Prediction of Total Nitrogen Distribution in Surface Soil Based on Multi-source Auxiliary Variables and Random Forest Approach

Yang Zhou\textsuperscript{1,2}, Xiaomin Zhao\textsuperscript{1,2}, and Xi Guo\textsuperscript{1,2}

\textsuperscript{1}Academy of Land Resource and Environment, Jiangxi Agricultural University, Nanchang 330045, China
\textsuperscript{2}Key Laboratory of Poyang Lake Watershed Agricultural Resources and Ecology of Jiangxi Province, Nanchang 330045, China

Soil total nitrogen is closely related to soil quality and fertility. It is of great significance to know the spatial distribution characteristics of soil total nitrogen for the implementation of precision agriculture management. The spatial distribution of total nitrogen in the surface soil of Xunwu County was predicted and mapped by using two methods: random forest and random forest plus residuals kriging. These methods were combined with multi-source auxiliary variables such as (i) terrain factors, (ii) geographical coordinate, (iii) remote sensing factors, (iv) climate factors, (v) distance factors, and (vi) soil physical or chemical factors. Also, the prediction accuracy of the two models was compared after 100 times of repeated operation. Our results show that the mean values of the decision coefficient ($R^2 = 0.6291$) and concordance correlation coefficient (CCC = 0.7613) of the random forest model were higher than those of the random forest plus residual kriging method ($R^2 = 0.5719$, CCC = 0.6881). Also, the mean values of the mean absolute error (MAE = 0.1570 g·kg\textsuperscript{-1}) and root mean squared error (RMSE = 0.2108 g·kg\textsuperscript{-1}) were lower than those of the random forest plus residual kriging method (MAE = 0.1682 g·kg\textsuperscript{-1}, RMSE = 0.2267 g·kg\textsuperscript{-1}). Importantly, adding residual to the random forest model did not improve its accuracy. These results suggest that the random forest model can be used as a new method for predicting soil properties, and it provides technical support for the implementation of agricultural management.