Geophysical Research Abstracts
Vol. 21, EGU2019-14320, 2019
EGU General Assembly 2019

Random forest based reclassification of soil texture for hydrological modelling

Milan Cisty, Veronika Soldanova, and Frantisek Cyprich

Slovak University of Technology Bratislava, Faculty of Civil Engineering, Department of Land and Water Resources Management, Bratislava, Slovakia (milan.cisty@stuba.sk)

When modelling hydrological processes in a watershed, it is usually necessary to define the soil properties in a given area to specify, e.g., the ratio of infiltration and runoff, evaporation, transpiration, and other processes in the river basin. The classification of soils according to their texture is one of the basic methods used for soil description. For this purpose, the soils are described using different classification systems.
Many environmental problems in which soil data serves as an input to simulation models are not restricted to national boundaries and therefore require international cooperation. However, only a few countries use the same particle–size fractions in their classification systems for soil textures. Therefore, the transformation of particle–size texture descriptions between various systems is needed. This study proposes a texture system reclassification by the proposed machine learning model which is verified by soil on a data set from Slovakia originally labelled by national texture classification system. Various theoretical parametric particle size distribution models exist which could help with soil texture classification transformations. However, there is usually no best parametric model that is superior under all circumstances. Various parametric models frequently show different degrees of precision in different soils and different environments. One of the possible solutions to this problem is the application of the ensemble machine learning methodology, which uses the best features of various parametric models for achieving more general results. The goal of the ensemble methodology (a random forest as a basic learning tool is used herein) is to combine the outputs of several models in order to improve the generalizability/robustness that could be obtained from any of the constituent models.
This paradigm was verified by using a soil dataset from Slovakia; however, the methodology proposed is also appropriate when dealing with soil texture classification systems used in other countries. An improvement in precision was demonstrated in the case study, and it is documented that the ensemble model worked better than any of its constituents.
Acknowledgements. This work was supported by the Slovak Research and Development Agency under Contract No. APVV-15-0489 and by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, Grant No. 1/0662/19.