First of all, we would like to express our sincere appreciation to Reviewer 1 for the positive overall assessment of our manuscript and for the critical questions that have led to an improved and clearer manuscript.

Please find below the detailed answers to the question:

- **33 - 58**: This seems to be a technical summary about SWAT. It is unclear why such detailed information of SWAT is provided at this position. What is the demand for further research regarding the general hydrological community?

The gwflow module is designed to replace the groundwater component's conceptual representation in SWAT+, and we focused on improving this new module. To provide the reader with the reason behind, we went from the basics so that every detail can be understood. It is also with an intention to take this article as a reference on how the coupling of SWAT progressed in time and its limitations and advancements. As you have stated in your subsequent question (question number 2), our focus is to deal with SWAT.

The major outcome of our research entails an important implication to the hydrological community, where not accounting for groundwater-soil interaction can be an essential factor for unsatisfactory hydrological simulations in hydrological models. Hence, other hydrological models should consider such interactions to pave the road for effective ground-surface water coupling.

- **59**: There are several efforts to improve groundwater representation in hydrological models. Since the authors are focused on SWAT, it would be scientifically sound to mention and discuss these efforts at least for current and previous SWAT versions and to work out the necessity to consider other approaches (e.g. the current study).
The current and previous SWAT versions have a similar approach when it comes to groundwater hydrology representation. They have lumped conceptual representation, which has only been changed after coupling techniques came into practice. However, since previous coupling techniques have significant limitations (as stated in lines 82-83), gwflow module is developed. Previous coupling techniques are discussed between lines 66 to 80.

- This seems to be the motivation of this study?

Yes, it is one of the main motivations behind this research where holistic approaches are needed to understand the overall geohydrology of a given catchment. In addition, a comparison of the standalone SWAT+ and SWAT+gwflow is made to show the limitation of the conceptual representation of groundwater hydrology by the standalone model. Moreover, soil-groundwater interactions, which are neglected in the previous version of gwflow are accounted which is the major contribution of this article.

- 93: Limitations of SWAT were also shown in Luo et al. (2012), Pfannerstill et al. (2013), Nguyen and Dietrich (2018), Shao et al. (2019). Additionally, these studies provided solutions to improve the representation of groundwater within SWAT.

Thank you for suggesting more references which we have not come across. The references you have suggested have not been implemented in the SWAT+ model so far, but in the future, we will take them as an input to modify the module or the standalone model even further. We would like to affirm that the main intention of this research is to modify the gwflow module and make a comparison with the standalone SWAT+ model. Hence, you may please take it as an extension of the gwflow module developed by Ryan T.Bailey (Bailey et.al 2020).

- 97: Please provide a clear structure: what are the gaps? which problems need to be solved? how are the problems solved? The introduction ends up with a rather technical information, which is not appropriate to guide readers through this study. It would be more helpful to provide the aim of this study at the end of the introduction.

We have modified the end of the introduction with the suggestions you have made. You may find them in the edited version of the manuscript from line 91 to 97.

- 151: The introduction ends up mentioning the integration of pumping. This part is not mentioned here?

That is correct, but the issue here is that we have not included pumping in the modelling scheme of the Dijle catchment due to data limitations, and we intend to include it in another paper to show the impact of pumping in the general hydrology. Nevertheless, we have tried it with a dummy number, and it worked well. The pumping term is included in equation 3 \((Q_{\text{pump}})\), where the groundwater change in volume with time is solved by including this term. Recently, Bailey et.al 2021 published a paper by including pumping in
a catchment located in the USA to investigate the impact of tile drain in the hydrology of
the catchment.

- **215: Please provide references.**

We have added the reference accordingly.

- **274, L.285 and L.286: It is unclear which parameters and which ranges were used for
the sensitivity analysis. This is very crucial to reproduce the results of this study and to
check if this method is appropriate.**

We have included the parameters used (in Table format) for calibration and the sensitivity
analysis results (in plot format) in the supplementary material. In addition, the calibrated
parameter sets is also included in the supplementary material.

Note that we had a master's student who tried to calibrate the SWAT+ model for the Dijle
catchment for almost five months, and it was impossible to achieve good agreement
between modelled and observed streamflow values, which is the primary motivation for
this research. The best parameter set we achieved was based on several simulations made
on that thesis (the student is also a co-author of this article - Lise Leda Piepers). The main
conclusion of the thesis was that the standalone model could not represent the hydrology
effectively due to limitations in the groundwater hydrology representation.

Moreover, to ensure transparency in our work, we are willing to share the model we
developed for the Dijle catchment. It is also a catchment we used to prepare the tutorials
for SWAT+gwflow model (previous version of the model) and recently we replaced it with
the case study we have from Tuscany, Italy – Ombrone catchment. Link for tutorial:
https://swat.tamu.edu/software/plus/gwflow/

- **302**

- Please provide a table with calibrated parameter values for both model setups.

- You may check the answer we provided above (#8).

- Please provide appropriate figures for simulated/observed discharge. Figure 5b does not
show any meaningful information since data is very compressed. A flow duration curve
would be helpful to show the deviation between simulated and observed discharge
magnitude.
Thank you for your suggestion, and we have included a flow duration curve to show the deviation. As the flow duration curve has indicated, the differences are significant during peak flows, but those events are few, with flow rate below 30 m$^3$/s in most events. The major player in this catchment is the groundwater hydrology, which is also indicated in the baseflow filtering discussed in question number 11 below.

355: The evaluation based on monthly discharge should play a minor role. Of course, groundwater processes are more time-delayed and could be checked at a monthly time scale. However, other crucial hydrological processes happen immediately on a precipitation event. For a consistent model behavior, all relevant processes need to be represented and of course they need to be evaluated. Consequently, it is necessary to focus to an appropriate extent on daily discharge.

It is correct that daily time step can be appropriate, and due to that, our assessment does not only rely on using monthly but also daily model outputs to investigate further. For instance, you may check Table 1, where objective functions are estimated for both daily and monthly time steps.

343: Are these model results realistic? Are there any options to evaluate this?

Yes they are realistic, and we can evaluate this by assessing the baseflow component of the streamflow using baseflow filtering techniques (e.g. WETSPRO - Willems 2009). This will allow us to understand how much of the total flow is accounted for by baseflow, interflow, and overland flow. We can infer from the filtering that baseflow accounts for the majority of the streamflow (around 80% of the total streamflow), indicating the prevalence of subsurface than surface processes. You may please check the WETSPRO result below.
Figure: The baseflow filtering parameters (top) where $w$ is the difference between 1 and the percentage of baseflow. The bottom plot shows the filtered baseflow for the catchment outlet.

For a detailed description of the baseflow filtering methodology, you may refer to the link: https://bwk.kuleuven.be/hydr/pwtools.htm#Wetspro. We have also included in the .zip file the filtering excel workbook.

- 356: Please provide discharge for the validation periods with an additional figure.

Thank you for your suggestion, we have included plots for monthly and flow duration curve for daily timestep in the supplementary material for the catchment outlet and additional gauging station inside the catchment.

- 392: This aspect is fully new and was not mentioned before.

One limitation of the SWAT hydrological model is the nonrepresentation of wetlands in the modelling scheme, which is clearly stated in several literature (Golden et.al 2014, Wellen et.al 2015). Especially during drought periods, wetlands rely on their interaction and the water they will receive from the groundwater and/or surface water bodies. On top of this recommendation, here in Belgium, the amount of drainage water from agricultural lands is not well known, which is a concern of the government, and we were asked to assess the amount and the impact it will have on water availability. This is the reason why we put this recommendation on the inclusion of tile drains in hydrological simulations to capture the actual hydrological behavior of a given catchment. We believe that including wetlands and tile drains (ditches) in the modelling scheme of the Dijle catchment can improve the hydrological simulation even further (on top of the improvement we made). To summarize, the points we raised in line 392 and further are recommendations.

Please also note the supplement to this comment: