Evaluating regression-kriging for mid-infrared spectroscopy prediction of soil properties in western Kenya-East Africa

Andrew Sila, Ganesh Pokhariyal, Keith Shepherd

PII: S2352-0094(16)30078-5
DOI: doi: 10.1016/j.geodrs.2017.04.003
Reference: GEODRS 120

To appear in: *Geoderma Regional*

Received date: 18 August 2016
Revised date: 10 February 2017
Accepted date: 3 April 2017

Please cite this article as: Andrew Sila, Ganesh Pokhariyal, Keith Shepherd , Evaluating regression-kriging for mid-infrared spectroscopy prediction of soil properties in western Kenya-East Africa. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Geodrs(2017), doi: 10.1016/j.geodrs.2017.04.003

Evaluating regression-kriging for mid-infrared spectroscopy prediction of soil properties in western Kenya- East Africa

Andrew Sila[a, b], Ganesh Pokhariyal[b] and Keith Shepherd[a]

[a] World *Agroforestry Centre, P.O. Box 30677 00100, Nairobi, Kenya*

[b] *School of Mathematics, University of Nairobi, P.O. Box 30197 00100, Nairobi, Kenya. (Corresponding author, Andrew Sila- Email: asila@cgiar.org; Phone: +254 736 363525)*

Highlights

1. Predictions of aluminium, copper, iron, zinc and boron were improved using the hybrid method.
2. Spherical and exponential methods gave lowest MSPE in most of the variogram models.
3. Soil carbon content was well predicted using the PLS regression method only.
4. Approach accounts for residual spatial correlation.

Abstract

Infrared diffuse reflectance spectroscopy has been widely applied in soil and environmental sciences to develop soil property calibration models. Quantitative relationships between soil properties are usually developed using different regression methods. These regression methods often rely on the assumption that the model residuals are independently and identically distributed without checking for spatial correlation. In this study, the utility of regression-kriging was investigated in building prediction models for soil properties using mid-infrared (7498 to 600 cm^{-1}) spectral data for soil samples collected from Nyando, Nzoia and Yala catchment areas in Kenya, sampled at 0-20 cm and 20-50 cm depths. Using a systematic technique, 158 samples were selected for analysis of a number of soil properties of interest using wet chemistry methods. We randomly divided the dataset into two groups: 118 samples in the calibration and 40 samples in the holdout validation set. The calibration set was first used to develop partial least squares regression (PLS) regression models for all the soil properties. Residuals from these models were used to generate semivariograms, which revealed a strong spatial dependence as determined by the ratio of nugget to sill for nitrogen, 9%; Al, 12%; and B, 36%, but with weak spatial dependence for exchangeable Ca (ExCa), 100%; and carbon, 76%. The fitted theoretical semivariograms were used to fit regression-kriging models. Lastly, both the PLS and regression-kriging models were assessed with the validation set and their prediction performance evaluated by R^2 and root mean square error (RMSE). The results showed that regression kriging method gave lower RMSE values for all the evaluated soil properties except for ExCa, B and exchangeable acidity, with the best predictions, compared with the PLS model, obtained for ExMg (R^2, 0.93 vs 0.88; RMSE, 6.1 vs 8.4 $cmol_c kg^{-1}$ and total nitrogen (R^2=0.92 vs R^2=0.74; RMSE, 0.11%, RMSE=0.2%). In this study, regression-kriging, which takes into account spatial variation normally ignored by other methods, improved use of infrared spectroscopy for predicting soil properties.

Keywords: Vertisols; spectroscopy; PLS; regression-kriging; chemometrics; cross-validation