

Interactive comment on “Systematic comparison of five machine-learning methods in classification and interpolation of soil particle size fractions using different transformed data” by Mo Zhang and Wenjiao Shi

Anonymous Referee #1

Received and published: 17 January 2020

1. when talking about the ilr transformation, authors did not go deep enough about the way of defining the basis. It is known that we can define different Sequential Binary Partition (SBP) matrices, and consequently, different ilr balances will be calculated. We can choose the best balances, among many possible choices. Actually, it is up to compositional data analyst. There are tricks to get to appropriate ilr balances; e.g. those balances with negligible cross-variogram values, to avoid performing co-kriging. The ilr related part of the manuscript in the current form is acceptable but it is suggested to mention that the used ilr balances are a possible set, among the others.

C1

2. The authors have used R-squared to measure the linear correlation of variables that are not linear. As the result, calculated R-squared values are not valid and they cause misinterpretation. R-squared must be replaced by a similar measure for non-linear data, such as Pearson rank correlation etc. R-squared values must remove from the manuscript or replace by an appropriate measure.

Please also note the supplement to this comment:

<https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-648/hess-2019-648-RC1-supplement.pdf>

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2019-648>, 2020.

C2