

## ***Interactive comment on “Comparing SWAT with SWAT-MODFLOW hydrological simulations when assessing the impacts of groundwater abstractions for irrigation and drinking water” by Wei Liu et al.***

**Wei Liu et al.**

liuwei.alan@gmail.com

Received and published: 17 October 2019

We thank the Anonymous Referee 2 for his/her thoughtful comments and efforts towards improving our manuscript.

### **General comments:**

Liu et al. present a comparison of SWAT with SWAT-MODFLOW for the Uggerby catchment, Denmark. The study is well written with an excellent level of detail provided in the method. I have no working knowledge of the models applied in this study, so my

C1

comments are high level, relating to the statistical interpretation of the results and their significance.

My main concern is that the conclusion that SWAT-MODFLOW is superior does not seem to be justified by model performances achieved in validation. Looking at Table 5, we see that the additional model flexibility offered by MODFLOW and MODFLOW with PEST significantly increase calibration performance without improving validation performance significantly. This suggests that these models are simply over-fitting. I don't think this paper requires any additional experiments to be run, but I do suggest that the results need to be interpreted accurately. Unless the authors offer a convincing reason for their current interpretation of the results (based on their validation), I would suggest that the whole discussion and conclusions need to be rewritten to be more reflective of an honest appraisal of the model performances. The other general issue is that the paper is very long. I think it can be shortened significantly without losing the key messages.

**Response:** We thank the reviewer for insightful comments. We have shortened the paper as much as we can (also in response to reviewer 1 comments). Regarding the model performance description, actually we evaluated the performance of the two models not only based on the statistic metrics values but also through visualization of hydrograph and according to the evaluation criteria recommended by Moriasi et al. (2015) which has widely been used for evaluating the performance of hydrological models. Although during the validation period, the percent bias (PBIAS) of SWAT-MODFLOW was slightly worse than SWAT; their values still fall in the same class according to the evaluation criteria recommended by Moriasi et al. (2015). However, according to the R2 and Nash–Sutcliffe efficiency (NSE) values, the evaluation criteria and visualization of the hydrograph (Fig. 6), we deem that the performance of SWAT-MODFLOW was overall better than SWAT. (Lines 445-453).

### **Specific comments:**

C2

1) Abstract - very long; considering shortening.

**Response:** Good suggestion. We have shortened it as much as we can, without compromising key outcomes.

2) Line 80 - no model considers the "entire" complexity. Please revise.

**Response:** Good point. We have rephrased the sentence in line 80 as follows:

Line 77-78: "In contrast, numerical, process-based models take into account more about the complexity and heterogeneity of river-aquifer systems."

3) Line 160 - please report

**Response:** Good suggestion. However, unfortunately, we have not found literature reporting the proportion of irrigation areas in Denmark. Instead, we found out that the annual irrigation amount during 1989-2007 was 175-259 million m<sup>3</sup> (Thorling et al., 2019).

Modifications: We have now added this information into the manuscript as follows:

Line 160-162: "In Denmark, approximately 800 million m<sup>3</sup> of water are abstracted annually and used for irrigation (175-259 million m<sup>3</sup> during 1989-2017) or drinking water (GEUS, 2009; Thorling et al., 2019), making the country highly dependent on groundwater."

4) Line 173 - The hypothesis that the "benefits of applying SWAT-MODFLOW outweigh the costs" is one that can be tested objectively (and is not answered in your results). I suggest reframing the study so that the aim is to explore the effects of introducing MODFLOW and MODFLOW pest into SWAT simulations in this particular catchment.

**Response:** The reviewer has raised a valid point. We have rephrased the hypothesis as follows: Line 28-31: We hypothesize that an integrated surface-subsurface model SWAT-MODFLOW performs better relative to a lumped semi-distributed catchment model SWAT when assessing the impacts of groundwater abstractions (for either

C3

irrigation or drinking water) on streamflow patterns.

Line 174-176: We hypothesize that the SWAT-MODFLOW performs better relative to SWAT when assessing the impacts of groundwater abstractions (for either irrigation or drinking water) on streamflow patterns.

5) Line 325 - Is the water stress threshold taken as a single value for the whole catchment? If so, what are the limitations of this assumption? Would the threshold vary according to crop type/soil type?

**Response:** No. Actually three and four values for water stress threshold were taken for the whole catchment after calibration, as shown in Table 3 and Table 4.

6) Page 12 - the reader does not need to know the names of your python scripts.

**Response:** Good point. We have now deleted the names of our python scripts.

7) Figure 5 - not particularly helpful. I think this can be omitted.

**Response:** Good suggestion. We have now deleted figure 5.

8) Line 478 (and throughout the results section)... lots of results reported in vague terms ("little higher", "much lower" ... etc). Please report

**Response:** Good suggestions! As suggested, we have replaced the vague terms by reporting

9) Line 552 - it does not reflect a shortcoming of the SWAT groundwater module if the improvements are simply overfitting.

**Response:** We have now made it clear that we refer to a shortcoming in the conceptual model of SWAT, as it ignores the variability in distributed parameters such as hydraulic conductivity and storage coefficients, lumps spatial detail within the groundwater domain of a subbasin, and contributes to the stream network as baseflow based on a linear reservoir approximation.

C4

Modifications: Line 545-548: "This reflects the shortcoming of the concept for SWAT groundwater module, which ignores the variability in distributed parameters such as hydraulic conductivity and storage coefficients, represents groundwater by a lumped module in individual subbasins, and contributes to the stream network as baseflow based on a linear reservoir approximation."

**10)** Line 560 - this conclusion is not warranted if the model has been overfitted, as is suggested by results reported in Table 5.

**Response:** We do not think the model was overfitted, but we made some changes in Line 560 to make the conclusion more convincing as follows:

Line 551-553: "The availability of spatial-temporal patterns of the groundwater head in SWAT-MODFLOW could significantly benefit groundwater resources management and provide the spatial explicitly water resources dynamics within a catchment."

**11)** Figure 6 - remove background shading.

**Response:** The reviewer has given us a good suggestion, but unfortunately, we are not able to remove the background shading in figure 6. There are some zone boundary lines inside the original shape file of layers (layer 1 and layer 3) showing the zones of different hydrological properties. If we remove the shading, those zone boundary lines will show up and then mix up with the contours, making the figure more blurry. Instead, the readers can zoom in the figure to read the figure more clearly.

**12)** Figure 10 - These are not promising results. Seasonal well drawdowns in the simulations do not occur in the observations. Why should this not be reported as evidence of the poor performance of SWAT-MODFLOW?

**Response:** Good point. We have rephrased the texts in lines 469-471 as follows:

Lines 469-471: There was generally a good agreement between the groundwater head level and dynamics simulated by SWAT-MODFLOW and that recorded at the two observation wells within the catchment, though the seasonal well drawdowns in Well A

C5

did not always occur in the observations (Fig. 10).

**13)** Lastly, the arbitrary labels attached to NSE scores ("satisfactory" etc) are inappropriate. Report the numbers, show the data, and let the reader decide what is satisfactory.

**Response:** Actually, the labels attached to the statistical metrics are not arbitrary. We evaluated them according to the criteria recommended by (Moriassi et al., 2015), which has been widely used to evaluate the performance of hydrological models.

## References

Moriassi, D. N., Gitau, M. W., Pai, N., and Daggupati, P.: Hydrologic and water quality models: Performance measures and evaluation criteria, *Transactions of the ASABE*, 58, 1763-1785, 2015.

Thorling, L., Albers, C., Ditlefsen, C., Ernstsen, V., Hansen, B., Johnsen, A., and Trolborg, L.: Grundvandsovervågning, Status og udvikling 1989–2017, GEUS, De Nationale Geologiske Undersøgelser for Danmark og Grønland Energi-, Forsynings- og Klimaministeriet, Copenhagen, 140, 2019.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2019-232>, 2019.