HESS 2021 168

Decision

Dear Dr. A. Jeantet

Thank you for submitted the revised version of your manuscript entitled "*Robustness of a parsimonious subsurface drainage model at the French national scale*".

You have addressed many of the issues raised by the reviewers. The manuscript seems considerably improved. Nevertheless, it was often not easy to identify the actual changes in the text because no track-change-mode document was provided. Please do so in the next round. In addition, there are still aspects that require substantial improvements. Some of them got only evident with the revised version. I list these issues below.

Language: Despite having the text read by a native speaker (according to your letter), there are still wording issues throughout the manuscript. I list some examples below to indicate the kind of problems that exist. Please pay due attention to this aspect:

- L. 29: What the meaning of *exacerbating any infiltration concern*?
- L. 31 - 33: The expression *stone aggregates* sounds strange. Actually, the sentence is not properly citing the literature. Actually the wording is literally copied from the reference [1] but it refers NOT to mole drains but to ditches separating the experimental plots. The article itself deals with mole drains and gravel mole drains. Accordingly, the sentence on L. 32 is factually wrong. Please correct.
- L. 52: *... thus complicating their parameterization on a large database.*. Weird sentence, which is not clear. Please rephrase.
- L. 275 - 275: This information is repeated for the third time. Please skip and avoid unnecessary repetitions throughout the manuscript.
- L. 301 - 309: This information has already be provided in the previous paragraph. Skip.
- L. 353 - 361: Not the Q_{05} are close to zero, but the respective biases! Please adapt the wording for all quantiles!
- L. 430: This sentence is not correct. I assume it should be: *Not all regions .... are well represented.*
- L. 452: *respected*: this is not the adequate expression here.
- L. 459: *respects*: Rephrase: simulates well the temporal ...
- L. 467: Reword. There are not just two extreme discharges.

- L. 490: Why should *model calibration (be) only relevant if the calibrated K and μ are probable according to the case study soil type*? The word *relevant* seems inappropriate here.
- L. 516: *relevant*: Not an appropriate term: it is also relevant if the performance is poor at a given site. Relevant has the meaning of be connected with the matter at hand (see e.g., `https://www.wordreference.com/definition/relevant`).
- L. 547: What is the robustness approach? Strange wording, please rephrase.
- L. 551: What is a conventional period?
- L. 566: *Adopting assumptions on model performance would then offer a relevant alternative*: What does this mean?
- L. 568: Reword: ... *that also limit the study.* (I assume that the assumptions were not made in order to serve as limitations!).
- L. 571: Re-word: *in every other context.* Your current version tells that nowhere else the assumption were true while you want to say that not everywhere it was true.
- L. 612: *matricial compartment*: replace by soil matrix.

Calibration:
- L. 65 - 72: What are the arguments for selecting KGE in your case?
- L. 227 - 231: Where can one see the priors? Provide these data.
- L. 441 - 442: Is it actually necessary to use these priors? Wouldn't the calibration process find the appropriate values anyway? Please test this aspect if haven't done already.
- L. 578: What about the priors?
- L. 586 - 587: Where can one see these data?
- L. 587 - 592: Are the calibrated parameters indeed constraint by the priors?

Further comments:
- L. 43: This re-design of drainage system is not further elaborated on in the manuscript. Can you comment on that aspect in the discussion when talking about the benefits of the model?
- L. 53: What is *semi-conceptual*? Please be consistent in your terminology. Often you describe the model as conceptual (e.g., L. 258; which seems clear to me).
- L. 54: You should make it explicitly clear that the model has six parameters. In your application described here, you simply kept two parameters fixed at values obtained from prior knowledge. This is important for applications elsewhere (see L. 173 - 188, also L. 439).
- L. 90: Provide examples of what you do not consider. Discuss what the implications could be.

- L. 95 - 96: But this would also be possible with other classifications?
- L. 108 - 109: *where drained soils are predominantly silty-clayey and composed of fine sediment with heavy clay*: this is confusing. Is the texture silty-clayey or heavy clay? It cannot be both at the same time, can it?
- L: 119: *drainage modalities*: what does this mean?
- L. 204 - 208: In your response you mention that the model could also deal with leaky subsoils. Please be specific about that and briefly describe how the model could represent this situation.
- L. 218 - 219: *In this context, the model calibration allows estimating parameters based on a comparison between observations and model simulations.* This is trivial and can be skipped.
- L. 322: Is this a general finding or site-specific? If the peaks are underestimated but the overall flux is ok, it implies that the recession curve overestimates water flow. Is this correct and a general observation or is it site-specific?
- Fig. 8. I assume you have two data points per site from the split test. Please indicate (by specific symbols) which data points belong to the same site.
- L. 437: Why is a simple model better for generalizing? A physics-based model should be better suited because it can accommodate more different situations while a conceptual model imposes more implicit model constraints. This view is supported by your description of the conceptual limitations of SIDRA-RU. On L. 521 - 522 you explicitly state that the model was primarily designed for silty soils and the results clearly demonstrate the limitations that go with that. You illustrate this also very clearly in Fig. 12. This all seems to contradict you claim that a simpler conceptual model is better suited for general applications.
- L. 451: *more local studies*: provide references.
- L. 471 - 472: *This rationale might partially explain the observed delay.*: You can simply test this explanation instead of being speculative.
- L. 500: What is the metric for consistency? There are quite some deviations.
- Fig. 11: If I understand correctly, you have a single parameter value for each site from the reference data base and the calibration (actually, from the calibration you should have two from the split test). Accordingly, it would be more informative to compare the site-specific values.
- L. 519: N. Jarvis commented on the good performance of MACRO on such soils. This has to be mentioned and referenced to avoid a wrong impression.

- L. 544: Why should that be the case? As you describe, SIDRA-RU is not that general because it performs not very well on heavy clay soils (due to conceptual constraints, see also L 521 - 522, Fig. 12). Why should MACRO perform poorly on non-heavy clay soils? The preferential flow part of MACRO is not a static feature but depends on site-specific parameters.
- L. 550 - 551: Why should the calibration period a dry period? Where can one see this?
- L. 552 - 554: This is a hypothesis that can be tested on your data. Please do so.
- L. 558 - 561: Above you argue that crops are not relevant. This is now confusing.
- L. 572 - 573: Is this already implemented or does it need a change in the model code?
- L. 604 - 605: *Thus, a good model can be used as a decision-making tool, for example to restrict pollutants' application during flow period for the case of pesticides*: It is a typical claim to state that a model is required to support decision making. But is it actually true? Practitioners generally know very well flow periods and if they were to restrict e.g., pesticide or fertilizer applications during such periods, models weren't necessary, I'd argue. In such a situation it most probably not the lack of knowledge regarding the flow regime that limits such restrictions but the crop-specific timing and needs for nutrients or crop protection. Please be more precise in describing what can be achieved for practice for which there is hardly an alternative for models.
- L. 608 - 609: How do transfer function solve this problem? You need to come up with an adequate transfer function model. A key aspect is that water flow is not identical to solute transport. This implies that conceptually, one has to add components such as to account for the transport aspect. Important in this context: fast transport is of high relevance for sorbing (and degrading) compounds such as P or pesticides. Hence, even if surface runoff may be irrelevant for simulating the water fluxes and the water balance, it may be essential to account for runoff (and preferential flow to tile drains) in the model concept. Often this implies that one has to introduce a shallow topsoil layer to account for the crucial processes (sorption, degradation, mobilisation) controlling the fate of agrochemicals. Refer to relevant conceptual models in the literature to provide some more depth to the discussion.
- L. 618: The term *exhaustive* seems inadequate given the model limitations on e.g., clay soil that are mentioned above.

Please address these issues in an adequate form.

Sincerely

Dr. Christian Stamm Editor HESS

References

[1] Tuohy, P., Humphreys, J., Holden, N. M., & Fenton, O. (2015). Mole drain performance in a clay loam soil in Ireland. Acta Agriculturae Scandinavica, Section B — Soil & Plant Science, 65(sup1), 2-13. doi:10.1080/09064710.2014.970664