First of all, we would like to express our sincere appreciation to Reviewer 2 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:

- *The authors declare that the primary novel side of this research article is representations of groundwater-soil interactions in the introduction. However, I did not see any details on the parameterization of groundwater-soil interactions in the model description. Moreover, detailed analysis on the influence of this process on the modeling result.*

In the original water balance equation of the gwflow module (Bailey et al., 2020), the interaction between SWAT+ HRU soil profiles and gwflow grid cells occurs only through the passing of deep percolation from the soil profile to groundwater, i.e. the water table elevation, and corresponding groundwater storage, simulated for each grid cell does not interact with nor affect the hydrological processes of the soil profile. This could lead to situations where the gwflow module simulates a water table within the soil profile, with all inherent groundwater calculations such as lateral flow, while at the same time, the HRU in the same vicinity as the grid cell is simulating percolation and soil lateral flow. This leads to duplicate processes, but by different equations. To remedy this flaw, in this study, we modify the gwflow module to allow the transfer of groundwater to the soil profile for grid cells where the water table rises into the soil profile. This is particularly important for groundwater-dominated catchments (i.e., high baseflow fractions) where the water table is shallow. For each cell, for each time step, the simulated water table elevation (i.e., groundwater head) is compared to the elevation of the soil profile bottom of the HRU to which it is connected spatially. If the water table elevation \( w_{t\text{elev}} \) (m) is above the elevation of the soil profile bottom \( s_{\text{elev}} \) (m), then the volume of groundwater to transfer to the soil profile \( Q_{gw\rightarrow soil} \) is calculated as:

\[
Q_{gw\rightarrow soil} = (w_{t\text{elev}} - s_{\text{elev}})A_{hru}S_y
\]
where $A_{\text{HRU}}$ is the area ($m^2$) of the HRU within the spatial area of the grid cell, and $S_y$ is specific yield of the aquifer. This volume of water is added to the HRU soil layers that are at or below the elevation of the water table. Including this transfer term can yield higher rates of recharge during a model simulation due to added water to the soil profile that can subsequently move downward to the water table. Therefore, when presenting and discussing model results or estimating recharge rates for a watershed or portion of a watershed, the “net” recharge should be estimated by subtracting the transfer volume from the simulated recharge volume, i.e., true recharge = $Q_{\text{rech}} - Q_{\text{gw-->soil}}$.

As for the result before and after accounting gwsoil interaction can be found in Table 1 inside the revised manuscript (also can be found below). It can be seen that when the groundwater – soil interaction is neglected, there is a lot of saturation excess flow (164.2 mm/yr), because the groundwater is allowed to rise to the ground surface and then discharge directly to nearby streams. When gwsoil is accounted for, then the groundwater is transferred to the soil, leading to higher ET, surface runoff, and lateral flow. Also, there is still recharge (net = 837 - 733 = ~104 mm/yr). This implies the high groundwater will be routed to the streams via soil lateral flow instead of direct excess flow that probably leads to a significant amount of water reaching the river.

**Table 1.** The water balance components (in mm) with and without applying groundwater – soil interaction.

<table>
<thead>
<tr>
<th>Inputs to Watershed</th>
<th>Without gwsoil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>838.9</td>
</tr>
<tr>
<td>Boundary Inflow</td>
<td>-5.5</td>
</tr>
<tr>
<td>Output</td>
<td>Value</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Lake seepage to groundwater</td>
<td>0.4</td>
</tr>
<tr>
<td>Outputs from Watershed</td>
<td></td>
</tr>
<tr>
<td>Surface ET</td>
<td>544.4</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>66.4</td>
</tr>
<tr>
<td>Soil lateral flow</td>
<td>10.6</td>
</tr>
<tr>
<td>Stream seepage to groundwater</td>
<td>-4.5</td>
</tr>
<tr>
<td>Saturation excess flow</td>
<td>164.2</td>
</tr>
<tr>
<td>Groundwater ET</td>
<td>0.2</td>
</tr>
</tbody>
</table>
On top of this, the streamflow simulation deteriorated when gwsoil interaction is not considered (NSE of 0.6 when accounting gwsoil interaction and NSE of 0.1 when it's neglected). All changes we made (the comparison between before and after applying the gwsoil interaction) can be found inside the revised document.

- Again, the current analysis simply compared two modeling results in modeling streamflow and simply showed the modeled groundwater balance. A scientific analysis on why the SWAT+gwflow model shows improvement in modeling streamflow (e.g., you can analyze the runoff components, the difference between two modeling results in water table depth, recharge) is needed.

The standalone SWAT+ model approach regarding groundwater hydrology representation has a shortcoming. It has lumped conceptual representation, which has only been changed after coupling techniques came into practice. This lumped representation limited us from having output such as the groundwater table depth or recharge. Due to this, we are not able to compare the results from SWAT+ and SWAT+gwflow based on recharge and water table depth. However, a detailed analysis of before and after accounting gwsoil interaction is now included inside the revised document and the comparison is also discussed below (on top of our answer for your first question).
Daily hydrologic fluxes of watershed inputs and outputs (Figure 1) show seasonal patterns of ET, runoff, and soil lateral flow. The components of water yield (surface runoff, soil lateral flow, saturation excess flow) are shown in Figure 7A, with saturation excess flow being zero, due to the inclusion of the gwsoil term, and hence the groundwater head as simulated by the gwflow module not allowed to rise to the ground surface and induce saturation excess flow, but instead groundwater is incorporated into the soil profile and managed by the HRU soil processes of runoff and soil lateral flow. For this condition, therefore, soil lateral flow acts as a pathway for groundwater to reach the stream network, and hence can be considered as baseflow. This allows an accurate depiction of baseflow, as shown by the streamflow comparison in Figure 2A. The inclusion of the gwsoil mechanisms results in net recharge (Figure 2C) that, when averaged over the entire watershed, is usually positive, but can be negative as groundwater is transferred to the soil profile but no recharge occurs.

**Figure 1.** Daily hydrologic fluxes (mm/day) for watershed inputs (precipitation, boundary inflow; positive value) and watershed outputs (negative values), for the period 1993-1996.

The average annual hydrologic fluxes (mm/yr) within the watershed system are listed in Table 1 above. For the simulation period, on average ET accounts for 67% of precipitation, and surface runoff, soil lateral flow, and net recharge account for 10%, 5%, and 12%, respectively. When neglecting the gwsoil mechanism (i.e., turning off this feature in the SWAT+ code), hydrologic flux rates change dramatically (Table 1). This condition yields a very high rate of saturation excess flow (164 mm), as groundwater head is allowed to rise to the ground surface, i.e. there is no interaction with the soil profiles of the HRUs. This high rate of saturation excess flow, when combined with surface runoff (66.4 mm) and soil lateral flow (10.6 mm) and accounting for stream seepage to groundwater (4.5 mm), results in a total water yield of 237 mm, which greatly overestimates streamflow at the gaging site. Net recharge (Figure 2D) is also much higher in this condition (204 mm), as groundwater is not removed from the aquifer and placed in the soil profile. The high recharge rates result in high groundwater head, which reaches the ground surface in many places of the watershed, resulting in saturation excess flow.
In addition, the current work also include the pumping in the SWAT+gwflow model. However, how much influence does the pumping process have? The relative importance between pumping and groundwater-soil interactions? Which factors contribute to the improved streamflow modeling (e.g., the original groundwater module, the groundwater-soil interactions, and the pumping) are not analyzed.

It is correct that we have included pumping inside the gwflow module, 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 study area 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 pump), where the groundwater change in volume with time is solved by including this term. Moreover, recently, Bailey et.al 2022 published a paper by including pumping while modelling (using SWAT+gwflow) a catchment located in the USA to investigate the impact of tile drain in the hydrology of the catchment. However, without the inclusion of pumping, the comparison made before and after including gwsoil interaction in our modelling scheme showed the importance of accounting such interactions (without gwsoil interaction, the streamflow simulation resulted in unsatisfactory NSE (0.1) while the inclusion resulted in NSE of 0.6).

The authors also emphasize that, compared with the precious study that used MODFLOW and SWAT, the current SWAT+gwflow model does not require any additional hydrological model. However, as the precious work obtained “nearly similar results”, I think the current point is relatively weak and is technical. Instead, it is more scientific to analyze the influence of direct coupling of groundwater module and SWAT on hydrological processes. For example, whether the SWAT model presents improved simulation of soil moisture, evapotranspiration after coupling the groundwater module and considering the groundwater-soil water interactions.

Thank you for suggesting further research points. The point we were inferring to was that the effort and time needed to build coupled surface water model and MODFLOW model is much higher than that of SWAT+gwflow model development. This is attributed to the need for building surface water model (e.g. WetSpaSS) to get recharge that forces MODFLOW. Contrarily, the SWAT+gwflow model does not require recharge from another hydrological model. Moreover, since we don’t have the model result of MODFLOW model (our comparison was based on the result presented in their paper), we can't compare the soil moisture, ET, and groundwater-soil interactions. Furthermore, as described in lines 301 to 304, their model setup is not SWAT+MODFLOW; instead, its MODFLOW model forced with recharge from WetSpaSS model.

I can assure you that this will be an essential part of a research we are currently working on where we are trying to compare SWAT+gwflow model setup versus SWAT+MODFLOW model setup for a catchment in Ethiopia. Hence, we would like to thank you for pointing out these fundamental research approaches.

The manuscript also needs improvement. For example,
- *I did not find any introductions to the streamflow data (including its record length, location) in the “Study area and data”.*

It is inside the methodology part, which is indicated in lines 247 to 253. Thank you for the note, and we have added the information about the streamflow data in the “Study area and data” section in line 127.

- Very detailed information is given in the study catchment in section 2, including the tributaries. However, the locations of tributaries are not given in Figure 1, which makes the reader confused.

Since the plot contains already congested information, we are forced not to include the tributaries inside the plot.

- What does the different colors in Figure 2 mean?

They are different geological units (layers), however, our main interest is the unconfined part of the geological unit. Hence, we indicated the impermeable layer along with the “Brussels sand” aquifer above it. We have noted that it can be misleading, thus, we have included the explanation in the caption.

- What does the “uncalibrated zone” mean in Figure 3? Why it is not calibrated?

The uncalibrated zones are explained from lines 212 to 221. The number of zones identified from the global dataset is 24, nevertheless, if we took all zones and assigned one hydraulic conductivity and specific yield parameters, then it will be difficult to calibrate the model (over-fitting will also be an issue). Due to this, we abstain from including those minor zones during calibration.

- Whose simulation results is shown in Figure? SWAT or SWAT+gwflow? Moreover, both SWAT and SWAT+gwflow should been shown in Figure 5 to help us better understand the improvement in SWAT+gwflow.

The results shown are for SWAT+gwflow. The result from the standalone SWAT+ model gave unsatisfactory result (NSE less than 0.5), hence, we decided to show result from only the coupled model. I.e., if we plot them together since the standalone model is giving results that are worse than the mean streamflow value, it will distort the view, thus, we removed it from the plot.

- The unit of PBIAS should be given in Table 1.

Well noted, and we now have included the unit inside the table.

- I cannot see the yellow, light blue and blue regions in the right panel of Figure 7. And what does the white area in the right panel of Figure 7 mean is not declared.

The yellow, light blue and blue regions correspond to few pixels and if you zoom in, you will find them. Most values are below 70 mm (yellow). The white area stands for areas with no groundwater-soil interaction. The red parts are for low flux values but due to significance level, they were plotted as zero. Now, since this can be misleading, we have replaced it with a new plot.

- L385 “MDOFLOW” should be “MODEFLOW”

Thank you, and it has been modified accordingly.