Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-635-AC3, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Predicting discharge capacity of vegetated compound channels: uncertainty and identifiability of 1D process-based models" *by* Adam Kiczko et al.

Adam Kiczko et al.

adam_kiczko@sggw.pl

Received and published: 26 March 2020

1 Major comments

First, we would like to express our gratitude for the reviewer's efforts and constructive remarks. We are grateful for the reviewer's view on the relevance and need of the conducted validation of the vegetation models at such smaller scale channels. The reviewer raises several important issues that we would like to address in our response with more detailed comment. We identified the following crucial points from the general and specific comments:

1. The drawbacks of the uncertainty analysis and the interpretation of obtained confidence intervals.
2. Definition of the identifiability, "parameter identifiability" versus "model identifiability".
3. A priori parameter distributions, ranges should be informative or uninformative?
4. Can we still consider a model as physic-based, if all essential parameters are being identified with very general assumptions on their variability ranges?
5. Insufficient evidence for findings presented in conclusions.
6. Insufficient literature review and positioning the study in the present state-of-the-art.
7. Presentation of the results for higher discharges
8. Data publishing.

At the end of our response, we also provided more brief answers on remaining reviewers specific comments.

Ad. 1. The drawbacks of the uncertainty analysis and the interpretation of obtained confidence intervals.

As the reviewer indicated, we used a term of the informal uncertainty estimation in the respect of applied GLUE methodology. We agree, that this might suggest that results of uncertainty are not reproducible, making comparison of models unreliable. We were aware of this problem and although GLUE method is in its general form informal and based on the modeler subjective assumptions on the model uncertainty, we took advantage of the approach presented by Romanowicz et al. (1996) and Romaowicz and

Beven (2002), equivalent to Bayes identification with a simple error model of a normally distributed white noise. The observation equation is then:

$\mathsf{H}(Q,\theta) = \hat{H} + \epsilon$

where ϵ stands for the noise, $H(Q,\theta)$ modeled water levels for given parameters θ and $\hat{H}$ observations. As steady state models are considered, the use of the white noise model is justified, as autocorrelation is not present. The noise is 0-mean with unknown variation. In common GLUE approaches, the variation is chosen subjectively by a choice of likelihood function and so-called "behavioral" parameters. In our study, we formalized this step using constraint on number of observations falling within confidence intervals (Eq. 5). Along with eq. 3 this allows to identify κ factor, which multiplied by σ^2 (variation of model residual for scaling) provides a variation of the noise. All together, assessing the widths of the a prior parameter distribution by "trial and error", to ensure that they are wide enough to explain model uncertainty and the Eq. 5-kind constraint, makes the approach similar to the adaptive Monte Carlo sampling demonstrated i.e. by Blasone et al. (2008). We have not applied automation in determination of a priori parameter ranges, although the concept is the same.

To reduce the effect of the a priori distributions (note, always present with finite bands) we used a relatively wide a priori parameter ranges. The goal was to ensure that a high probability region is enclosed within Monte Carlo sample. This was tested by setting such ranges, although keeping their physical interpretation, that widths of confidence intervals were sensitive for κ, above values found by minimization task given by eqs. 3-5. So, it was always possible finding wider confidence intervals and the a posteriori 95% quantiles of computed water levels were noticeably narrower than the spread of MC sample. This ensures that within region of interest (explaining the model uncertainty) the solution becomes insensitive to the spread of a priori distribution.

With a given error model, as the reviewer mentioned in his comments, applied analysis accounts only for the total uncertainty. It is not possible distinguishing other sources

like measurement error. We think however, that providing additional term e.g. for measurement errors would make the comparison of discharge models difficult, as we can expect different estimations of measurement uncertainty for different methods but for the same data sets. For sure this leads to the overestimation of parameter variability ranges, but it applies to all methods in similar way.

We hope above explanation will satisfy the reviewer, as our approach to the uncertainty estimation is based on formal assumptions and allows for comparison of different models. We thank the reviewer for pointing out that such description is missing in the manuscript and should be given in the revised version.

Ad 2. Definition of the identifiability, "parameter identifiability" versus "model identifiability"

We thank the reviewer for raising the issue that the "parameter identifiability" was in some parts of the text erroneously used when "model identifiability" should have been used. The reviewer suggested, that the term identifiability should be clearly defined. In the previous version of the manuscript we gave a remark on our understanding of this term (line 350, sec 3.2.: "The model identifiability is understood here as the ability to determine the parameter a posteriori distribution that explains the model uncertainty in relation to observations. This is satisfied by meeting the constraint given in Equation 5"), although we agree that this definition should be developed. Applying the probabilistic identification problem we consider that the model is identifiable, if it is possible to find a posteriori parameter distribution that explains the total model uncertainty in the respect of observations. The criterion for identifiability is the constraint given by Eq. 5 – 95% confidence intervals should enclose not less than 95% of observations.

The other issue, that reviewer pointed out, is that if we consider model identifiability or parameter identifiability. So, the question is about the aim of the analysis: obtaining a model with a good prediction skills or estimate parameter values that agree with measured values. Our concept for the article was however different, although the reviewers' comments have helped us to realize that we did not present our concept well. With the inverse identification problem, the goal is always to identify the model and it was the same in our study. Having identified the model, in the second stage, it is possible to analyze, if obtained parameters agree with their true values. Our results (for Pasche and GTLM) indicate, that for many parameters (but not for all) the median of the solution noticeably differs from measured values. The obvious problem is of course, if a model that well explains the rating curve, but its parameters might be different from their real physical values, should be used for water level predictions? At this point, probably not emphasized in the text, instead of advocating for our approach, we would like to present a discussion on that issues, as in our opinion the model identifiability is not the written assumption in many studies. With our article we would like to provide cons but also pros for the identification of models with a strong physical interpretation. Because, although parameters might be different, which sparks an impression of black-box modeling (more comments on that issue in the section on black-box modeling), differences are usually interpretable. The shift in a given parameter is compensated by others, i.e. the large stem diameter comes along with too large spacing of plants for the Pasche method. In our opinion, the ability for such interpretation might be considered as an advantage of the more physics-based models over simpler ones, as if modeler is aware of parameter interactions and can decide, if e.g. given before discrepancies in vegetation characteristics are important in analyzed case. Moreover, having easily interpretable parameters values, in contrary to e.g. Manning coefficient, it might be possible to recognize their unrealistic values, resulting from other model errors like improper representation of a geometry: tree trunks with diameters over several meters are much more evident than Manning roughness coefficients ranging for the floodplain at 0.1-0.2 $m^{-1/3}s$. However, we do not to show that it is possible to identify model parameters but show what is the effect of model identification, also in terms of estimated parameters.

If the manuscript is considered for the revision, we would like to improve the text by providing a clear definition of the identifiability and explain that our aim is the model identification, the output of which are also parameter values. We would also provide a

discussion on the use of models, the parameters of which might be different from the real ones.

Ad. 3. A priori parameter distributions, ranges should be informative or uninformative?

We have not introduced terms of informative and informative a prior parameter distribution. Our idea was to formulate identification problem the same way that it is done in most practical case studies of flood flow modeling, where usually there is very limited data on vegetation properties in a river channel. So our a prior distributions are uninformative and provide a wide region, where the solution for a general case, different channels, might be found. Moreover, using informative parameter bands would introduce a subjectivity to the study and it would be impossible to compare different methods. I.e. how to apply similar constraints on stem diameters in Pasche method and Manning roughness coefficients in DCM?

The reviewer's idea to take advantage of known parameter values is however attractive. It might be interesting to investigate how parameter identifiability and uncertainty estimates are affected when a prior parameter variation is reduced with additional knowledge. We think however, that this could be the scope of another study.

We would like to address the reviewer's remark on the type of the a priori distribution by directly defining that we use uninformative a prior distribution and explain more clearly our concept at this point.

Ad. 4. Can we still consider a model as physic based, if all essential parameters are being identified with very general assumptions on their variability ranges?

The reviewer raises an important issue: is it still a process based model, and not a black-box, if most of its physical variables are identified through an inverse problem (calibration)? The process-based methods are indeed functions with large number of parameters and when their physical interpretation is neglected, they indeed might take a form of formulas rather explaining the data, than providing an insight into the process itself. This would make the task similar to the problem of estimating the rating curve, as in studies mentioned by the reviewer. To maintain the physical interpretation of models it is necessary to ensure that parameter values are restricted by physical constraints. In our study we used non-informative a prior parameter distribution, although within physically possible ranges of a given parameter. We analyzed, if these physical constraints allow eliminating an inappropriate method and in the case of the two-layer and Pasche approaches we succeeded: GTLM had very poor prediction skills when applied to rigid-unsubmerged vegetation, while Pasche, Mertens methods were unidentifiable for flexible submerged vegetation cases.

However, physical constraints on parameters values might be insufficient, as it is possible to identify values much different from real, measured ones. So, i.e. does the Pasche model maintain its physical consistency, when used with much larger spacing and larger stem diameters, than those measured from flume experiments? As we already mentioned in the comment on the identifiability, these are issues, which we wanted to demonstrate with the study (parameter interpretation section), rather than answer directly. Please note, that possible conclusions apply as well to Pasche or GTLM, like to the Manning formula. The difference is that in the first case the interpretation is obvious, whiles large values of Manning coefficients are common in practice. So maybe this is an advantage of process based approaches, where parameters are easily interpretable? They can be identified and modeler can validate if obtained values follows his/her exceptions.

On the other hand, the use of the inverse problem to determine parameters values, is not uncommon approach even for "very" process-based approaches, like i.e. Shino-Knight model, where it is necessary to identify turbulence parameters (like turbulent viscosity and secondary/advection flow term) and it is known that outcomes are affected by high equifinality (Knight et al. 2007). In our opinion, such problem will apply to all process-based methods, when applied in general practice task. Therefore,

we would like to avoid impression that we suggest using the physic-based models as black-box ones, but our aim was to investigate the usually unstated assumption on the identifiability of such models.

The discussion of these issues was missing in previous version of the manuscript, but we would like to include it in the revision.

Ad. 5. Insufficient evidence for findings presented in conclusions.

The reviewer indicated that conclusion are not well-supported with the manuscript text. In the case of conclusion points (claims) 1 and 2, the issue will be clarified by specifying that the primary goal is the model identification (please see the note on identifiability), as the reviewer suggested.

We agree with the reviewers remarks on conclusion points 3-4. With our approach we are unable to separate sources of uncertainty, including the equifinality. The equifinality is present, and it can be seen in parameter distributions, where wide regions in parameter space can be considered highly probable (high values of likelihood measure). In the case of uncertainty and the number of observation it is not a matter of equifinality but the ill-posed inverse problem (insufficient number of observations). The conclusion points 3-4 should be revised as follows:

3. The uncertainty related to the **ill-posed inverse problem** is noticeable only when a small number of observations is used in parameter identification.

4. The parameters obtained through the identification differ from their measured physical values, which results from the parameter equifinality. The equifinality does not, however, affect the uncertainty of a model.

The way, how this effect can be traced should be explained. It can be done by interpreting obtained average confidence widths as a function of the number of observation points used for identification (Appendix A, Figures A1-A25). Wide confidence intervals, and their spread for the small observation number n=1 (following the reviewers remarks the symbol should be changed) should be attributed to the ill-posed inverse problem. Additional data points allow to narrow confidence intervals and reduce their spread, among different observation sets. In the manuscript text we provided for each method the number of observations, at which the width of confidence interval stabilizes. Additional observations affect the solutions but not widths of confidence intervals. Note, that our analysis at this point is rather descriptive. We understand that additional observations (of water levels) do not noticeably affect the estimates of uncertainty bands, so the effect of ill-posed inverse problem becomes negligible.

The reviewer remarks will be addressed in the text by clarifying our understanding of the model identification and in the case of "parameter equifinality" correcting conclusions as presented and providing a discussion of the ill-posed inverse problem.

Ad. 6. Insufficient literature review and positioning the study in the present state-of-the-art.

We agree with the reviewer, that the literature review can be improved and references to studies such as these provided by the reviewer as an example, should be included in the manuscript. However, in our opinion the research problem is different than in given examples. In our case we address the parameter identification problem with variables having specific physical meaning (i.e. Manning roughness coefficient or vegetation height, density). The point was not only to obtain an efficient estimator of the water level-flow dependency, but investigate, if it can be obtained using physically interpretable models and then if interpretability is maintained in terms of parameters determined through the inverse problem. Examples of such approaches might be found in hydrology, as we hoped we presented in the manuscript, but not in hydraulics. The only exception might be a study of Berends (2019), analyzing inverse identification of Delft3D model, which we found after submitting our manuscript. This is the most similar study, although focused on the single but distributed model, and we will cite that publication.

If applicable, in the revised version of the manuscript the description of how we consider our study novel should be improved. Also the review of the state-of-arts should be improved and positioning of our study in the respect of rating-curve fitting should be discussed.

Ad. 7. Presentation of the results for higher discharges

The reviewer indicates that for the scope of research, it would be interesting to present the results, when the model is verified for higher discharges. We think it is a very good idea and if applicable we would like to add an additional section to the results, where i.e. discharge curves obtained with models identified using lower flows are analyzed in respect of the accuracy in predicting higher flows.

Ad. 8 Data publishing.

The reviewer suggests, that the revised version of the manuscript should include data we used in computations. We agree at this point, and we are ready to include our data sets.

2 Responses to specific comments

1. P1-L10: 'quasi-Bayes'. The only time this term is mentioned is in the abstract. Perhaps a definition could be included in the manuscript.

   Response: The term "quasi" should be removed, as it applies to common form of the GLUE.

2. P3-L58 the authors use 'process-based' models and 'physically-based' models seemingly interchangeably. I recommend choosing either'process-based'or'physics-based' (not 'physically-based', which is admittedly used throughout literature)

Response: Difficult choice, we think that process-based suits better and this term will be used in the revised manuscript.

3. P3-L58 I do not share the authors' broad assertion that physics-based approaches are unpopular in practical applications.

   Response: It should be specified, that this statement applies to flood hazard assessments, where other methods than the DCM are very rare.

4. P3-L62-64 The authors skip over many other possibilities by jumping from the detailed parameters of the Vastila and Jarvela models to the Manning coefficient. A (physics based) model with fewer parameters would be an option. Numerous studies can be found in literature where two-dimensional models use spatially distributed information on vegetation, often based on remote sensing techniques.

   Response: The reviewer is right, we should mention methods that allow to estimate the vegetative roughness based on e.g. treating vegetation as rigid cylinders (with drag coefficient and frontal area/density derived from remote sensing). Although we would like to indicate that these approaches present rather the state-of-art than the present practice in several countries, including Finland and Poland where the authors come from, where the lumped approach to resistance parameters prevails. We will highlight in the introduction that we are focusing on 1D models for practical applications.

5. P3-L72 This is a very valid point, and in my view the most important objective of this study

   Response: Thank you for acknowledging this point which we agree is the main objective of the work. If the article is considered for revision, we would like to emphasize this issue throughout the manuscript and present this earlier in the Introduction.

6. P3-L74 "Any method can ... a parameter calibration". I reject this statement. Do the authors like cake only if all its ingredients can be individually tasted?

   Response: We had such impression analyzing the problem based on our experience on flood hazard assessments. We understand that we should not generalize, so we consider softing this statement, by i.e.: "Usually a method is widely applied in practice if all its parameters can be identified as the solution to the inverse problem – a parameter calibration."

7. P3-L77 'predictive uncertainty' is a technical term used differently by different authors. The authors should define their use of the term.

   Response: We agree and will define "the predictive uncertainty as the estimated total uncertainty of the modeled variable".

8. P3-L79 'As one of the first works, this paper' What paper do the authors refer to? (If they mean their manuscript, see my general remark on novelty).

   Response: We meant our manuscript, in this point we still consider our approach novel, with comments given in general responses.

   "As one of the first works, this paper evaluates the uncertainty of chosen 1D state-of-art methods for predicting the influence of complex vegetation on the discharge capacity (understood as the dependency between water level and discharge) in compound channels where vegetative flow resistance dominates".

9. P3-L82 "most of the previous studies": the literature review by the authors cites dated literature. To give confidence in this statement, the authors might provide a review of more recent literature.

   Response: The literature review should be updated with e.g. studies on discharge curve fitting. Please see also our general comment on the literature review and position of our study.

10. P3-L89-90 I don't think this is a valid contrast. Morphologic and hydraulic modelling are very different challenges and many existing vegetation models are not suited for morphological modelling.

    Response: We disagree at this point. Warmink et al. (2013) analyzed the hydraulic resistance, parametrized with bed morphological features. This makes the study similar to the parametrization of resistance using vegetation characteristics. To avoid future confusions, we would like to develop this comment on Warmink et al. (2013) article.

11. P4-L95 The authors compare 'explicit' and 'implicit' uncertainty analysis. I'm not sure this terminology is commonplace, and in any case requires explanation.

    Response: The reviewer is right, we will develop the explanations concerning these terms.

12. P4-L116-L120 I'm unsure why the authors use the terms 'minor values' (text), 'minor parameters' (figure), or 'conservative approach'. To me this sounds derogative, as if to discredit this approach in favour of their proposed alternative, although I readily assume the authors do not intend this. For instance, I do not see why surface roughness is in any way 'minor'. In fact, if this parameter is used for calibration it is very likely the most sensitive parameter, so by all accounts should be labelled 'major'. Nor do I see why the first approach should be labelled 'conservative' (what is conserved? Do the authors mean 'traditional'?). Second (and this point was raised by another referee as well), the 'conservative approach' is surely preferable - if reasonable estimates of the uncalibrated parameters are available - over treating vegetation models as black boxes.

    Response: We thank the reviewer for raising these points and want to highlight that it was not our intention to discredit approaches where measured values of vegetation are available, but to analyze if the the use of physic-based approaches provide reliable estimates for small channels without a prior knowledge on vege-

tation properties. We will clarify this scope of the manuscript more clearly when describing the goal of the paper at the end of Introduction section. Following the reviewer's suggestion, we will use the term “traditional approach” when referring to the way how it is usually done (Fig. 1a), as the opposition to the new (proposed) approach (Fig. 1b). Further, we will highlight that “traditional” approach is preferable when vegetation data is available. In the case black-box issue, please refer to our broad answer on that issue.

    In the case of “minor values”, we meant those parameters of process-based methods that have less importance on the conveyance estimation of compound channels compared to vegetation characteristics; herein, please note that (P4, L100) states “This work focuses on one-dimensional methods for compound channels with a significant share of the flow resistance generated by vegetation.”. We will express this issue more clearly in the revised manuscript. We agree that the term “minor” is not precise, as these parameters are nevertheless significant. So we will replace the word “minor” by “parameters other than vegetation properties”. We agree that the term “minor” may have sounded derogative as we did not express ourselves clearly, although this was not our intention.

13. P5 L134-135: For its merits in popularizing uncertainty analysis, the GLUE method is (in)famous for the liberal use of the likelihood measure, which does not agree with Bayes' theorem. The authors choose to use a so-called 'informal' likelihood measure with a scaling factor that controls the uncertainty. The authors then force the model uncertainty to include at least the right amount of data points through equation 5. This approach, inspired by Bayes theorem, has the known disadvantage that predictive and model uncertainty are lumped (the authors approximate the total uncertainty), that parameter uncertainty tends to be overestimated, and that the choice for a likelihood measure is arbitrary (i.e. not following from the error model, as is the case in a proper Bayesian approach). A defense by the authors on their choice for an informal over a formal approach

would be appreciated for readers unfamiliar with this distinction. Also, given that the authors use an informal method, I'm interested whether a behavourial threshold was used(the authors mention the need for this, but not whether it was used. I assume none was used given the scaling factore.)

Response: The response on the uncertainty analysis was given before in general comments. In the case of behavioral threshold, the reviewer is right, we have not used it, because of the applied likelihood function. Such comment will be added to the manuscript after defining the likelihood function.

14. P7-L194 Please elaborate which resistance (all,only the bed,only the imaginary wall?)

    Response: The information in the manuscript is not precise: bed, imaginary wall and also vegetation stems.

15. P11-L297 Is the Ritobacken Brook free [FB02?]owing? Can uniform [FB02?]ow be reasonably assumed? P13 In general, I suggest adding the first section of chapter 3 to the method section, as new methodology is introduced here.

    Response: The Ritobacken Brook is free flowing in that there are no hydraulic structures affecting the flow at the investigated discharges and water levels. No changes were observed in the rating curve at the downstream culvert (downstream of which there is a forested section with steep slope). At very high discharges, the culvert at the downstream end of the study reach will start to dam the flow, but no such high flows were recorded in the present data. The flow at Ritobacken is gradually varied (we will replace the term "non-uniform flow" by "gradually varied"), and therefore we used the energy slope instead of bed slope(L312). In the case of the first section of the chapter 3, we agree on moving lines 315-324 to methodology, of course keeping the tables in results.

16. P13-L315 By'trial and error' choosing the sampling size of Monte Carlo, do the

authors mean increasing the sample until convergence is observed? What convergence criteria is used? Which sampling method is used?

Response: We thank the reviewer for pointing out that this information is missing in the article. For the convergence we used two criteria: mean of computed water levels and good fit (in deterministic manner) of the rating curves determined for each combination of observations. For the sampling we used Latin Hypercube method. The information will be included in the text.

17. P13-L315 "In a similar way"; are a priori distributions chosen by trial and error? In principle a priori distributions are either informative, based on prior knowledge, or uninformative. Here an uninformative uniform prior is chosen, but I have to learn this from the captions in table 1. It would be helpful if this is explicitly added to the text as well. I would also appreciate a brief exposition of the choice for a noninformative prior. Given the models are physics-based, and the authors have a pretty good estimation of their likely values, it seems more logical to use informed priors.

    Response: Please, see general comments: Ad. 3.

18. P17-L365 Here the authors define "model identifiability" as (I paraphrase): "it is identifiable if it is fittable". The authors then admit that their approach would allow even poor models to [fit well, while "the only limitation could be the physical meaning of the parameters". It is unclear whether the authors did indeed let themselves be restrained by the physical meaning of those parameters. One may remark here that minor changes to their chosen approach (i.e. a formal Bayesian approach and informed priors) would be expected to alleviate some of these problems.

    Response: This issue was addressed in general comments.

19. At this point it is also good to remark on a different, perhaps more fundamental point. The authors go into depth into 'model identifiability' but the reader was led

to believe that 'parameter identifiability' was the objective of the study. Yet most results and almost all figures focus on the question 'will the model fit' - which is not a very interesting point to stress given the objective of the study. The only figures that support 'parameter identifiability' are 14 and 15, but those are currently insufficient to support the claims made in the conclusions; it would be helpful to plot the a priori cdfs as well, so as to see how they were constrained a posteriori.

Response: This issue of model/parameter identiafiablity was addressed in general comments. The prior distributions are uniform, so the way how they are constrained with likelihood function can be presented using dot-plots (parameter value vs likelihood measure) or like in the case of Fig. 14-15. We prefer the second option, as the figure is much more readable. As the reviewer found Fig 14-15 interesting, we think that such plots can be presented for other methods as well.

20. P31-L495 - The 'trial and error' a priori distribution estimations bothers me a bit when claiming objectivity. Given the limited number of observations, the a priori distribution is expected to affect the output.

    Response: Explanation given in a broad comment on the uncertainty estimation.

21. *P31-L500 'It was possible to identify...one (DCM)'. This should not be surprising. It is in general easier to fit a model with more parameters than one with fewer. However, the more impressive claim would be that the parameters are identifiable as well. I refer to the work of Werner et al. (2005, doi:10.1016/j.jhydrol.2005.03.012), to illustrate that challenge.*

    Response: This links with the problem statement: model identification vs parameter identification. As we mentioned in general comments, we would like to present the model identification and then analyze it outcomes, also in terms of parameter identification. The reviewer's idea, given with the Werner et al. (2005) is very interesting, as it could be analyzed in terms of uncertainty estimation with

increasing knowledge on parameter variability. However, we are afraid, that it would be hard to address this issue in a single article.

22. P32-L528 'Thus, our result... resistance dominated'. I do not see how this follows from your results nor from the previous sentence.

    Response: We agree that this information was not directly presented. The Ritobacken case was also monitored with the absence of vegetation as cited in L28: "particularly in small to medium-sized channels where up to 90 per cent of the [FB02?]ow resistance can be caused by plants (e.g. Västilä et al., 2016)." We will add this result from the field to the appropriate place in the revised manuscript. The reviewer is right, we have not performed studies without vegetation so the claim is unsupported with the results. The results show, that the choice was important for analyzed cases and in this way, the sentence should be rephrased.

23. P33 Conclusions. The authors conclude the article with 8 claims.

    Claim 1. The authors claim it is possible to identify the parameters of physics-based models, even if those models have many parameters. This is an unlikely claim (given the number of data points and the number of model parameters), but may follow from a confused definition of 'parameter identifiability' versus 'model identifiability'. If the parameters are 'identifiable', I would expect narrow a-posteriori distributions compared to the a-priori distributions. Figures 14-15 do not show a-priori distributions. Although it is difficult to judge whether the parameter distributions are meaningfully narrow, it seems only C is well defined. If so, it would be interesting to re[FB02?]ect on figure 1. The second claim, like the first, seems to only apply to model identifiability, not parameter identifiability. The authors might spend some words on how they perceive the model to be used-does it matter if the models are physics-based? Or any(data-based) model that fits the rating curve applicable?

    Claim 3: Perhaps the authors could explain how uncertainty relating to parameter equifinality can be distinguished from other uncertainty.

Claim 4, first sentence: Would not a better explanation be that the model is insufficient in some way, and that model parameters differ to account for this? Second sentence: I don't understand this in relation to the third claim.

Response: we have addressed this remarks in general comments (Ad. 5).

Literature

- Berends, K. D. (2019) Human intervention in rivers, quantifying the uncertainty of hydraulic model predictions, PhD thesis, University of Twente, Netherlands.
- Blasone, R., Vrugt, J., Madsen, H., Rosbjerg, D., Robinson, B., and Zyvoloski, G. (2008). "Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling." Adv. Water Resour., 31(4), 630–648.
- Knight, D. W., Omran, M., and Tang, X.: Modeling depth-averaged velocity and boundary shear in trapezoidal channels with secondary flows, Journal of Hydraulic Engineering, 133, 39–47, 2007.
- Romanowicz RJ, Beven KJ, Tawn J. Bayesian calibration of flood inundation models. In: Anderson MG, Walling DE, editors. Floodplain processes. Chichester: Wiley; 1996. p. 333–60.
- Romanowicz, R. J., & Beven, K. J. (2006). Comments on generalised likelihood uncertainty estimation. Reliability Engineering & System Safety, 91(10-11), 1315-1321.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-635, 2020.