Geophysical Research Abstracts
Vol. 17, EGU2015-9886, 2015
EGU General Assembly 2015

Using boosted regression trees to predict the near-saturated hydraulic conductivity of undisturbed soils

John Koestel (1), Michel Bechtold (2), Helena Jorda (3), and Nicholas Jarvis (1)

(1) Swedish University of Agricultural Sciences, Soil and Environment, Uppsala, Sweden (john.koestel@slu.se), (2) Thünen Institute of Climate-Smart Agriculture, Bundesallee 50, 38116 Braunschweig, Germany, (3) KU Leuven, Division of Soil and Water Management, Celestijnenlaan 200e, P.O. Box 2411, 3001 Leuven, Belgium

The saturated and near-saturated hydraulic conductivity of soil is of key importance for modelling water and solute fluxes in the vadose zone. Hydraulic conductivity measurements are cumbersome at the Darcy scale and practically impossible at larger scales where water and solute transport models are mostly applied. Hydraulic conductivity must therefore be estimated from proxy variables. Such pedotransfer functions are known to work decently well for e.g. water retention curves but rather poorly for near-saturated and saturated hydraulic conductivities. Recently, Weynants et al. (2009, Revisiting Vereecken pedotransfer functions: Introducing a closed-form hydraulic model. Vadose Zone Journal, 8, 86-95) reported a coefficients of determination of 0.25 (validation with an independent data set) for the saturated hydraulic conductivity from lab-measurements of Belgian soil samples.

In our study, we trained boosted regression trees on a global meta-database containing tension-disk infiltrometer data (see Jarvis et al. 2013. Influence of soil, land use and climatic factors on the hydraulic conductivity of soil. Hydrology & Earth System Sciences, 17, 5185-5195) to predict the saturated hydraulic conductivity (Ks) and the conductivity at a tension of 10 cm (K10). We found coefficients of determination of 0.39 and 0.62 under a simple 10-fold cross-validation for Ks and K10. When carrying out the validation folded over the data-sources, i.e. the source publications, we found that the corresponding coefficients of determination reduced to 0.15 and 0.36, respectively. We conclude that the stricter source-wise cross-validation should be applied in future pedotransfer studies to prevent overly optimistic validation results.

The boosted regression trees also allowed for an investigation of relevant predictors for estimating the near-saturated hydraulic conductivity. We found that land use and bulk density were most important to predict Ks. We also observed that Ks is large in fine and coarse textured soils and smaller in medium textured soils. Completely different predictors were important for appraising K10, where the soil macropore system is air-filled and therefore inactive. Here, the average annual temperature and precipitation where most important. The reasons for this are unclear and require further research. The clay content and the organic matter content were also important predictors of K10.

We suggest that a larger and more complete database may help to improve the prediction of K10, whereas it may be more fruitful to estimate Ks statistics of sampling sites instead of individual values since the Ks is highly variable over very short distances.