Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-36-AC1, 2020

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Updated European hydraulic pedotransfer functions with communicated uncertainties in the predicted variables (euptfv2)" *by* Brigitta Szabó et al.

Brigitta Szabó et al.

tothbrigitta.mail@gmail.com

Received and published: 27 July 2020

Response to Anonymous Referee #2

Thank you for the review and constructive comments. We will address the comments in a revised version of the article. Below we give details on exactly how we address the concerns raised by anonymous referee 2. Please note the following during reading the responses: - the supplement file includes the structured and formatted version of the responses, - the responses are in blue regular font and follow the referee's questions (RC2), - new text parts that will be added to the manuscript are in blue italic font, - the reference to the lines and pages relates to the discussion paper available from:

Printer-friendly version

Discussion paper

https://gmd.copernicus.org/preprints/gmd-2020-36/gmd-2020-36.pdf .

RC2: Soil pedotransfer functions are important when used for estimation of soil hydraulic parameters in catchment, regional, or continental scale applications. This manuscript improves the estimation of euptfv1 and provides information about prediction uncertainty, and can be applied for more predictor variable combinations than the euptfv1. Overall, the manuscript is interesting, important, well written, and organized in a logical well. Therefore, I recommend accepting this manuscript after minor revisions that are required to address the general and specific comments provided below.

A: Thank you for the positive general comment.

RC2: 1. The authors compared the estimation of water content at saturation, field capacity, wilting point, plant available water content, saturated hydraulic conductivity, etc., individually. I think these sections are somewhat lengthy. However, the most interesting part of the comparisons between point and parameter predictions and euptfv1 and v2 are very short. Is it possible to extend the comparisons and the discussion?

AGREED.

A1: Regarding comparison between point and parameter predictions, we will be more specific by adding the following:

P12 L5: " . . . more accurate and for further 8 cases RMSE were smaller."

P12 L6: "The reason for higher RMSE in parameter estimation can be that the VG model does not always adequately describe the measured MRC data (Weber et al., 2019). Therefore, when THS, FC, FC_2 and WP are computed with parameter estimation those are not only affected by the uncertainty of the prediction of VG parameters but by the goodness of VG model fit as well."

P12 L8: will rephrase the sentence to make it clearer, which now reads: "For THS point estimation performed better than parameter estimation. When the moisture retention curve is not needed, but only THS and/or FC/FC_2 and/or WP, we recommend to compute those with . . .”

In order to include the suggested comparison, between euptfv1 and v2, we will include the following sentences

P12 L12: “The most important reason for it can be that the interaction between the target variable and the predictors is more complex for the cases of predicting FC or VG parameters – to describe the MRC –, which can be untangled using random forest. This may provide a reasons the random forest algorithm performed significantly better than the PTFs derived with linear regression or a simple regression tree.”

P12 L14: "The RMSE of THS prediction was somewhat lower for euptfv1 than for euptfv2, but the difference was not significant. It could be due to the close to linear relationship between THS and BD and high relative importance of BD in THS prediction (84 %). This way their interaction can be efficiently described with the linear regression which is capable to extrapolate as well. Extrapolation with the random forest algorithm is not possible outside the training data, which can limit its performance. The general improvement of the PTFs in euptfv2 is threefold, it is due to i) using random forest instead of single regression tree or linear regression, ii) including more detailed information on soil sampling depth, not only distinguishing topsoils and subsoils and iii) providing information on prediction uncertainty.

Regarding the description of the individual point and parameter estimations we will keep the details because we think it is instructive to provide information about the importance of specific predictors.

RC2: 2. The authors listed so many PTFs. When I was reading the conclusion part, I cannot find which PTF I should use. Is it possible to make some concluding remarks regarding which PTFs should be used for corresponding predictors? I think this will be very helpful for future readers.

AGREED.

A2: Thank you very much for this very helpful comment. Indeed, it is very important that users should easily understand which PTF to select and apply. To achieve this, 1.) we will add a dedicated paragraph on it above the Conclusion section, 2.) will highlight in the abstract and short summary that this section is provided, 3) we will move Table S3 from supplementary material to the manuscript as Table 11.

The new paragraph 4 reads:

“4. Practical guidance on how to use the PTFs The minimum input requirements for all PTFs are sand, silt and clay content, and soil depth. Soil depth needs to be considered in regard to the depth of the other input properties and soil hydraulic data needs, e.g. if the soil hydraulic properties of the top 20 cm (0-20 cm) is needed, then depth needs to be set at 10 cm in the input data of the prediction. If only soil texture information is available for the predictions, the class PTFs from euptfv1 could be applied (Tóth et al., 2015). We emphasise that: 1. the units of input soil properties (predictors) have to be the same as indicated in the text and that the sand, silt, and clay are defined by the following particle diameters: clay < 2 μm, silt between 2 and 50 μm, and sand between 50 and 2000 μm, 2. when only specific water content values at saturation, field capacity or wilting point are required (ie. THS, FC_2, FC, WP) it is recommended to use point PTFs. This is also true for the prediction of KS, 3. for AWC, the most accurate way is by first predicting FC and WP with the point predictions and then compute AWC using Eq. (1), and similarly for AWC_2 using FC_2 and Eq. (2), 4. it is recommended to do the VG prediction if only moisture retention curve parameters are needed, and 5. the MVG prediction when both moisture retention and hydraulic conductivity parameters are required. The VG algorithms predict the following van Genuchten model parameters: the residual water content ÏŚr (cm^3 cm-3), the saturated water content ÏŚs (cm^3 cm-3), and shape parameters α (cm-1) and n (-). Parameter m is provided based on m=1-1/n (van Genuchten, 1980), and for the hydraulic conductivity curve, the two additional parameters: K0 (cm day-1) the hydraulic conductivity acting as a matching point at saturation and L, the shape parameter related to pore tortuosity (-).

Table 11 shows the recommended PTFs for each predicted soil hydraulic property and available predictor variables. The users need to check which basic soil properties are available for the predictions, then look in Table 11 which PTF is recommended to use. The algorithms have been implemented in a web interface to facilitate the use of the PTFs, where the PTFs' selection is automated based on soil properties available for the predictions and required soil hydraulic property. The Code and data availability section provides information on how to access this resource."

The additional text in the short summary and the abstract is given by:

Short summary: "... The influence of predictor variables on predicted soil hydraulic properties is explored and practical guidance on how to use the derived PTFs is provided. . . ."

Abstract: ". . . for the prediction of water content at -100 cm matric potential head and plant available water content. A practical guidance on how to use the derived PTFs is provided."

SPECIFIC COMMENTS:

RC2: 1. Figures 2, 5, and 6: Is it possible to include R2 in these figures? This will make the comparison between different figures easier.

AGREED

A1: We will add R2 to Figures 2, 5, 6 and S1, e.g.: Fig_1_response.

RC2: 2. In the abstract and conclusion sections: -15.000 should be -15,000

AGREED

A2: Thank you for noting it, we will correct it in the entire text.

RC2: 3. Page 6, line 4: why did the authors utilize median values instead of mean values? Nothing changed.

A3: Our aim was to provide information about the uncertainty of the predictions, therefore we applied quantile regression forests. This way the most probable predicted response value is at the 50th percentile, i.e. the median, which is considered more robust against the outliers than the mean. In this way we decided to use the median as the predicted value (yhat) rather than the mean.

RC2: 4. Page 7, line 19: “in the study of (Khodaverdiloo et al., 2011)” should be “in the study of Khodaverdiloo et al. (2011)”

AGREED

A4: Thank you, we will correct it.

RC2: 5. Page 10, line 4: “and RMSE” should be “an RMSE”

AGREED

A5: Thank you, we will correct it.

RC2: 6. Page 10, Line 27: “;” should be “,”

AGREED

A6: Thank you, we will correct it.

RC2: 7. Page 12, line 14: add a connection/linking word before “it is due to”

AGREED

A7: The text will be rephrased to: “The improvement of the PTFs is twofold, the better performance is due to . . .”.

Please also note the supplement to this comment:
https://gmd.copernicus.org/preprints/gmd-2020-36/gmd-2020-36-AC1-supplement.pdf

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-36,

2020.

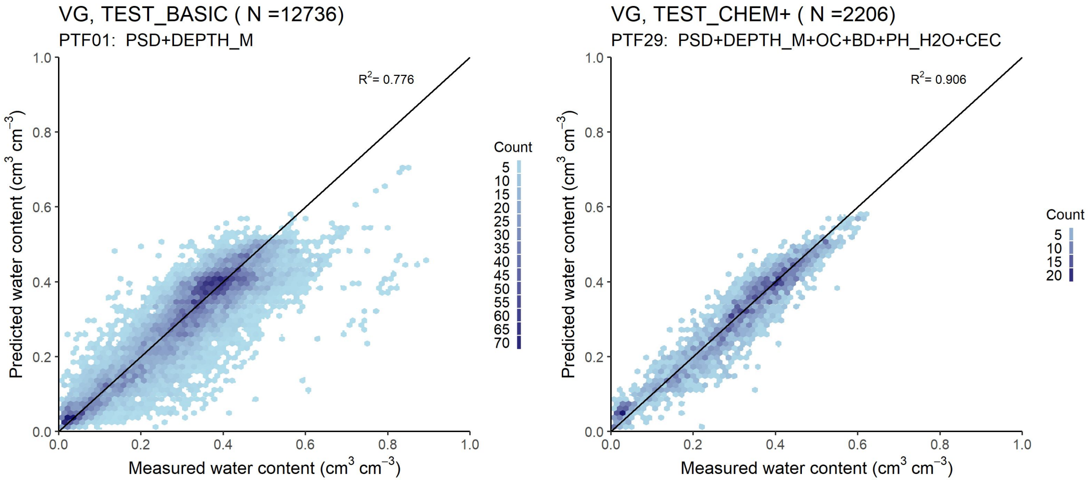

Fig. 1. Fig_1_response

Printer-friendly version

Discussion paper