January 6, 2019

Dear Reviewers,

Please find enclosed a revised version of our manuscript *"Inverse modeling of a single ring infiltration experiment: a key for analyzing unsaturated hydraulic properties of soils?"*.

We would like to thank both referees and editors for their valuable and constructive remarks and comments. All remarks and proposals have been addressed in the current revision. We believe that the revision according to both review reports has increased the level of the paper. First of all we should highlight here, that the title of our paper has been updated for

"Automated calibration methodology to avoid convergence issues during inverse identification of soil unsaturated hydraulic properties."

since we would like to underline the most import contribution of our work – the calibration methodology.

For your convenience, we have highlighted all changes in magenta in the revised manuscript, and we further include bulleted list detailing all changes in the manuscript according the referees' remarks and comments.

Yours sincerely,

Michal Kuraz, Lukas Jacka, Johanna Blöcher, and Matej Leps

Response to the referee's comments

Referee #1

1. *The abstract is very long; I suggest shortening it to convey the major points of this research.*
   answer: Following your advice the abstract has been significantly reduced and rewritten.

2. *Typically, the last paragraph of the introduction section is reserved for describing the aim, innovation, general methodology of the research. This is missing from this manuscript. I like the fact that the authors present their research questions, but they also need to briefly describe the methodology used to answer them and how it is novel.*
   answer: We completely agree with this remark, an extra section according to your advice has been added.

3. *The authors should state the country where the study area is located. Also, a short description of the synthetic problem should go under section 1.2*
   answer: The country of origin has been added to site description in Section 2.2.1, on line 101.

4. *How was the boundary condition of H=-280 cm obtained? Was this based on experimental data?*
   answer: The estimate of the groundwater table is rather an overhead limit, which is still a sufficiently distant from the top soil layer – the layer, which hydraulic properties are investigated. This overhead limit was assumed from the near surface tensionmeter measurements and with the assumption of a hydrostatic state in the unsaturated zone. We are VERY aware, that this estimation of the groundwater table is a subject of a HUGE uncertainty. However, such position of the groundwater table is VERY distant from the evaluated layer, and effects of such a deeply located groundwater table on the infiltration experiment vanish. An explanation has been added in Section 2.2.1, line 103.

5. *The authors have assumed that the water table remains constant during the experiment. What are the potential implications of this assumption?*
   answer: We agree, that this assumption has also several limitations. An explanation was added at line 189, section 2.4.1. The potential implication of this assumption is that the groundwater aquifer adsorbs the entire infiltrated amount without affecting its volume. In theory, this assumption is non-physical. However, for a short time experiment, such as the evaluated infiltration experiment, when the infiltrated amount within the simulation time even doesn't reach the groundwater aquifer, and also with the consideration, that the infiltrated amount is completely negligible in comparison to the expected groundwater aquifer volume, this assumption becomes acceptable.

6. *There is an issue on line 186*
   answer: Thank you for this remark, we have now double-checked the latex compilation log for undefined and multiple defined references.

7. *The main issue I have with the methodology is related to the use of a multi-objective algorithm to solve the inverse problem. The argument that the authors make for using a multi-objective algorithm is not convincing. According to the multi-objective optimization definition the objective functions should be competing. I do not quite understand how the multi-objective optimization can work with non-competing objectives.*
   answer: Using a non-competing multi-objective algorithm is not a standard approach, however it has been proved in the given references, that this strategy can be beneficial also for typically single objective problems. However, we understand, that this strategy should be carefully explained to all readers of our manuscript, and thus a detailed

explanation has been added in between lines 294 – 313, section 2.5.1. We are also aware that we should explain the reader our motivation for such selection of the error criteria, and thus an extra text has been added on lines 286 – 293, section 2.5.1.

8. *More details are needed for the genetic algorithm (about mutation, cross-over parameters, the choice of stopping criteria etc). How were these choices made?*
answer: See updated text in between lines 317 – 325, section 2.5.2. Settings of the genetic algorithm originate from the references given in this section.

9. *On line 337 and elsewhere the authors state that expert knowledge is needed. I understand that expert knowledge is an important part of this methodology, but very little details are given throughout the manuscript about this. Please elaborate on the use of expert knowledge and how it will be incorporated in the methodology.*
answer: Incorporation of the expert knowledge is now explained in section 3, lines 425 – 430.

10. *Line 352 is not clear. Please explain in a more clear way your point about specific storage. I do not see what you explain reflected in figure 8.*
answer: Finally, the specific storage was wiped out from the model due our misinterpretation of objective function response plots. This mistake was now corrected, and the non-zero specific storage forms a part of the parametric set.

Referee #2

Major issues:

I *As this study aims to identify non-uniqueness of parameters it is critical that the inverse problem is solved properly and converged solutions are obtained. It is clear from a number of the objective vs parameter plots that the GA obtained solutions have not converged. What is currently claimed as non-uniqueness could be merely a characterization of the poor convergence of the optimizer as opposed something inherent to the problem. This is confirmed with the statement that the GA strategy may compute dominated solutions, this is not sensible at all as these clearly indicate sub-optimality. Non converged solutions should be refined e.g. GA + Nelder-Mead / gradient based strategy to obtain converged solutions, converged solutions must be ensured at all cost before this work can be published. The authors have to do due diligence here to avoid making improper conclusions about the nature of the problem.*
answer: The problem was treated with multicriteria objective function definition, where the objectives are not competing, see the explanation given on lines 294 – 313, section 2.5.1. In case that our model is not capable to perfectly fit the experimental data so that all error functions are zeroes, which can be achieved on synthetic problems only, then the optimal position at the Pareto front doesn't necessarily need to point to minimum at each objective function. The response plots depict only the first objective function (13). However, respecting your comment, we used the multiobjective genetic algorithm to obtain initial parameter estimates, and then we switched into single objective definition, where only the RMSE error function (13) was minimized, and for this final parameter optimization we employed gradient method – boxed Newton method with constraints (Byrd et al., 1995). With this treatment, we have always achieved fully converged solution for the first error function (13).

II *The authors consider the discretization error and it would significantly enhance the quality of the manuscript if they properly quantify the effect of the discretization error for all local minima and not only selected ones. Hence, parameters for r0, r1 and r2 should be presented (it would be convenient to display them on a single graph) and the variance in the parameters quantified.*
answer: Following your recommendation all response plots are now plotted in a single

graph, so that the effect of the discretization error for the local extremes is now apparent. On the other hand, providing response plots for all extremes except the ones with good fitting qualities is rather counterproductive. The local minima, that were identified with the employed genetic algorithm typically point to parametric subspace with a severe model misfit. In this parametric subregions the objective function can exhibit complicated topology, and the identified local minima are rather *false minima*. Due to huge model misfit, it is not worse to further explore the objective function topology in this misfitting parametric subregions.

III *A number of model simplifications are present in the work and their role on the variance of parameters should be quantified to complete the study. These include:*

1) The ring thickness is modelled as 25mm model in the experiment it is 2mm. It is speculated that the results are insensitive to this simplification. This should be shown and quantified and not merely left to speculation, i.e. half and double the ring diameter and show that the results are insensitive.

2) The affect of the distance from the axi-symmetry (rererred to as axis of anisotropy in the paper) 2cm?. Again half and double to quantify the effect of this simplification on the estimated parameters.

answer: Following you advice this model simplification were validated, the oversized ring thickness was validated in between lines $205 - 221$, and the detached domain was validated in between lines $237 - 258$. However, validating this simplification becomes cumbersome due different triangularizations, which are required for different domain shapes. In order to avoid effects of different domain trigularization, domains were discretized with extremely refined mesh densities.

IV *Deterministic versus statistic: 22 SR experiments are conducted and reduced to an expected value using a smoothing operation. Including the variance of experimental data in this study would significantly enhance the merit of the work. Suggestions is to show the parameters for the expected value, as well as one standard deviation away from the expected value to quantify the variance of the parameters based on the experimental data, this would indicate valid ranges of the parameters that one can expect and allow for a holistic view on ill-posed nature of the problem when including uncertainty in the experimental data.*

answer: We are aware that real world experimental data manipulation must be accompanied with a proper statistical analyses. However, the data used in our paper were already fully analyzed in another data paper (Jačka et al., 2016). However, for the demonstrative purposes of this contribution, we have significantly reduced this data part, and only the mean values, which were given by Jačka et al. (2016) as a representative data set, are analyzed here.

Since this paper is concerned to calibration methodology, data error propagation and further statistical analyzes were not given here, since we presume that such analyses is beyond the scope of this paper. We presume that potential readers of this paper will be interested in how we proceeded with the calibration, and proper uncertainty analyzes would disrupt the reader attention, especially since the current paper extension is exceeding 30 pages. Moreover, we plan to employ the given methodology for analyzing all 22 input data to obtain SHP properties ranges in a standalone paper.

V *Convergence of solver Convergence for certain parameter sets is known to be an issue as highlighted in the paper - quantify the success rate of the simulations in the paper and which parameters failed to converge. A significant amount of effort then it is required that focusses on those designs to improve convergence, it is not evident from the paper that the choices made reflect this.*

answer: With the porposed methodology and our numerical implementation we haven't experienced any single convergence issue on the broad parametric space. During the

paper preparation the computations were repeated for several times. For the initial genetic algorithm 40.000 solver runs are required, however, due our repeated computations, over 500.000 solver runs were executed. In this paper and extra section 4.3 on line 497 was added, were we mention just over 40.000 runs.

Minor issues:

1. *Table 1 explain and define the variables in the text for the benefit of readers not familiar with this field.*
   answer: An extra section 2.1 on line 86 was added.

2. *Double check: should axis of anisotropy (p 7) not be axis of axi-symmetry*
   answer: Double checked, corrected. Thank you for this feedback, I was probably overroccupied with another system, which was anisotropic, while preparing this manuscript.

3. *GAMMA1 and GAMMA2 is not defined until Figure 4 but used in the beginning of Section 2.3.3.*
   answer: Reference to the figure added on the top of the section 2.4.1 on line 159.

4. *Mathematical notation: Eq (11) dH/dn is vector, hence the RHS must also be a vector not a scalar 0.*
   answer: Thank you for this feedback, corrected for L_2 norm.

5. *On page 9 the Section number is missing (??)*
   answer: Thank you for this remark, we have now double-checked the latex compilation log for undefined and multiple defined references.

6. *Check all references to figures e.g. Figure 4.2 Figure 10.*
   answer: Double-checked.

7. PHI *depicted as the objective function in the figures is not defined only* `PHI_1, PHI_2 and PHI_3`. *Define PHI in text.*

   8. P19 `PHI_1 on figure should be PHI}`
   answer: Double-checked and fixed figure description.

8. *Pareto-dominance should be non-dominance.*
   answer: Fixed.