An automated machine learning based ensemble approach for improving estimates of soil water retention parameters

Hao Chen¹,², Tiejun Wang¹,², and Yonggen Zhang¹,²

¹Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Tianjin, China
²Tianjin Key Laboratory of Earth Critical Zone Science and Sustainable Development in Bohai Rim, Tianjin University, Tianjin, China

Accurately mapping soil water retention parameters is vital for modeling atmosphere-land interactions but is challenged by limited measurements and simulations globally. Ensemble pedotransfer functions (PTFs) have been highly recommended for use due to the higher reliability of ensemble models and the error compensation among ensemble members. However, conventional ensemble approaches assign a fixed weight to each PTF and may not fully utilize the strengths of individual PTFs. In this work, we developed a new ensemble approach based on an automated machine learning workflow to assign varying weights to assemble 13 widely used PTFs. The AutoML-assisted ensemble approach (AutoML-Ens), as well as the simple average (MEAN), Bayesian Model Average (BMA), and the hierarchical multi-model ensemble approach (HMME), were evaluated using the global coverage National Cooperative Soil Survey (NCSS) Soil Characterization Database. Results indicate that AutoML-Ens approach performs better than the conventional approaches in terms of the coefficient of determination ($R^2$) and root mean square error (RMSE). Three soil hydraulic parameters, i.e., saturated water content, field capacity, and wilting points, and their corresponding uncertainties, were further derived through the AutoML-Ens approach at a 30”x30” geographical spatial resolution based on a global soil composition database (SoilGrids), which can be applied in the Earth System Modeling. This study demonstrated the necessity of dynamic weights assigning in ensemble approaches and the great potential of coupling data-driven (here, the AutoML) and modeling (empirically or physically-based PTFs) approaches in mapping global soil water retention-like parameters.