Reply to the reviewer comments RC1: 'Review of the manuscript by Schürz et al.' by Björn Guse

Summary

In this manuscript by Schürz et al., a detailed analysis of the impact of scenario simulations on hydrological variables is presented. In a sensitivity analysis, the sensitivities of five groups are separated. These groups are three types of scenarios (land use, point source and climate) and two model-specific groups (model set-up, model parameterisation). In the analysis, the impact of the input variables and of the uncertainty of the selection of scenarios or model characteristics is presented.

Overall, I like the manuscript. However, I see still potential for improvements to increase the understanding of the manuscript.

We would like to thank Björn Guse for his very constructive review and the valuable comments to improve the quality of the manuscript. We appreciate the positive feedback on the manuscript. In the following, we addressed each comment individually. The comments made by Björn Guse are printed in *serif, italic font*. Our replies to the comments are written in black, non serif font and our suggestions to revise the manuscript according to a comment are highlighted with the colors blue for insertions and red for deletions.

Major comments

From my perspective, the readability of the manuscript can be increased by a clear separation in the two major points of the article. First, the impact of the input variables is analysed to show which input variables are more relevant for discharge and nitrate. Second, it is analysed how the selection of a scenario or model characteristic controls the target variables (uncertainty analysis). I think that the article would be easier to understand if these two aspects are clearly separated. This comment is mainly related to abstract, introduction and discussion. In contrast, these two aspects are already clearly separated in the conclusion.

We strived for a consistent structure in the manuscript by presenting the two separate blocks: 1) the sensitivity analysis of the model inputs and the model setup and parametrization; 2) and the uncertainty analysis together with the visual analysis.
With your comment in mind, we see the equivocality in the outline of the manuscript. In the current form the manuscript outline can be interpreted as if the performed sensitivity analysis was the actual focus of this work and the visual analysis of the uncertainties are a mere by-product (e.g. p.1L15ff in Abstract, p.5L7-16 in the Introduction).
We agree, that emphasizing the value of the visual analysis and treating it as an individual part of this work increases structure and consistency of the manuscript (e.g. being consistent to section 2.6 p.12L18ff where we already clearly separate the two goals of the performed analyses).

We suggest the make following changes in the updated version of the manuscript:
Exemplary suggested changes in the abstract, p.1L15ff:

The analysis of the 7000 generated model combinations of both case studies had two main goals; i) to identify the dominant controls on the simulation of discharge and $NO_3^- - N$ loads in the two case studies and ii) to assess how the considered inputs control the simulation of discharge and $NO_3^- - N$ loads. ~~In both case studies we employed global sensitivity analysis (GSA) t~~ To ~~identify~~ assess the impact of the input scenarios ~~scenario inputs~~, the model setup and the parametrization on the simulation of discharge and $NO_3^- - N$ loads we employed methods of global sensitivity analysis (GSA). The uncertainties in the simulation of discharge and $NO_3^- - N$ loads that resulted from the 7000 SWAT model combinations were evaluated visually. We present approaches for the visualization of the simulation uncertainties that proved to be a powerful diagnostic tool in this study to assess how the analyzed inputs affected the simulations. ~~We accompanied the GSA with a visual analysis of the simulation outputs and their associated uncertainties that resulted from the simulations of the 7000 SWAT model combinations. We present visualizations of the results of the GSA and the simulation uncertainty bands that proved to be powerful diagnostic tools in this study.~~

Following the suggested changes for the abstract we suggest to apply the same ideas for revising the introduction. Here we will focus on the sections p.3L21-29 and p.4L7-16 in particular.
Concerning the discussion we would prefer to keep the approach we currently follow in the manuscript, in which we discussed our findings and related them to other literature. Clearly separating the findings concerning the sensitivity analysis and the uncertainty analysis is difficult to facilitate for the larger part of the discussion. A separation might lead to a lot of repetition in the discussion.

P.11, L: 1: Is it correct that you have identified 43 and 52 behavioural parameter sets out of 100.000 model simulations? If it is true than the number of behavioural parameter sets is rather low. How is than the impact on the sensitivity analysis meaning that most of the parameter sets are unbehavioural?

The decision whether a parameter set is considered to be behavioral is highly subjective, as the objective criteria that are applied to evaluate the model simulations and the thresholds for these objective criteria that define a simulation as “good” or “insufficient” are individual decisions. The decisions made in the presented work are listed on p.10L6-10.
We agree, that a different definition of a behavioral parameter set would affect the influence that the model parametrization has on the analyzed model outputs. The effect of the assumptions made in such an impact assessment were therefore discussed in section 4.2. Thus, the low number of behavioral parameter sets does not per se affect the sensitivity of the model outputs on the model parametrization.
The low number of behavioral parameter sets in this specific case results from the study design. A model parameter set was considered as a behavioral parameter set, if the simulations performed with **all** model setups using that parameter set fulfilled the applied objective criteria (which means a model parameter set implemented in the Raab case study had to meet the thresholds for the objective criteria in all six implemented model setups).

This design decision may influence the resulting sensitivities of the model outputs, as the impact of the model setup and the model parametrization combined can be greater than the effect resulting from the illustrated study design (All individual model setups together with their model parametrizations may result in behavioral model/parameter setups, that are not considered here). The implemented setup however, isolates the effect of the model setups and the effect of the model parametrizations for model setups that were calibrated for a reference period and are applied for future changing conditions. In the context of an environmental impact study, to assess their individual effects is highly relevant in our opinion.
As a result of your comment (and similar comments in other reviews) we see a requirement to clarify the the evaluation of the parameter sets on p.10L6ff.

Figure 4: It is very hard to understand this figure. In my understanding the results from Fig. 3 are shown again and in addition to that the variations evoked by changes in land use or point emissions. Is it maybe better to present this as relative change to the lines in Figure 3? Or only as line and not as coloured area?

The Figures 4 to 8 follow the same pattern in the analysis that they illustrate. The Figures 4 to 8 indeed present the results shown in Figure 2 in a modified way. While Figure 3 shows the uncertainty bands of the 7000 simulations performed in each case study implementing the different combinations of input scenarios and model setups, the following figures separate the resulting uncertainty bands with respect to the discrete realizations of the individual model inputs and model setups. During the compilation of the manuscript we tested different ways to communicate the information we wanted to convey (e.g. plot all 7000 simulations as lines, analyze relative changes, etc.). We concluded however that the selected visualizations were the most suitable ones that supported our findings best. Thus, we prefer to remain with the presented figures. We see however the need to clarify the explanations regarding the Figures 3 to 8.

Discussion: One idea is to add a table or figure as an overview in the discussion to show which of the five criteria has a dominant impact on discharge and nitrate and which criteria are uncertain. I think that the article would benefit from a clear and easy understandable presentation at the end as a kind of take-home-message. I have in mind a figure which summarize all results in relative values. To understand the overall idea of summary figures see for example Figure 9 in Herman et al., 2013.

We highly appreciate this comment and thank you for the link to the publication by Herman et al. (2013). Herman et al. (2013) used the summary figure as a very effective tool to summarize their findings. We were discussing how to implement this tool to summarize our findings in the manuscript. So far however, we were not able to come up with a good solution that would add value to the manuscript and facilitate interpretation for the reader. Thus, unless we come up with an appropriate illustration during the revision of the manuscript, we prefer to not add an additional figure.

Specific comments

P.1, L. 8: I suggest to modify to: "In impacts studies in two Austrian catchments, …

We prefer your suggestion over the phrase in the manuscript. The text will be changed accordingly.

P. 1, L.13: I suggest to write "for each catchments" instead of "for both catchments".

Together with other changes the section p.L11-14 will be updated as follows:
We developed scenarios of future changes for land use, point source emissions, and climate. The developed input ~~and implemented the~~ scenarios were implemented in ~~realizations in the different~~ SWAT model setups with different spatial aggregations and employing different model parametrizations that were able to adequately reproduce historical observations of discharge and $NO_3^- - N$ loads. ~~, which resulted in~~ In total 7000 combinations of scenarios and model setups were used to ~~both catchments. With all model combinations we~~ simulate~~d~~ daily discharge and $NO_3^- - N$ loads at the catchment outlets of each catchment.

P.2, L: 17: I suggest to add: "using a set of different climate input data for hydrological models" at the end of this sentence (or a similar statement).

In this particular section we wanted to keep the statement more general (the statement is also true for land use change, or any other change process expressed with discrete scenarios). Thus, we prefer to keep the general phrase with the following example of climate change scenarios, as written.

P.2, L.27: The discussion on equifinality is not well motivated. I miss a sentence to relate both paragraphs.

We suggest the following modification of this section in the updated version of the manuscript:
To simulate the development of hydrological variables under changing conditions, the developed scenarios are implemented in hydrological models that are calibrated for historic conditions. Yet, often different model setups and different sets of parameters in a model can perform equally well to reproduce historical observations of the variables of interest. Equifinality is a well-known issue in hydrologic modeling that has been extensively addressed in the literature...

P.3, L. 5: I suggest to add a sentence at the beginning of the paragraph similar to "Sensitivity analysis can be used to derive the impact of different input variables on hydrological target variables" to make clear why you have selected this method.

The sentence will be added accordingly.

P. 6, L. 11: fertilizer

This will be corrected accordingly.

P. 8, L. 5: Please avoid one-sentence-paragraphs

The sentence will be added to the previous paragraph.

P. 9, L: 14: I suggest to write: "applied a GSA on discharge and nitrate...".

To consider your suggestion and the suggestions made by other reviewers on this sentence it will be changed as follows:
In a ~~pre-analysis step~~ parameter screening, we applied a GSA to the simulations of discharge and $NO_3^- - N$ at the catchment outlets using all SWAT model setups ~~individually~~ to identify the relevant model parameters.

Table 3: Is the sensitivity related to discharge or nitrate or both?

We appreciate this comment and think that this is valuable information for the reader. Thus, we suggest to modify Table 3 and differentiate between parameters that were influential for discharge related processes and $NO_3^- - N$ related processes.

P.15, L. 14: You may add that this result could be expected since the model structure is known to be of higher importance for low flows since high flows are strongly driven by the precipitation (observations).

We addressed this issue already in the discussion to some extent. We suggest however to stress this issue more and to clarify this point in the discussion. In contrast to your statement, the study design (that tried to assess the individual effects of the model setup and the parametrization) clearly show that the model setup has a stronger influence on large and medium discharges, whereas the model parametrization greatly affects the low flow.

P. 15, L. 31-34: For me, it seems to be that in Fig. 3, spring is the dominant season in the upper left subplot.

We agree with your observation that spring is little more dominant in Schwechat. Thus, we will mention this fact as well.

Figure 4: The legend needs to be explained in the figure caption.

We agree that the used abbreviations are not self explaining in this figure. Hence, we will add an explanation of the abbreviations in the figure caption.

P. 19, L. 5: Could you add in which subplot you can see this drastic change?

Actually, that finding is supported by all subplots. The anomalies in precipitation affect the simulated long term monthly averages of discharge and $NO_3^- - N$ loads as well as the different segments of the illustrated flow duration curves of both catchments.

P. 19, L. 11: Have you an explanation for this?

We think that the precipitation anomalies explain these findings, to a large extent. Increases in mean annual precipitation increase the discharge and $NO_3^- - N$ loads, while a simulated reduction of mean annual precipitation in the future results in a reduction of discharge and

$NO_3^- - N$ loads. This explanation is, in our opinion, provided in the text of the manuscript. We suggest however to revise this section and try to specify the statement more precisely.

P. 25, L. 3: I suggest to add "The selection of" before "climate scenarios".

This will be changed accordingly.

P. 26, section 4.2: You may add a statement similar to "This analysis shows again that a clear description of the selected scenarios is mandatory for impact studies."

We appreciate this comment and such a statement will be added to the text.

References

Herman, J.D.; Kollat, J.B.; Reed, P.M.; Wagener, T. (2013): From maps to movies: high resolution time-varying sensitivity analysis for spatially distributed watershed models, Hydrol. Earth Syst. Sci., 17, 5109-5125.