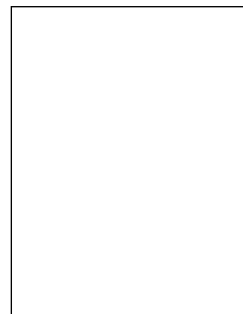


Accepted Manuscript

Title: An Insight into Machine Learning Algorithms to Map the Occurrence of Soil
Matter Horizon in the Northeastern Qinghai-Tibetan Plateau

Author: ZHI Junjun, ZHANG Ganlin, YANG Renmin, YANG Fei, JIN Chengwei,
LIU Feng, SONG Xiaodong, ZHAO Yuguo, LI Decheng



PII: S1002-0160(17)60481-8
DOI: 10.1016/S1002-0160(17)60481-8
Reference: NA

To appear in:

Received date: NA
Revised date: NA
Accepted date: NA

Please cite this article as: ZHI Junjun, ZHANG Ganlin, YANG Renmin, YANG Fei, JIN Chengwei, LIU Feng, SONG Xiaodong, ZHAO Yuguo, LI Decheng, An Insight into Machine Learning Algorithms to Map the Occurrence of Soil Matter Horizon in the Northeastern Qinghai-Tibetan Plateau, *Pedosphere* (2017), 10.1016/S1002-0160(17)60481-8.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

An Insight into Machine Learning Algorithms to Map the Occurrence of Soil Mattic Horizon in the Northeastern Qinghai-Tibetan Plateau

ZHI Junjun¹, ZHANG Ganlin^{1,2,*}, YANG Renmin^{1,2}, YANG Fei^{1,2}, JIN Chengwei¹, LIU Feng¹, SONG Xiaodong¹, ZHAO Yuguang¹, LI Decheng¹

¹State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008(China)

²University of the Chinese Academy of Sciences, Beijing 100049(China)

ABSTRACT

Soil diagnostic horizons, which have a set of quantified properties, play a key role in soil classification. However, they are difficult to predict, and little attempt has been made to map their spatial occurrence. We evaluated and compared four machine learning algorithms, namely classification and regression tree (CART), random forest (RF), boosted regression trees (BRT), and support vector machine (SVM), to map the occurrence of soil mattic horizon in the northeastern Qinghai-Tibetan Plateau using easily accessible ancillary data. The mechanism of resampling and ensemble techniques significantly improved prediction accuracies (measured by area under the Receiver Operator Characteristic curve score, AUC), and produced more stable results of the BRT (AUC of 0.921 ± 0.012 , mean \pm standard deviation) and RF (0.908 ± 0.013) models compared with CART (0.784 ± 0.012) which is the most commonly used machine learning method. Although the SVM model yielded a comparable AUC value (0.906 ± 0.006) to RF and BRT, it is sensitive to parameter settings which are extremely time-consuming, thus we consider it inadequate for occurrence-distribution modeling. Considering the obvious advantages in high prediction accuracy, robust to parameter settings, the ability to estimate uncertainty in prediction, and easy interpretation of predictor variables, BRT seems to be the most adequate method. The outcomes provide an insight into the use of machine learning algorithms to map the mattic horizon, and perhaps others.

Key Words: digital soil mapping; machine learning algorithm; Qinghai-Tibetan Plateau; soil diagnostic horizons

INTRODUCTION

Soil diagnostic horizons provide baseline information for soil classification, agricultural activities, and ecological management. On the Qinghai-Tibetan Plateau, low temperatures and water stagnation in the

*Corresponding author. E-mail: glzhang@issas.ac.cn.

Download English Version:

<https://daneshyari.com/en/article/11012938>

Download Persian Version:

<https://daneshyari.com/article/11012938>

[Daneshyari.com](https://daneshyari.com)