SOIL Discuss.,
https://doi.org/10.5194/soil-2019-72-SC1, 2019

Interactive comment on "Development of pedotransfer functions for tropical mountain soilscapes: Spotlight on parameter tuning in machine learning" *by* Anika Gebauer et al.

Alexandre Wadoux

alexandre.wadoux@sydney.edu.au

Received and published: 19 November 2019

I agree with the authors that a grid-search process is by far the most used method for parameter tuning in Pedometrics, but claiming that (L. 64) differential evolution has been applied for the first time in Gebauer et al. (2019) is not correct. Differential evolution is routinely used in Pedometrics since several years, in particular to find optimal values of parameters, see for example https://doi.org/10.1016/j.geoderma.2018.03.010 or https://doi.org/10.1016/j.catena.2016.02.016. For parameter tuning in ML applied to soil mapping, Wu et al. (2016) (10.1007/s11368-016-1374-9) compared a genetic algorithm, Particle Swarm optimization and a grid search process to find optimal ML tuning

parameters.

Without any surprise an optimization algorithm leads to more optimal parameter values than a grid-search process. This is obvious because a global optimization algorithm searches for any possible value within pre-defined boundaries while a grid-search is limited to a user-defined number of values. It should be noted that grid-search parameter tuning is by far the most used because the user knows in advance the number of iterations that will be needed. This is impossible to estimate with differential evolution, even though user-defined values in differential evolution can make the optimization to converge faster. This is a major limitation and the main reason why differential evolution (or any global optimization algorithm such as SA or PSO) are not routinely used for ML tuning parameter optimization.

Parameter tuning of ML models is computationally expensive and in most cases differential evolution will be too slow. In my experience differential evolution can need several hundreds to several thousands of iterations to find a global optimum.

For this reason, when ML parameters need to be tuned, other more efficient algorithms are used in the ML literature. Bayesian optimization is one of them. Bayesian optimization has been designed for parameter tuning of ML models but is much faster than other global optimization algorithms. Bayesian optimization finds the optimal tuning parameter values in very few iterations. Another advantage is that the algorithm does not need specific pre-defined boundaries. I personally applied it for ML parameters tuning in https://doi.org/10.5194/soil-5-107-2019.

Can the authors make a plot with in the x axis the number of iterations and in the y axis the value of the tuning parameters? This would be useful to see how the algorithm converges.

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2019-72, 2019.