reflectance mode from two fields in the UK. After dividing spectra into calibration (75%) and validation (25%) sets, spectra in the calibration set were subjected RF and PLSR with leave-one-out cross-validation to establish calibration models of TN and OC. On-line predicted values of TN and OC were used to develop maps using ordinary kriging in ArcGIS software (ESRI, USA). Results showed that RF outperformed PLSR models for both datasets used, whereas the lowest model performance was obtained with the local dataset. The effect of spiking local samples into the ECD was significant, and resulted in high coefficients of determination (R2) values of 0.97 and 0.98, low root mean square error (RMSE) of 0.01 and 0.11, and high residual prediction deviations (RPD) of 5.78 and 7.17, for TN and TC, respectively. The on-line predicted maps showed better spatial similarities between laboratory measured and RF predicted maps, as compared to PLSR predicted maps. Therefore, these results suggest that ECD vis-NIR RF calibration models can be successfully used to predict TN and OC under on-line measurement conditions.

keywords: *Vis-NIR spectroscopy, spiking, random forest, partial least squares regression, soil mapping*