Draft: Response to the editor (deadline Oct.18, 2013)
Edited by G.Blöschl, H.Flühler, H.Holländer, H. Bormann
In italics (font 10): Points raised by the reviewers
In regular fonts (plain text font 11): Response by author and co-authors

Interactive comment on “Impact of modellers’ decisions on hydrological a priori predictions” by H. M. Holländer et al.

Anonymous Referee #1

Received and published: 9 August 2013

We would like to thank the reviewer for his/her insightful comments. Review comments under inverted commas, response in plain text.

“This manuscript is the follow-up to a previous study and further investigates the ‘human’ effect in modeling with the many decisions one has to take when applying a model to a new study site. This is a very important topic that certainly needs more attention and I absolutely want to see this work being published. Having said this, I also have to admit that I was a bit disappointed when reading the manuscript and have several concerns with the manuscript in its current form:

1) Study design: When different people, with different backgrounds, apply different models with different types of information, it is difficult, if not impossible, to really assign the reasons for any differences to one source. From a study design view point, it would have been better to use one model applied by different persons. I can see the argument that everyone applied a model he/she was familiar with, but obviously this makes it difficult to distinguish between model and modeler effects. Can we really rule out that the differences are mainly due to different model structures? On page 8896, 20 the authors state that the differences could mainly be attribute to different modeler decisions. While I would also feel so, I do not see this being really demonstrated in this study. It would be valuable if at least one model would be applied by different modelers, or one modeler would apply different models.”

We agree that there are a number of alternatives to the study design and this is perhaps obvious in hindsight. At the time the study was started, the focus was indeed on model parameters and model structure choice which both turned out to be dominated by the modellers’ experience. Hence the focus of this paper changes in the course of the study. We still believe that very useful, if preliminary, conclusions can be drawn from the study.

The point the reviewer is making on whether one can really rule out that the differences are mainly due to different model structures is an important one. According to Holländer et al. (2009) the process equations are quite similar between most of the models, therefore dimensionality and parameterisation (both mainly modeller’s decisions) remain as relevant differences. The models used in the comparison are all quite flexible because of their high dimensionality. Virtually any hydrological response can therefore be simulated with appropriately selected parameters. We therefore think that, at the end, the main differences are due to model parameter choice (in the context of each model) and the parameters have been subjectively chosen by the modellers. This is why we believe that the differences are indeed due

to the modellers rather than the structure. It would be straightforward to fit all models to the observed hydrographs with close to perfect match but we think this is not needed as the high dimensionality implies that this is possible.

To address this review comment we will revise the manuscript to better bring out the role of model structure versus modeller. We will also stress that these are preliminary results and a more refined study design would be interesting to be pursued in the future.

"2) Also related to the study design, I see several reasons why models where changed in round 2 and 3. These are as mentioned the field visit and the additional data and, mentioned less clearly, of course also the interaction among the modelers (as far as I understand the field visit where done by all scientist together, i.e. there certainly was informal information exchange to some degree). It would be quite interesting to distinguish between these points, but with the given study design, I am not sure, to which degree this might be possible."

Yes, this is correct. We will address this point in the revised manuscript. Of course, distinguishing between these points in a quantitative way will be difficult, but we think this should be discussed in the paper.

"3) Are the modelers being studied or are they studying? The text is written in a way as if the modelers were studied, at the same time, they seem to be co-authors (i.e., studying). In hydrology we are not used to study human beings and to reflect on the researchers own role in the process. This study certainly would benefit from looking at other sciences, where the situation of being involved in the process one wants to study, is more common. The text is also formulated in a way that tries to not reveal who the respective modelers were, but of course it is easy from the author list and the models to guess, who was who, especially n the case where the modeler is a 'she'.

I feel it would be better to include the names in a table, rather than have the readers guessing. Also it has to clarified more clearly what the contribution of the different coauthors was. Did the 'modelers' just do their respective modeling or where they also involved in analyzing results and in the evaluations, discussions & writing? When it comes to evaluating what a modeler was thinking, for instance, it is crucial to know whether he/she had any influence on what was written."

Again, this is a good point. From a formal (social science) perspective, the modellers, as an objective of study, should not be part of doing the assessment. A more comprehensive study design should take this aspect into account. In practice, this was perhaps not as critical as it seems, as the observations on the modellers were mainly made by the two senior authors involved who did not do any modelling themselves. Perhaps this should be more clearly stated in the manuscript to clarify the individual roles of the authors. Regarding disclosing the identity of the modellers, the idea of not doing this was to avoid a 'beauty contest'. Often, modellers feel attached to the models they have built or are using (the Pygmalion effect) and highlighting the names associated with each model would perhaps bring this point to the fore. Instead of a beauty contest, the main focus of the paper is on the underlying mechanisms of why model parameters are chosen in a particular way and their effect on the modelling results. We also

carefully avoided any modifications and additional model runs after completing the individual prediction steps. Illustrating the progress in prediction quality was the ultimate goal of this study. The finding that progress was at best moderate was a surprise to the group and so was the modellers's role in the process, so we hope these findings are of interest to the readership.

"4) The study is limited by the number of models/modelers. Conclusions on the effect of prior knowledge etc. are therefore difficult. While I can see that it is difficult to get a much larger number of modelers, this is an issue that needs to be clarified more clearly."

We fully agree. The number of participating modellers was below what one would require for the results to be statistically significant. Hence, statistically, the graphs regarding prior experience or amount of information do not formally support the hypothesis that system knowledge is a good basis for choosing the right model structure and parameter estimation method. Rather, these results SUGGEST that this is the case, paving the way for more detailed studies in the future. We are convinced that this study will trigger more research along these lines. In the revised manuscript we will better bring out the preliminary character of the study.

"5) Paper structure: to be honest, I got lost when reading the manuscript. The structure needs to be improved, currently there is a mixture of methods, results and discussion at many places. Also I find the manuscript in several parts too anecdotic, I found it very hard to follow the text and the see how did what and why. At some parts it is also unclear, what is a description of the results of the previous study and what are new results. With many outstanding scientists on the co-author list, I am a bit surprised that the presentation quality is not as good as it could be."

We appreciate the concern that the paper is unusually long and the structure has potential for improvement. We will move part of the complete decision process into an additional appendix to make it more digestible by better separating the documentation from the discussion. In addition, we will streamline the text by eliminating less important details.

In conclusion, we are grateful for the thoughtful remarks of this reviewer and we are convinced that addressing them will strengthen the paper.