



Soil organic carbon mapping using state-of-the-art machine learning algorithms and deep neural networks in different climatic regions of Iran

Ruhollah Taghizadeh-Mehrjardi (1,4), Karsten Schmidt (1,5), Mojtaba Zeraatpisheh (2), Thorsten Behrens (1), Roozbeh Valavi (3), Thomas Scholten (1,5)

(1) Eberhard Karls University Tübingen, Institute of Geography, Soil Science and Geomorphology, Tübingen, Germany (ruhollah.taghizadeh-mehrjardi@mnf.uni-tuebingen.de), (2) Department of Soil Science, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz, Iran (zeraatpishem@yahoo.com), (3) The Quantitative & Applied Ecology Group, School of BioSciences, The University of Melbourne, Australia (valavi.r@gmail.com), (4) Faculty of Agriculture and Natural Resources, Ardakan University, Ardakan, Iran (rtaghizadeh@ardakan.ac.ir), (5) CRC 1070 RessourceCulture, University of Tübingen, Gartenstr. 29, Tübingen, Germany

Abstract

Soil Organic Carbon (SOC) is a major source of terrestrial carbon and also a key element for soil quality and fertility, as it improves aggregate stability, water retention capacity, and nutrient availability. This study aims to evaluate the performance of eight state-of-the-art machine learning algorithms including: artificial neural network (ANN), random forest (RF), Cubist, support vector regression (SVR), regularized linear regression (lasso), k nearest neighbour (kNN), XGBoost, and deep neural networks (DL) to predict SOC concentration at five standard depth intervals (0-5, 5-15, 15-30, 30-60 and 60-100 cm) using environmental covariates as predictors. Three study areas in Iran from different climatic regions including arid, semi-arid, and humid climate regimes were selected with an overall sample count of 185, 188 and 122, respectively. The SOC content ranges between 0.03 and 12.33%. In all areas, the amount of SOC decreased with increasing soil depths. The same holds true for the performance of all tested prediction techniques, calculated as spatial cross-validation (0.19% to 1.90%). Furthermore, results showed that all machine-learning techniques had better performances in arid region compared to semi-arid and humid climate regimes. Moreover, DL outperformed the commonly used machine learning algorithms in digital soil mapping (0.15% to 1.10%). The superiority of DL-models was most evident for deeper soil increments. We conclude that DL has a high potential in soil organic carbon mapping and could be an alternative to the standard algorithms to predict SOC content especially with regard to larger soil depth.

Keywords: Soil organic carbon mapping, machine learning, climate conditions, deep learning, Iran

Acknowledgements

Ruhollah Taghizadeh-Mehrjardi has been supported by the Alexander von Humboldt Foundation under the grant number: Ref 3.4 - 1164573 - IRN - GFHERMES-P. Roozbeh Valavi is supported by an Australian Government Research Training Program Scholarship and a Rowden White Scholarship.