SOILD

Interactive
comment

Interactive comment on "Development of pedotransfer functions for tropical mountain soilscapes: Spotlight on parameter tuning in machine learning" *by* Anika Gebauer et al.

Anonymous Referee #2

Received and published: 14 January 2020

The objective of the study is to develop pedotransfer function for water contents at four pressure heads (PF 0, PF 0.5, PF 1.5, PF 2.5) for two tropical mountain regions with high soil organic carbon content. Boosted regression tree technique was used to fit the models for both areas considering two tuning procedures to determine the regression tree-model parameters (n.tree, shrinkage, interation depth, bag fraction): grid search and differential evolution, the latter showing better results on the water retention estimates for the both areas. The work also compared the performance of the proposed PTFs with other PTFs from literature confirming the better performance of the proposed models. I congratulate the authors for the effort in collecting physico-chemical and hydrological data in such atypical soils and for using innovative techniques, such

Printer-friendly version

Discussion paper

as the differential evolution, in order to get better results on the models fits. I also congratulate them for developing PTFs for organic soils which are not so common in the literature. The work in well written and structured and the subject is well posed Some general and specific comments are summarized below: a) Line 45: In organic Finnland soils? b) It was not clear in the text why you have chosen the boosted regression tree models (Lines 72-73); c) The sentence in line 97-98 should be reformulated ("It allows representing a research. . ..to the accessible area"). The way it was written was unclear to me. d) Line 105. Sometimes outliers should carry important information from the studied area. You should detail the reason of removing them. e) The description of the boosted regression tree should be improved by describing clearer its fitting procedure (Lines 110-115). f) Line 144: What the word "respectively" is related to? g) Line 182: After explaining the reasons for excluding the outliers it would be interesting to inform the range of their values for each soil property; h) Lines 182-190: What is right and left- skewed distribution? It is not clear. i) Line 199: I suggest to correct this sentence: "..organic matter being characterized by a high water holding capacity" to *organic matter which is associated with soils with a high water holding capacity* j) Line 218: How the scaled water retention value was defined? This needs to be clarified; k) Line 230: Change Fig.11 and 12 to Fig 8 and 9; l) Line 240: Change Section 3.2 to Section 3.3; m) Line 245: "PSD measurements were not included..in this area". This sentence should go to Section 3.1 when you call Fig.4 in the text. n) Line 234: Change Fig.9 a-d and 10 a-d to Fig.8 a-d to Fig.9 a-d; o) Lines 253-263: âĂć Did you apply the PTFs from the literature to the Laipura soils considering their range values applicability? âĂć The test Laipura soils were included in the calibration of the proposed PTFs (BRT PTF – Table 2)? This need to be clarified. p) Avoid vague sentences: ex: "these values are.."(Line 268), "in this case"(Line 320), "this might.." (line 320), "This might also result" (lines 325-326); q) The code of the proposed models should be presented; r) Is it possible to provide the study database to the readers?

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2019-72, 2019.