SOIL Discuss.,
https://doi.org/10.5194/soil-2020-62-RC1, 2020

Interactive comment on "Improved calibration of Green-Ampt infiltration in the EROSION-2D/3D model using a rainfall-runoff experiment database" *by* Hana Beitlerová et al.

Mehdi Rahmati (Referee)

mehdirmti@gmail.com

Received and published: 26 November 2020

Review Report on Manuscript "soil-2020-62" entitled "Improved calibration of Green-Ampt infiltration in the EROSION-2D/3D model using a rainfall-runoff experiment". The paper is dealing with infiltration submodule of the EROSION-2D/3D model. Authors are trying to provide some pedo-transfer functions (PTFs) to optimize the prediction of the needed calibration factor in infiltration submodule, named skinfactor. The skinfactor can be determined experimentally from rainfall-runoff or infiltration experiments with the hillslope simulation tool EROSION-2D. However, authors believe that the process is time consuming, labor-extensive, and limited from several aspects. So, the determina-

tion of the skinfactor may restrict the usage of the EROSION-2D/3D model in the cases where the experimentally data of rainfall-runoff or infiltration are missed. So, the parameter catalogue of the model provides extra studies to predict skinfactor from several readily-available parameters. However, authors believe that the studies are limited and provide data for limited conditions. Authors also criticize the provided regression models for their weak determination coefficients. Therefore, authors provide a R package enabling automated and batch determination of the skinfactor for, what they called, an extensive rainfall-runoff infiltration experiment dataset. They used database containing 464 experiments only, from which around 191 experiments are excluded from further analysis. Overall, I found the study very interesting and worthwhile. However, some questions arise when reading the paper: 1) The first question is that what is the reason authors have limited themselves to limited number of the experiments while they themselves are criticizing the model's provider providing skinfactor prediction with limited number of experiments (I think, 116 experiments). Since the skinfactor is predictable from both rainfall-runoff and infiltration experiments, why not to use infiltration experiments which are available in higher numbers. For example, SWIG database (Rahmati et all. 2018 , https://essd.copernicus.org/articles/10/1237/2018/) provides more than 5000 infiltration experiments (including 374 rainfall simulator experiments) from all around the world that can be used to provide a global PTF for skinfactor predictions. 2) The second question is that why the linear fixed-effect model is used to develop the PTF and why nonlinear methods (or let say machine learning methods) are not used? I understand that linear mixed-effect method is much stronger than simple regression methods since it accounts for both explained and non-explained variations in independent variable. However, the relationship between skinfactor and soil readily-available parameters seems more nonlinear to me and I think machine learning method may act much better than the linear mixed-effect method. 3) I believe that the reason why the working group plays an important role (as input parameter) in prediction of skinfactor is that the used database is not global enough. I believe if we use a larger database, we can simply provide a global PTF being free from working groups effects. 4) In the

MM section, please clarify that how you have determined experimentally the skinfactor for PTFs development. 5) In model selection section, I see authors have correctly divided the database into two groups of training and validation subsets. However, there should be one more step to assess the reliability of the developed PTFs. What authors have done is only assessing the accuracy of the models. However, the accuracy may be rooted in chance since you divide the database into training and validation subsets randomly. So, I suggest authors to repeat the process of splitting data into training and validation subsets 10 times (at least) and calculate the criterions. Finally, check the STD between obtained results of 10 times. It will give you a better understanding of the reliability of the PTFs.

Sincerely Yours, Mehdi Rahmati

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2020-62, 2020.