EGU2020-22569
https://doi.org/10.5194/egusphere-egu2020-22569
EGU General Assembly 2020

Development of a Swiss National Soil Spectral Model Library using data-driven modeling

Philipp Baumann, Anatol Helfenstein[2], Andreas Gubler[3], Reto Meuli[3], Armin Keller[4], Juhwan Lee[5], Raphael A. Viscarra Rossel[5], and Johan Six[1]

[1]Institute of Agricultural Sciences, Department of Environmental Systems Science (D-USYS), ETH Zürich, Switzerland
[2]Soil Geography and Landscape Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands
[3]Swiss Soil Monitoring Network (NABO), Agroscope, Reckenholzstrasse 191, 8046 Zürich, Switzerland
[4]Swiss Soil Competence Center (KOBO), School of Agricultural, Forest and Food Sciences HAFL, Bern University of Applied Sciences BFH, Bern, Switzerland
[5]Soil and Landscape Science, School of Molecular and Life Sciences, Curtin University, GPO Box U1987, Perth WA 6845, Australia

Soil data at different scales are needed for assessments and monitoring of soil functions. Soil diffuse reflectance spectroscopy using visible–Near Infrared and mid-Infrared energies can be used to estimate a range of soil properties, rapidly and inexpensively. However the spectroscopic modeling is challenging because of the large soil diversity and its complex composition. We developed a National Soil Spectral library (SSL) (n = 4339) using samples from (i) the Swiss Soil Monitoring Network (NABO; 7 sampling campaigns at 71 agricultural locations since 1985, n = 592) and (ii) the National Biodiversity Monitoring (BDM) Program (n = 4295, 1094 locations across a 5x5 km grid). The SSL will provide spectroscopic models for estimation of functional soil properties at different scales (e.g. total carbon (C) and nitrogen, organic C, texture, pH and cation exchange capacity). We used a rule-based algorithm, Cubist, for the modelling. The models were tuned across full combinations of {5, 10, 20, 50, 100} committees and {2, 5, 7, 9} neighbors, using 5 times repeated 10-fold cross-validation grouped by location. Further, transfer learning with RS-LOCAL tuning was performed for each of the 71 monitoring sites separately by a hold out approach in order to select optimal instances from the remaining SSL. Total soil C in the reference data ranged from 0.1% to 58.3% C and the best Cubist model had a cross-validated RMSE of 0.82% C. The RS-LOCAL approach ($RMSE_{mean}$ = 0.14 %) was on average 2.5 times more accurate for the estimation of C over time at each of the 71 NABO sites compared to the general Cubist approach. Our results suggest that data-driven selection of SSL instances targeted to closely related soils produces less biased estimation of soil properties over time at smaller geographic extents. The general Cubist calibration models are useful when reference analyses in a new study area are scarce. In conclusion, the Swiss SSL models can be used to cost-efficiently estimate a range of soil properties for a diverse applications and purposes in Switzerland.