Dear Mr. Wadoux,

Thank you for your time and valuable comments.
Please find our replies below each comment. They are displayed in blue. Line numbers in the replies refer to the revised manuscript with tracked changes. Line numbers in the reviewer comments refer to the first manuscript version.

Kind regards,
Anika Gebauer

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 .
Further applications of the differential evolution algorithm in soil-related research questions were added to lines 70 ff.. Nevertheless, we would like to emphasize that applications for machine learning parameter tuning in pedometrics are scarce.

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.
The results of Wu et al. (2016) show that mathematical optimization algorithms do not necessarily outperform grid search: tuning two real-valued parameters using grid search led to better results than tuning the same parameters applying optimization algorithms. This shows that a lucky selection of predefined parameter combinations to be tested can result in grid search outperforming optimization algorithms – in particular, if the number of optimization iterations is small. The more values are tested during parameter tuning, the higher the probability of finding the global optimum. According to Wu et al. (2016), the better performance of grid search was caused by grid search exhaustively evaluating every predefined parameter value, while optimization algorithms can get stuck in local optima. Wu et al. (2016) did not mention the number of iterations of the optimization algorithms, but we assume that increasing the number of iterations would have led to results that are at least as good as those achieved by grid search.
Tuning three parameters with the genetic algorithm resulted in a better model performance than tuning two parameters with grid search, the genetic algorithm or the particle swarm optimization (Wu et al. (2016), Fig. 2). The different number of parameters hampers the direct comparison of the applied tuning techniques. The difference in model performance is most likely caused by tuning the third parameter and not by the parameter tuning technique itself.
Even though the benefits of optimization algorithms towards grid search are obvious, further direct comparisons of mathematical optimization algorithms and grid search applied for machine learning parameter tuning in soil related research questions are necessary. We compared the differential evolution algorithm to grid search because it is superior to various optimization algorithms including particle swarm optimization and genetic algorithms (lines 62 ff.).

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.
We agree that it is beneficial to know the number of tuning iterations in advance. However, the differential evolution algorithm allows defining a stopping criterion. In this case, we stopped the differential evolution optimization either after 10 iterations without improvement or a maximum number of 200 iterations (lines 183 f.).
Regardless of the chosen parameter tuning technique, restricting the number of parameter values to be tested reduces the probability of finding the global optimum. For tuning real-valued parameters, it is impossible to know the necessary number of iterations in advance. A trade-off between computing time and the probability of finding the global optimum has to be made for any parameter tuning technique.
We agree that parameter tuning of machine learning models requires computing power. But with computers becoming more efficient and the possibility of parallelization, a parameter tuning technique should be judged based on the predictive power of the resulting machine learning model and not on the required computing resources.
It requires more computational resources to find the global optimum of real-valued parameters by testing every possible parameter combination during grid search instead of converging towards the optimum using a well-adjusted optimization algorithm. Making the wrong decisions in predefining the parameter values to be tested might even prohibit grid search from ever reaching the global optimum in tuning real-valued parameters.

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 .
Thank you for recommending Bayesian optimization. We see the advantages of the proposed parameter tuning technique. However, comparisons of both algorithms led to contradictory results: while some studies found Bayesian approaches to be superior (e.g. Carr et al. 2016*) others reported the opposite result (e.g. Schmidt et al. 2019**; Salt & Howard 2015***). Further comparisons of Bayesian optimization and differential evolution in machine learning parameter tuning applications are necessary concerning computing requirements and the resulting model performance.

* Carr, Garnett and Lo. “BASC: Applying Bayesian Optimization to the Search for Global Minima on Potential Energy Surfaces.” ICML (2016).

** Schmidt, Safarani, Gastinger, Jacobs, Nicolas and Schülke, "On the Performance of Differential Evolution for Hyperparameter Tuning". *International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-8. (2019) doi: 10.1109/IJCNN.2019.8851978

*** Salt, Howard, Indivera, Sandamirskaya, “Differential Evolution and Bayesian Optimisation for Hyper-Parameter Selection in Mixed-Signal Neuromorphic Circuits Applied to UAV Obstacle Avoidance”. CoRR (2015). https://arxiv.org/abs/1704.04853

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.
In accordance with comment C42 of the anonymous referee #1, the number of differential evolution iterations was added to lines 397 ff. The suggested plot was not drawn, as a figure showing the converging algorithm for each tuning parameter and each cross validation fold would be rather confusing.