Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-552-AC2, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Mapping soil hydraulic properties using random forest based pedotransfer functions and geostatistics" *by* Brigitta Tóth et al.

Brigitta Tóth et al.

szatmari@rissac.hu

Received and published: 11 February 2019

General comments

This is an interesting manuscript investigating an important topic. The manuscript is well structured, but it is difficult to follow in places. Especially the methods need further clarification and details.

A: Thank you for the review and suggestion for clarification. We will improve the intelligibility of the text, especially of the methods' description. Please find our answers and proposed solutions for your specific comments hereinafter.

Specific comments

Q1: Lines 11-13: very long sentence

A1: We will rephrase it:

Spatial 3D information on soil hydraulic properties for areas larger than plot scale are usually derived by indirect methods, such as pedotransfer functions (PTFs) due to lacking measured information on those. PTFs describe the relationship between the desired soil hydraulic parameter and easily available soil properties based on a soil hydraulic reference dataset. Soil hydraulic properties of a catchment or region can be calculated with applying PTFs on available soil maps.

Q2: Lines 13-16: not clear

A2: Thank you for highlighting it, we will rephrase the sentence according to the followings:

Our aim was to analyse the performance and spatial distribution of soil hydraulic properties derived with (i) indirect (using PTFs) and (ii) direct (geostatistical) mapping methods. We performed the study on Balaton catchment in Hungary, where density of measured soil hydraulic data fulfils the requirements of geostatistical methods.

Q3: Abstract: in general: a lot of acronyms for an abstract. Not clear and difficult to understand what direct and indirect are.

A3: We will eliminate the following abbreviations: THS, FC, WP, MARTHA. The direct mapping method is based on geostatistical analysis, in this study we used random forest with kriging. In the indirect mapping method first we derive the prediction function of soil hydraulic properties. We will try to make it clearer with small edits in the abstract: "... soil hydraulic properties derived with (i) indirect (using PTFs) and (ii) direct (geostatistical) mapping methods. ... As a direct, thus geostatistical method random forest combined with kriging (RFK) was applied ..."

Q4: Introduction: I think a language revision is needed. The first paragraph for example is difficult to understand.

A4:. Thank you for the suggestion. Along the suggestions of the two reviewers we made a lot of corrections for the improvement of the manuscript. If necessary, we will turn to a native English speaker for further corrections and improvements.

Q5: Lines 19-22: this is not always true. Please discuss.

A5: Thank you for highlighting it. The sentence will be modified similarly to the following: “It has been shown, that often but not always the more models are combined for the prediction the more accurate the result is (Baker and Ellison, 2008; Cichota et al., 2013; Nussbaum et al., 2018; Wu et al., 2018). Although significancy of improvement is often not tested.”

Q6: Section 2.2: the description of the soil hydraulic dataset is not clear. The split between test and training in particular.

A6: Thank you for highlighting it. We will rewrite the description on splitting the dataset into train and test similarly to the following: “ For the construction of PTFs those samples were selected from the MARTHA dataset which have measured values of soil horizons or layers considered as dependent and independent variables. We needed two kinds of predictions: (1) for topsoils where we could include organic matter content, calcium carbonate content and pH among the predictors and (2) for subsoils without the above soil chemical parameters, because those are not available for the 30-60 and 60-90 cm soil depths on the Balaton catchment. First we randomly selected 67% of the samples from those which have data on the dependent and all the independent variables available on the catchment area to derive the PTFs. The rest 33% was used to compare the performance of the PTFs, this we called TEST_CHEM set. In the second step we needed a training (67% of data) and test set (33% of data) also for subsoil prediction for which we didn’t have to apply the restriction on the soil chemical properties, therefore we could include more samples for the analysis. As test set we used

the samples of the TEST_CHEM set and further added cases to reach the 33% of the complete data appropriate for subsoil predictions. Again the left 67% was used for the training. Number of samples used to train and test the PTFs was 8157 and 12039 for THS, 8051 and 11931 for FC, 8195 and 12036 for WP, with and without soil chemical properties respectively.”

We will move the above paragraph under 2.4.1 to clarify that this splitting was performed for the PTF approach.

Under 2.2 we will include information only on the soil hydraulic dataset, adding some clarification about what kind of data is included in the dataset: "For the prediction of soil hydraulic properties based on soil properties and other environmental variables we used the Hungarian Detailed Soil Hydrophysical Database (Makó et al., 2010) extended with topographical, meteorological, geological information and remotely sensed vegetation properties (Table 1), called MARTHA ver 3.0 (acronym of the Hungarian name of the dataset). MARTHA consists of 15142 soil horizons’ data belonging to 3970 soil profiles. The samples in it have measured information on basic soil properties – e.g. soil depth, organic matter content, clay, silt and sand content, calcium carbonate content, pH, etc. – and also on soil hydraulic properties such as soil water retention at different matric potential values.”

Q7: The results are well presented, but there are minor (and less minor) problems with typos and structure of the sentences."

A7: Thank you. We will correct the following typos related to RMSE and R2 values similarly to it:

"RMSE values calculated on the test sets for RF were between 0.042 and 0.045 cm3 cm-3 for THS, 0.039 and 0.042 cm3 cm-3 for FC, 0.035 and 0.038 cm3 cm-3 for WP, which is close to the performance of other internationally accepted PTFs (e.g. Botula et al. (2013), Román Dobarco et al. (2019), Zhang and Schaap (2017)). R2 was 0.408-0.487, 0.746-0.766 and 0.737-0.762 for THS, FC and WP respectively on test

sets in the case of RF."

Further typos and structure of the sentences will be checked by native speaker.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-552, 2018.