



Advances in Soil Mapping: Exploring Machine Learning Techniques and Their Applications

Ian Eyre*

Department of Agricultural Engineering, The University of Queensland, Gatton, Australia

DESCRIPTION

Soil mapping is the process of creating spatial representations of the distribution and characteristics of soil. Soil maps are useful for various applications, such as agriculture, environmental management, land use planning, and climate change adaptation. However, soil mapping is often challenging due to the complexity and variability of soil, the scarcity and uncertainty of soil data, and the high cost and time required for soil sampling and analysis. Machine learning (ML) is a branch of artificial intelligence that involves developing algorithms that can learn from data and make predictions or decisions. ML has been increasingly applied to soil mapping to overcome some of the limitations of traditional methods, such as statistical regression or expert knowledge. ML can handle large and complex datasets, capture non-linear and spatial relationships, and provide accurate and robust predictions of soil properties and classes. There are many ML methods that have been used for soil mapping, each with its own advantages and disadvantages.

Common machine learning methods for soil mapping

Support Vector Machines (SVM): SVM is a supervised learning method that aims to find the optimal hyper plane that separates the data into different classes or predicts a continuous value. SVM can handle high-dimensional and non-linear data, but it requires careful selection of parameters and kernels, and it may be computationally intensive for large datasets.

Multivariate Regression (MVR): MVR is a statistical method that models the relationship between one or more explanatory variables (covariates) and one or more response variables (soil properties or classes). MVR can be linear or non-linear, depending on the functional form of the model. MVR is simple and interpretable, but it may suffer from multicollinearity, over fitting, or under fitting.

Regression Trees (RT): RT is a supervised learning method that

recursively partitions the data into homogeneous subsets based on a series of binary splits on the covariates. The prediction for each subset is the mean or mode of the response variable in that subset. RT can handle non-linear and categorical data, but it may be prone to over fitting or instability.

Cubist: Cubist is a supervised learning method that combines RT and MVR. It builds a tree-like structure of rules that define linear regression models for each subset of the data. Cubist can handle non-linear and categorical data, and it can provide prediction intervals and variable importance measures.

Random Forest (RF): RF is an ensemble learning method that combines multiple RTs into a single model. Each tree is grown on a bootstrap sample of the data and a random subset of the covariates. The prediction for each observation is the average or majority vote of the predictions from all trees. RF can handle non-linear and categorical data, and it can provide measures of prediction accuracy and variable importance.

Gradient Boosting Machines (GBM): GBM is an ensemble learning method that iteratively fits weak learners (usually RTs) to the residuals of the previous learners. The final prediction is the weighted sum of all learners. GBM can handle non-linear and categorical data, and it can provide measures of prediction accuracy and variable importance. However, GBM requires careful tuning of parameters and may be sensitive to outliers.

Applying machine learning methods to soil mapping

- Generating soil databases for the target soil property or class
- Deriving and selecting environmental covariates based on the SCORPAN framework (soil, climate, organisms, relief, parent material, age, and spatial position)
- Model calibration, validation, and parameter tuning using cross-validation or independent datasets
- Spatial prediction using the calibrated model
- Interpolation or extrapolation of the prediction function if required
- Accuracy assessment using statistical or graphical methods

Correspondence to: Ian Eyre, Department of Agricultural Engineering, The University of Queensland, Gatton, Australia, E-mail: yreian@qa.edu.au

Received: 04-Aug-2023, Manuscript No. AGT-23-23031; **Editor assigned:** 07-Aug-2023, PreQC No. AGT-23-23031; **Reviewed:** 21-Aug-2023, QC No. AGT-23-23031; **Revised:** 28-Aug-2023, Manuscript No. AGT-23-23031; **Published:** 04-Sep-2023, DOI:10.35248/2168-9881.23.12.333

Citation: Eyre I (2023) Advances in Soil Mapping: Exploring Machine Learning Techniques and Their Applications. Agrotechnology. 12:333.

Copyright: © 2023 Eyre I. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

- Using deep learning (a type of ML) method for digital soil mapping in Chile. The authors used a Convolutional Neural Network (CNN) model that takes inputs as images of covariates and explores spatial contextual information by finding non-linear local spatial relationships of neighboring pixels. The CNN model reduced the error by 30% compared with conventional techniques that only used point information of covariates.
- Using ensemble ML method used for digital mapping of soil properties in Ethiopia. The authors compared six ML methods (SVM, MVR, RT, Cubist, RF, and GBM) for predicting six soil properties (pH, organic carbon, clay content, sand content, cation exchange capacity, and base saturation) at four depths (0 cm-10 cm, 10 cm-30 cm, 30 cm-60 cm, and 60cm-100cm). The results showed that GBM performed best for most soil properties and depths, followed by RF and Cubist.

- Using ML-based pedo-transfer functions for estimating soil hydraulic properties in Germany. The authors used four ML methods (SVM, MVR, RT, and RF) to estimate soil water retention and hydraulic conductivity from basic soil properties (texture, organic carbon, and bulk density). The results showed that RF performed best for both soil hydraulic properties, followed by SVM and RT.

Therefore, ML methods should be used with caution and in combination with other sources of information, such as expert knowledge, soil surveys, and field observations. ML methods should also be evaluated and updated regularly to ensure their validity and applicability. ML methods should not be seen as a substitute for soil mapping, but rather as a tool to enhance and complement it.