Geophysical Research Abstracts
Vol. 20, EGU2018-11169, 2018
EGU General Assembly 2018

Soil sample optimization for precision agriculture

Yones Khaledian (1) and Bradley Miller (1,2)

(1) Department of Agronomy, Iowa State University, Ames, Iowa, USA (ykh@iastate.edu), (2) Institute of Soil Landscape Research, Leibniz-Centre for Agricultural Landscape Research (ZALF) e.V., Müncheberg, Germany (miller@zalf.de)

Grid, random, and stratified sampling each come with unique advantages and disadvantages. A common sampling strategy in Iowa farm fields is grid sampling at intervals equivalent to approximately 50, 75, or 100 m. However, more efficient sampling could be accomplished with a more strategic sampling design. Using data from multiple fields that were sampled using the standard grid approach, we selected different subsets of those points by gradually reducing the number of model training points to identify the optimal sample size. We define optimal for this study as the minimum number of samples needed to achieve a similar spatial modeling accuracy as the original grid sampling approach. In this study, we consistently used the same covariates and model building method to predict soil organic carbon. Specifically, we used Cubist to develop rule-based, multiple linear regression models. The accuracy of the models was assessed using the R2 and RMSE of the cross-validation. For reducing the quantity of samples used to train the model, we compared the selection of points from the existing grid samples by random, match-stratified, and stratified sample selection. The 'random' sample sets were generated by randomly selecting the needed number of samples from the grid sample points. The matched-stratified sampling set was generated by matching each location selected by conditioned Latin hypercube sampling (cLHS) with the nearest grid sampling point. For the stratified sample sets, the points from the grid were categorized by slope gradient quantiles and then randomly selected within each quantile to obtain the needed number of points. Our results indicated that for a regression-based model, there was high variability and no clear pattern in model accuracy as the quantity of randomly selected samples was reduced. In contrast, stratified strategies can maintain similar prediction accuracy with much fewer samples. The benefit of maintaining prediction accuracy with fewer samples is reduced labor and sample analysis costs.