EGU22-10847, updated on 31 May 2022
https://doi.org/10.5194/egusphere-egu22-10847
EGU General Assembly 2022

Random Forests to classify salt-affected soils from soluble salt ions

Demis Andrade Foronda
Universidad Mayor de San Simón, Facultad de Ciencias Agrícolas y Pecuarias, Cochabamba, Bolivia
(demis.andrade@gmail.com)

Salt-affected soils are characterized by an excess of soluble salts and/or sodium. The widely used US Salinity Lab (USSL) classification considers the exchangeable sodium percentage (ESP), electrical conductivity (EC) and pH. Breiman and Cutler's random forests (RF) algorithm chooses the most voted class over all the trees at training time. The aim was to model and predict the USSL salt-term categories from soluble ions by applying RF model classification. Topsoil samples (110) were collected from the High Valley of Cochabamba (Bolivia) and were analyzed to measure the soluble cations (Na^+, K^+, Ca^{2+}, Mg^{2+}) and anions (HCO_3^-, Cl^-, CO_3^{2-}, SO_4^{2-}), in addition to the required salinity parameters to classify the samples according to the thresholds: ESP of 5%, EC_e of 4 dSm^{-1} and pH of 8.2. No samples matched in the saline category. The overall out-of-bag error was 17.4% and according to the confusion matrix, the class errors for normal, saline-sodic and sodic soil were 0.12, 0.00 and 0.36, respectively. The variables with higher estimated importance and also selected by RF backward elimination were: Na^+, Cl^-, Ca^{2+} and HCO_3^-. Additional sampling might be useful in order to reduce the errors and misclassifications, as well as to improve the selection of variables.