

Interactive comment on “EstSoil-EH v1.0: An eco-hydrological modelling parameters dataset derived from the Soil Map of Estonia” by Alexander Kmoch et al.

Anonymous Referee #2

Received and published: 21 December 2019

Summary

In the manuscript by Kmoch et al. a new countrywide soil dataset for Estonia at 1:10000 scale is presented. Those soil properties are provided which are the most frequently required soil input variables for eco-hydrological modelling, focusing on providing soil data for the SWAT model. The data originates from the Soil Map of Estonia vector dataset (1:10000), which includes information on soil types according to Estonian soil classification, soil quality, number and depth of soil layers, information on coarse fragments and Estonian texture classes. Numerical soil properties are derived or through using characteristic values of certain soil groups or computing them from available in-

C1

formation, or if data is not available for calculation, data of external dataset is used.

General comments

The scale of the presented soil dataset is outstanding. Detailed information about coarse fragments is unique. Descriptive or categorical type information originating from soil survey is very valuable even if uncertainty is generated when those are converted into quantitative data. The manuscript presents method to derive input information from soil survey data for those models, which require quantitative information about soil properties. This kind of data transformation has several difficulties which authors had to face. Significant amount of work has been put into the construction of the presented dataset, which has to be acknowledged. The work deserves to be published after major revision. Please find hereinafter suggestions for consideration.

Terminology used in international literature should be adapted in the manuscript.

It is not clear what authors mean by “complex text codes” in the abstract.

Please provide more precise information about the meaning of “soil profiles (e.g., layers, depths)” “layer information”, which is mentioned in the abstract and introduction.

Under materials and methods section authors mention that potential fertility was mapped, in the abstract and introduction soil quality is mentioned. It has to be clarified which soil property with which method was mapped, and reference or detailed description on how it was derived is needed. A table including metadata would be very informative in the manuscript, in which variable name, file name, description of variable, units of measure, reference, etc. could be included, e.g. meta file of SoilGrids. The “EstSoil-EH_v1.0_attribute_fields.txt” file could be a starting point for that.

The authors could put into context the novelty of providing data at 1:10000 scale – which scale is outstanding. Information on other national soil datasets – which are considered detailed or high-resolution e.g. <https://dl.sciencesocieties.org/publications/sssaj/pdfs/82/1/186>, etc. – could be

C2

referenced, and the progress presented by EstSoil-EH v1.0 could be highlighted.

Regarding the mapped soil properties, the following specific comments could be considered for the manuscript:

1. Soil type: Is it not clear why new soil types were added to the original dataset, how soil type was extended, e.g. was original soil type recoded based on soil profile information included in the dataset? How were Estonian soil types translated into WRB reference groups? Is there a reference document for it? Based on which soil classification system did you add new soil types and how? Please write down how many soil types were included initially and how many soil types were added. It is not clear how you got 7067 soil types in the attribute table if 120 soil types exist in Estonia. Maybe you meant something different. P4 L28: why "Overall soil type group" is differentiated from "Soil type" which is in L20?

2. Texture classes: Clarification is needed on how USDA soil textural classes and then sand, silt and clay content were derived. Based on present manuscript Estonian soil textural classes were available from the official 1:10000 scale National Soil Map of Estonia. Estonian soil texture class names were translated using USDA terminology. Based on the Estonian texture class names average sand, silt and clay content were added to each soil layers. Please consider to add USDA texture class names based on the average sand, silt and clay content which characterize the Estonian texture classes. Please provide reference for the definition of the Estonian texture classes.

3. Coarse fragments content: It is not clear how - "skeleton indicator number" was derived from the shape and size of the stones and - "inferred rock content (% of volume)" was derived from "skeleton indicator number".

4. Soil organic carbon content: It has to be described why measured SOC data was averaged by soil units in the training dataset for deriving SOC prediction. Was not it possible to use soil profile data to derive the prediction? Predictors used in the random forest method could be listed under materials and methods section. Performance of

C3

SOC prediction could be included in a table. Variable importance could be shown in a figure.

5. Bulk density: It is mentioned that BD is calculated based on texture and SOM, but texture is not included in Equation 4. It has to be considered that moist bulk density is required for SWAT.

6. Potential fertility: It is listed under materials and methods, but not included under the results. Reference or description for the computation would be needed.

7. Organic horizon thickness: Similarly to potential fertility, it is mentioned under materials and methods, but not discussed in results section. Do you mean thickness of A horizon or thickness of soil horizon with accumulation of humified organic matter? Please add reference.

8. Please clearly state for which soil properties the performance could not be analysed because of lacking measured data.

Some parts of the manuscript could be simplified by decreasing mainly technical descriptions, e.g.: P6 L3-P7 L3.

More descriptive plots and tables could be provided for the readers about the derived dataset.

Language revision of the manuscript would improve its readability.

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-192>, 2019.

C4