HESS 2017-314

Data-based mechanistic model of catchment phosphorus load improves predictions of storm transfers and annual loads in surface waters

Editor comments to the response to the reviews

8.9.2017

Dear Dr. M. C. Ockenden

Thank you for the responses to the comments provided by the reviewers.

In most cases, I consider the suggested modifications of the manuscript and/or the replies as satisfactory. However, regarding the issue of model uncertainty the revision needs to go beyond what you suggested. Reviewer 1 and 2 raised some important questions that need more in-depth responses and analysis in the manuscript.

In the Introduction you mention explicitly that even though the model structure is largely data driven, structural model errors will remain (p. 5, L. 10). You also discuss limitations of model structures in section 3.5 (p. 12, L: 6 – 15). Despite acknowledging model structure as a source of uncertainty, you almost completely ignore this aspect when responding to comments by the reviewers addressing this aspect. Reviewer 1 for example comments "…*However, when the uncertainty bands do not encompass the measurements, it's not really better situation than having a large parameter uncertainty. The model is either missing an important process or measurement uncertainty is not accounted for.".*

You reply that *"…*The model fits the data well, so the covariance matrix is small (in L2 sense), and the uncertainty of the model is limited to its parametric uncertainty."

First, you do not comment at all at the correct observation by the reviewer that for a substantial fraction of time the uncertainty bands do not enclose the observations (see Fig. 4 – 6). Second, you explicitly claim that parameter uncertainty dominates the model predictive uncertainty ("…The model fits the data well, so the covariance matrix is small (in L2 sense), and the uncertainty of the model is limited to its parametric uncertainty."). Implicitly you also suggest that by showing the measurement uncertainty the deviations between observations and simulations can be explained. However, you do not provide any arguments why model structure was not a relevant source of uncertainty. In addition, one has to consider that you actually skipped the error term in Eq. 2. What are the implications for uncertainty quantification?

In addition to this lack of actual evidence for your statements, it also contradicts other statement in the text and the data you present. On p. 5, L. 30 – 31 you write: “Prediction bounds for the model can be calculated by adding the residual uncertainty and the parameter uncertainty.” Comparing the deviations between simulations and observations in Fig. 4 for example with the magnitude of the indicated parameter uncertainty makes it hard to reconcile them with your statements above. This holds true even when considering the aspect of measurement uncertainty. You briefly touch upon that issue in the text and suggest to add this information to the (figures in) the revised version. While this a very valuable suggestion, a closer visual inspection of the data casts doubts whether measurement uncertainty can fully explain the mismatch between observations and model simulations.

In summary, there are two aspects where you do not provide a satisfactory answer to important questions of the reviewers regarding model uncertainty. First, throughout the text you deal with the different sources of uncertainty in an inconsistent manner. In some paragraphs, model structure is considered, in others not, at some places residual errors are explicitly mentioned, later on they are completely ignored. Second, there are obvious discrepancies between the model predictions and simulations that ask for i) a proper presentation of the relevant data (e.g. the residuals) and ii) a coherent discussion of possible sources of uncertainty. Please note, the issue is not that the model results were not of sufficient quality. It is only about the presentation and the discussion of the actual uncertainty.

To address these issues, I ask you to provide the following data and information:

- Refer to the different sources of uncertainty in the Introduction (and Method section) and explain how you have quantified and /or accounted for them in the context of this paper. Please also refer to input uncertainty.
- Description how the measurement uncertainties were quantified. This has to include an explanation how you dealt with the mix of systematic and random measurement errors and how you accounted for the temporal auto-correlation of discharge errors.
- Provide the uncertainty bands for the observations (as suggested in your response).
- Provide an analysis of the residuals (for discharge and TP) by showing time series of the residuals and the residuals as a function of the observed values. This can go to the SI but the reviewers and readers need this data to judge the model quality.
- Explain how measurement uncertainty can explain systematic deviations between simulations and observations given the fact that you fit the model to these observations.

- When making statement about the origin of uncertainty, please refer to data such that readers can follow your arguments by referring to actual data (e.g. the residual analysis).
-

Based on the visual inspection of e.g., Fig. 4 or 6, I have the impression that peaks (of a certain magnitude) are systematically underestimated (I am happy to see if your data proofs me wrong). Again, should this be the case, I don't see any problem in that for the manuscript, but think that this was an important aspect to know. First, from a theoretical point of view one would ask why a DBM could not capture that? Would even longer time series be needed? Such aspects would nicely fit into Chap. 3.5. But also from in practical terms it might useful because it could pinpoint hydrological conditions during which pronounced TP loads/concentrations occur (even if they are not fully reproduced by the model).

There were two additional comment that were only partially answered. Reviewer 2 asked "*What is exactly the meaning and the implications of not using a noise model in Eq 2? This should be explained in more detail. Any inference algorithm (in this case probably the RIV(C)BJ), needs to make assumptions about the errors to estimate parameters. Does not using a noise model mean that you assume the errors to be uncorrelated? Or is the error model inferred by the algorithm itself? The assumptions made in the inference process should be clearly stated and checked."*

I think your response does not really provide the answer to what the reviewer wanted to know. As I interpret the comment it was not just about why you skipped the error term in Eq. 2. The point is that in order to estimate your model parameters the simulated and the observed values are compared. If you use all the hourly values to minimise your objective function you implicitly assume that all these data points are independent and uncorrelated. In reality however, the observed and simulated values are highly auto-correlated (for some typical time scale). If you model prediction is overpredicting discharge at time x, it is highly probable that the same holds true for the next time step x+1 (as a consequence of your high temporal resolution). For these reasons, people have tried to base their parameter estimates on innovation for example (e.g., Yang, Reichert et al. 2007). I think the reviewer wanted a clarification of that aspect.

Reviewer 2 also stated: "Eq 4 leads to significant violation of the mass balance w.r.t. water if Q(t-1) is larger than 1 (this depends strongly on the units of Q) and beta is larger than 0. This should be clearly stated, and then briefly mentioned why this is not a problem in this case (if it is not)."

I have to admit that I could simply not follow your argument and see how the water balance problem is avoided. Can you rephrase?

In summary, I ask you to revise the manuscript according your responses that you have provided to each review and to additionally address the issues that I have listed above.

Sincerely

Christian Stamm

Editor HESS

References:

Yang, J., P. Reichert, K. C. Abbaspour and H. Yang (2007). "Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference." Journal of Hydrology **340**(3-4): 167-182.