

The reviewer comments are in **Bold black**, while the authors' responses are in **bold blue** text.

**Full Title:**

**Unravelling groundwater's role in soil-plant-atmosphere continuum: Integrated ecohydrological modelling approach using STEMMUS-SCOPE and MODFLOW 6**

**Overall evaluation:**

**This study investigates the role of groundwater in the soil-plant-atmosphere continuum by coupling the STEMMUS-SCOPE SPAC model with MODFLOW 6 to simulate key variables such as soil moisture, temperature, groundwater levels, and ecosystem processes at three locations in the Netherlands. While the research addresses an important topic in ecohydrological modeling, there are notable limitations that need addressing : 1) although the coupled model aims to improve accuracy in vegetation carbon dynamics through groundwater integration, indicators like GPP show no significant improvement compared to single-model results, raising questions about the necessity of using a coupled approach. 2) the one-dimensional nature of the model limits its ability to capture lateral runoff dynamics, particularly in agricultural landscapes where water level fluctuations are significant. 3) It is important to improve result presentation, as current figures lack clarity. Overall, while the study has potential significance, it requires substantial revisions to address these issues and better communicate findings.**

**Dear Reviewer,**

**We would like to sincerely thank you for your precious time in reviewing our manuscript and providing valuable comments that led to possible improvements in the current version. The authors have carefully considered the comments and tried our best to address every one of them. We have restructured the manuscript to clearly emphasize the scientific contributions enabled by the new modelling framework. We hope the revised version meets your expectations. Please, have a look at our responses to your comments.**

**Major concerns:**

- 1. It is widely recognized that explicitly representing groundwater mass and energy within an IEM framework improves the quantification of the SPAC. A more carefully formulated hypothesis could better inspire readers.**

**Okay, the research hypothesis (lines 189-196) is expanded to state the study's contributions.**

**Lines 189-196 in revised version:**

***"The hypothesis of this study is that explicit representation of groundwater mass and energy in an IEM framework would improve quantification of the soil–plant–atmosphere continuum by: a) improving the simulation of soil moisture and temperature profiles, b) enabling two-way feedbacks between the unsaturated and saturated zone and their further impact on water fluxes, and c) revealing the groundwater's influence on carbon fluxes and vegetation physiological processes such as root water uptake and fluorescence. Collectively, testing these hypotheses would unravel***

*the groundwater role in shaping the SPAC dynamics and deepen the understanding of ecosystem functioning.”*

2. Three field sites and corresponding observational data were used to apply and test the coupled model. However, the rationale for selecting these specific sites and the underlying motivations for conducting the modelling work there are not detailed.

Okay, we have answered this in our reply to the detailed comment number 1.

3. While the effects of groundwater are noted, the discussion regarding its influence on the soil profile, water, and carbon fluxes could be expanded. Providing specific results and illustrative figures would greatly clarify this crucial aspect, which is fundamental to readers' interpretation of the work.

We thank the Reviewer for this constructive comment. To address this, the Discussion section has been significantly expanded and reorganized to better highlight the main contributions of the study and to position the findings within the broader state of knowledge. In particular:

- We have revised the Discussion sections into 5 sub-sections:
    - Soil moisture (Section 4.1.)
    - Soil temperature (Section 4.2.), including three separate subsections
      - Groundwater temperature and soil temperature (4.2.1.)
      - Groundwater level and soil temperature (4.2.2.)
      - Groundwater and surface soil temperature and canopy temperature (4.2.3.)
    - Evapotranspiration components (Section 4.3.), including three separate subsections
      - Subsurface evaporation (4.3.1.)
      - Root soil water uptake (4.3.2)
      - Root groundwater uptake (4.3.3.)
    - Carbon fluxes and SIF (Section 4.4).
  - Within each Section, we have substantially rewritten the text to better highlight the contribution of the groundwater.
  - We also added one Figure (Figure 9) in Section 4.2.3.
  - We have also strengthened the discussion with relevant literature and, where appropriate, include comparisons with previous studies in order to place the findings within the broader ecohydrological modelling context.
4. The manuscript does not detail how soil heterogeneity in the profiles was addressed in the modelling. Clarifying how this critical factor is accounted for is essential for assessing the model's applicability and accuracy.

We apologize for the lack of clarity. The soil profile was discretized into a number of several vertical layers at each site as explained in lines 471-489. Before the start of the simulations, all layers were assigned uniform values. Then, during the calibration, the soil parameters were updated to reflect the soil heterogeneity and improve the match between observed and simulated values. To clarify:

- A new table (Table 2 in the revised version) is added that summarizes the sites' characteristics and discretization of the model layers.
- The table of the parameters (Table 4 in the revised version) is updated with the ranges of the calibrated values instead of the previous one dominant value.
- Relevant text is added in lines 517-522.

Lines 517-522 in revised version:

*“As explained in Section 2.4.1, the soil profiles at each site were discretized into multiple vertical layers (Table 2) to account for soil heterogeneity. Initially, all layers were assigned uniform values for each parameter. During the calibration procedure, these parameters were updated individually for each layer to reflect the vertical variability of the  $\theta$  and  $T_s$  measurements. The ranges of the calibrated values for each parameter across the layers are summarized in Table 4.”*

5. The study's primary focus appears ambiguous: is it to present a new modelling framework, or to elucidate the specific role of groundwater within the SPAC system? Clarifying this objective would strengthen the paper's contribution. If the former, the model structure requires a more detailed and systematic introduction. If the latter, the analysis of groundwater's effects should be substantially deepened and discussed with greater insight.

The study focused on the groundwater contribution to SPAC modelling rather than presenting a new modelling framework. To address this, the Discussion section has been significantly expanded and reorganized to better highlight the main contributions of the study and to position the findings within the broader state of knowledge, as we have explained in detail in our reply to major comment 3.

Detailed comments:

1. The choice of the three field sites (Loobos, Cabauw, Veenkampen) lacks justification. Merely noting that results differ across sites is insufficient; the significance of these differences for the study's conclusions remains unclear. Explaining why these sites were chosen and what their comparative results demonstrate would greatly enhance the impact of the work.

Okay, we have added details about the motivation behind the selection of these sites (lines 397-405).

Lines 397-405 in revised version:

*“Those three sites were selected because: a) they provide comprehensive in-situ measurements spanning over the water, energy, and carbon cycles, including meteorological forcing, different observations such as soil moisture and temperature profiles, Eddy covariance flux measurements (e.g. ET, GPP, and NEE), groundwater levels and temperature, and vegetation characteristics (e.g. SIF). Such data density is essential for rigorously testing the hypothesis of this study and evaluating the developed coupled soil–plant–atmosphere–groundwater modelling framework, and b) the sites represent two dominant Dutch landscapes (forest on sandy soil; meadow on clay/peat), allowing preliminary assessment of model transferability across ecosystem types.”*

2. While it is true that groundwater contributions are inherently variable, presenting this as a key finding is too general. The abstract could be more compelling by specifying how the variability manifests in your study systems (e.g., its primary drivers or magnitude) and what new understanding it provides.

We thank the Reviewer for this comment. To address this, we have tried to be more specific and highlight the exact groundwater contributions in the abstract. Below is the revised sentence

Lines 49-54 in revised version:

*“Notably, benefits of ST-SC-MF6 simulation were particularly prominent during both wet and dry periods, i.e. improving the representation of: a) the  $\theta$  dynamics during wet conditions, when hydraulic connectivity between the phreatic surface and the unsaturated zone is strong, b) the root water uptake and carbon assimilation during dry periods, when shallow groundwater mitigated vegetative water stress. and c) the  $T_s$  dynamics during both wet and dry periods, influenced by both GT and GWL, and further propagated to other  $T_s$  dependent processes.”*

3. The statement attributes the benefits during dry periods to groundwater mitigating water stress. To substantiate this key conclusion, it is necessary to clarify: how was this causal relationship established? What analysis directly connects the simulated benefits to the reduction of vegetative water stress?

We thank the Reviewer for this comment. To address this, we have refined one of the Discussion subsections to focus on the groundwater contribution to evapotranspiration components. In subsections 4.3.2. and 4.3.3., we have substantially rewritten the text to highlight the contribution of the groundwater in root soil water uptake (lines 990-1004) and root groundwater uptake (lines 1006-1026) respectively. Additionally, we strengthened the discussion with relevant literature and, where appropriate, with comparisons with previous studies.

4. Lines 41-42: The performance description here would be more concrete if supported by a specific numerical value, such as a key performance indicator or an error metric.

Agreed, we have updated the relevant sentence (lines 46-49) with the difference in the KGE values (which was also added to Table 5) between the two setups (with and without groundwater).

Lines 46-49 in revised version:

*“The ST-SC-MF6 showed better agreement with observations than the ST-SC, improving the KGE values for: a)  $T_s$  (+0.19), and ET (+0.35) at Loobos, b)  $\theta$  (+0.30), ET (+0.11), NEE (+0.18), and SIF (+0.05) at Cabauw, and c)  $\theta$  (+0.21), and ET (+0.15) at Veenkampen.”*

5. The summary should move beyond stating the known importance of groundwater. Its stronger point is the demonstrated need to include groundwater processes in models, and this focus should be sharpened.

Okay, the short summary is rephrased to demonstrate the importance of including groundwater in soil-plant-atmosphere continuum models (lines 59-64).

Lines 59-64 in revised version:

*“Groundwater is ignored in soil-plant-atmosphere continuum (SPAC) models, potentially underestimating its impact on soil and vegetation simulation. To address this, an integrated ecohydrological model is developed by coupling a SPAC model to a hydrological model. The coupled model improved simulation of soil moisture, temperature, evapotranspiration, carbon fluxes and fluorescence. The findings highlight groundwater role in ecosystem dynamics and advance water, energy and carbon cycle modelling.”*

6. The descriptions of the various model types are somewhat repetitive. Condensing this section would improve clarity and focus.

Okay, the introduction is shortened. The description of the models is moved to Methods (lines 237-256).

7. Lines 61-62: The jump to ECV indicators here is logically abrupt. Consider combining this sentence with the next and adding a brief transition.

Okay, the emphasize of the essential climate variables is removed, and this paragraph is rephrased. In the revised version, we only give examples of the variables (in particular, the vegetation and soil variables) that are used to monitor and evaluate ecosystem functioning (Lines 76-84).

Lines 76-84 in revised version:

*“To assess and monitor ecosystem functioning, a range of different physical variables are recognized (Bojinski et al., 2014; GCOS, 2025). Examples of such variables, in particular the vegetation ones, are: leaf area index, above- and below-ground biomass (Atkins et al., 2018; de Conto et al., 2024), gross primary productivity, net ecosystem exchange, evapotranspiration, plant phenology (Shi et al., 2023) and more recently the sun-induced chlorophyll fluorescence (SIF), which serves as a proxy for tracking the photosynthesis process (Sun et al., 2023). Additionally, other soil-related variables play a significant role in characterizing the ecosystems’ functioning, such as soil moisture, surface and subsurface temperature, and evaporation (Bojinski et al., 2014; Zeng et al., 2019; GCOS, 2025).”*

8. Lines 81-85: Presenting model developments requires more than an isolated list of additions. A concise review of the baseline SPAC model’s history and key features is needed to root these advancements in a clear foundation.

We thank the reviewer for this helpful suggestion. We agree that presenting model developments should not appear as an isolated list of additions without sufficient context. However, our aim is not to present detailed model developments, but rather to highlight the limitations of such SPAC models since SPAC models either ignore or oversimplify the underlying groundwater system as stated in lines 146-167. We then go towards the available IEMs that overcome the SPAC models' limitations and present the capabilities of the available IEMs in Table 1, concluding with the limitations of the available IEMs, as none of the available IEMs includes simulation of groundwater heat processes and SIF (as stated in lines 185-187). The limitation of SPAC models and available IEMs explicitly motivates the research gap, leading to our research objective (which was extended by adding specific objectives in lines 200-205). Therefore, we respectfully prefer not to expand the

text about the developments of the SPAC models and IEMs to avoid substantial expanding of the manuscript and shifting the focus away from the methodological contribution of the present work.

9. Lines 94-96: The causal claim here is unsupported. An increase in groundwater discharge does not logically equate to being “exposed” to climate change impacts. Please rephrase to accurately describe the relationship.

Okay, the sentence is rephrased (lines 117-121) to be:

*"Groundwater-dependent ecosystems can be exposed to surface water – groundwater interactions, through altered recharge or discharge rates. Such changes may impair their ecological function if groundwater availability declines, particularly under future climate regimes (Otoo et al., 2025)."*

10. Lines 100–103: The supporting role of groundwater in vegetation physiology is attributed primarily to its relative hydrological stability, not to a lag phenomenon. The wording should be adjusted to reflect this correct causal mechanism.

Okay, the sentence is rephrased (lines 125-129) to be:

*"Additionally, groundwater systems typically respond more slowly to meteorological forcing, as compared to the surface water or the near-surface soil moisture, due to their larger storage and longer flow paths. This relative hydrological stability allows groundwater to continue supporting vegetation in mitigating water stress effects during prolonged drought periods after soil moisture has been depleted."*

11. Lines 110–111: This sentence appears logically disconnected from both the preceding and following content. Its relevance to the narrative flow should be clarified or it should be rephrased to establish a clearer connection.

Okay, the sentence is removed.

12. Lines 157–158: The statement notes that existing models “all lack the representation of groundwater heat processes and groundwater temperature simulation.” To strengthen the manuscript, please clarify whether your study addresses these specific gaps. If it does, highlighting this contribution in the abstract would underscore its significance. If not, consider phrasing the limitation in a more measured, forward-looking manner (e.g., “few studies have incorporated...”).

Yes, our manuscript addresses this gap as mentioned in the next lines of the same paragraph (lines 187-189). To clarify this point, we have added it explicitly in the study objectives (lines 200-205) and also in the abstract (lines 32-35).

Lines 187-189 in revised version:

*"Hence, a new IEM is needed that can physically-based integrate the groundwater processes (mass and energy) with detailed energy, carbon and SIF fluxes"*

In addition, we have added a new section in the Discussion (Section 4.2), with three subsections, to highlight:

- The impact of groundwater temperature on soil temperature (subsection 4.2.1).

- The impact of groundwater level on soil temperature (subsection 4.2.2).
- The impact of groundwater on surface soil temperature and canopy temperature (subsection 4.2.3).

13. Lines 165–172: This paragraph does not effectively articulate or foreground the key research contributions of this study. It should be revised to clearly underscore the novel points being presented.

Okay, the research objectives (lines 200-206) are extended to state the novel contributions.

Lines 200-206 in revised version:

*“This study aims to investigate the role of groundwater in the soil-plant-atmosphere continuum with a focus on assessing the benefits of explicitly incorporating the groundwater component in modelling soil-plant-atmosphere processes. In particular, the specific objectives are to investigate the impacts of the groundwater component (mass and energy) on: 1) regulating the soil profile, in terms of moisture and temperature, 2) the dynamics of the unsaturated/saturated zones and its further impact on water fluxes, and 3) the carbon fluxes and vegetation physiological processes (i.e. root water uptake and fluorescence).”*

14. Lines 169–170: The parenthetical remark “(which will be extended to 2D/3D in a follow-up study)” pertains to future work rather than the present study’s contributions. Unless critical for contextualizing current limitations, it can be removed to maintain focus on the reported research.

Agreed and removed.

15. Lines 173–192: This discussion would be more logically positioned within the earlier section that introduces the first two model types, thereby improving the overall structural coherence.

Okay, the description of the two models is moved to Methods (Lines 237-256) to be positioned immediately before the coupling of the two models (as also suggested by Reviewer 1).

16. A clear scientific hypothesis is needed in the Introduction to establish focus and testability.

Okay, the hypothesis is expanded with details (lines 189-196) to align also with the specific objectives (lines 200-206), as also mentioned in our reply to detailed comment 13.

Lines 189-196 in revised version:

*“The hypothesis of this study is that explicit representation of groundwater mass and energy in an IEM framework would improve quantification of the soil–plant–atmosphere continuum by: a) improving the simulation of soil moisture and temperature profiles, b) enabling two-way feedbacks between the unsaturated and saturated zone and their further impact on water fluxes, and c) revealing the groundwater’s influence on carbon fluxes and vegetation physiological processes such as root water uptake and fluorescence. Collectively, testing these hypotheses would unravel the groundwater role in shaping the SPAC dynamics and deepen the understanding of ecosystem functioning.”*

17. **Study Sites and Data:** The rationale for selecting these three specific locations to apply the model requires a more objective and well-reasoned explanation. Please clarify the scientific basis for their selection (e.g., representing a gradient of key conditions, data availability, or specific hypotheses to be tested) to justify the study design.

Okay, we have explained the motivations behind the selection of the study sites in our reply to detailed comment 1.

18. **Lines 368–371:** The content of this paragraph is more closely aligned with the site or data description. For better organizational flow, it is recommended to relocate it to Section 2.1.

Okay, we have moved this paragraph to lines 205-212 to be positioned after the study objectives.

Lines 205-212 in revised version:

*“To this end, a new IEM framework is developed by coupling the STEMMUS-SCOPE SPAC model to the MODFLOW 6 IHM model and applied to three sites in the Netherlands (i.e., Loobos, Cabauw, and Veenkampen). Since STEMMUS-SCOPE is a point-based model, MODFLOW 6 is a grid-based model, and this study is the first example of coupling these two models, the ST-SC-MF6 models were set up in a vertically-stacked orientation (1-dimensional, 1D). The 1D models were assumed to be representative of the ecosystem of each of the corresponding sites.”*

19. **Lines 426–427:** The criteria for deeming the listed RMSE values as “acceptable” are not self-evident. Please provide a justification or cite relevant references that establish these specific thresholds as benchmarks for model performance acceptance in similar studies. This is essential for assessing the rigor of your model calibration.

Okay, we have added multiple references which used similar RMSE values (lines 531-535).

Lines 531-535 in revised version:

*“The RMSE value was deemed acceptable when it was minimized to  $0.1 \text{ m}^3 \text{ m}^{-3}$ ,  $2^\circ\text{C}$ ,  $0.5 \text{ m}$ ,  $1^\circ\text{C}$ ,  $1 \text{ mm day}^{-1}$ ,  $3 \text{ gC m}^{-2} \text{ day}^{-1}$ ,  $5 \text{ gC m}^{-2} \text{ day}^{-1}$ , and  $0.25 \text{ mW m}^{-2} \text{ um}^{-1} \text{ sr}^{-1}$  for  $\theta$ ,  $T_s$ ,  $\text{GWL}$ ,  $\text{GT}$ ,  $\text{ET}$ ,  $\text{GPP}$ ,  $\text{NEE}$ , and  $\text{SIF}$ , respectively, as similar ranges were indicated in former studies (Lekula & Lubczynski, 2019; Pacheco-Labrador et al., 2019; Thum et al., 2019; Van der Velde et al., 2023; Benz et al., 2024; Qian et al., 2024; Yang & Lei, 2024; Wang et al., 2026).”*

20. The results section contains excessive descriptive text that merely reiterates what is already visible in the figures. To strengthen this section, focus on concisely presenting the key findings and their interpretation, rather than describing the data trends in detail.

Okay, we have restructured the Results section completely based on themes (as below) rather than study site to focus on key findings and their interpretations and also for easier comparison between the three sites. The new themes are:

- 1) Overall model's performance
- 2) Saturated zone dynamics
- 3) Unsaturated zone dynamics (soil moisture and soil temperature)
- 4) Evapotranspiration

## 5) Carbon fluxes and SIF

Additionally, Results tables (Tables 5 and 6 in the revised version) were supported by additional information (i.e. differences in KGE and RMSE values) to help for quick and easier comparisons and reduce the corresponding text.

**21. Lines 460-463:** The results indicate that the model coupling did not improve the accuracy of GPP simulations. This appears to contradict the central hypothesis regarding groundwater's significant influence on vegetation. Please discuss this discrepancy. Does it suggest the influence is negligible under these conditions, or are there other explanatory factors (e.g., model limitations, data scale)?

We thank the Reviewer for this comment. We acknowledge that in Loobos, the GPP simulations of both cases (with and without groundwater) aligned well with the observations. However, this was a specific case due to the characteristics of the Loobos site, which has relatively moderate groundwater levels and generally low values of soil moisture (max observed value was  $0.16 \text{ m}^3\text{m}^{-3}$ ). However, in Cabauw, we observed a significant improvement in the NEE simulation (where GPP was not available in Cabauw) for the coupled ST-SC-MF6 model than the ST-SC model.

- To clarify, we have split and expanded the related text in the Discussion into four paragraphs:
- Paragraph 1 (lines 1031-1041) describes the physical relationship between groundwater and carbon fluxes.
- Paragraph 2 (lines 1042-1056) explains the Cabauw case.
- Paragraph 3 (lines 1057-1062) explains the Loobos case.
- Paragraph 4 (lines 1063-1070) explains the difference between Cabauw and Loobos.

### Lines 1042-1056 in revised version:

*“At Cabauw, where GWL is relatively shallow, exerting strong control on the  $\theta$  dynamics, the ST-SC-M6 showed improved simulation of both  $\theta$  and NEE. In particular, the NEE was increased during the dry periods (Figs. 7e and f) due to enhancement in the simulated  $\theta$  resulting from dynamically updated GWL values. This simulated mechanism aligns with recent findings that groundwater depth is a critical control on ecosystem productivity, with higher carbon assimilation occurring under shallow groundwater conditions (Glanville et al., 2023). Specifically, in meadow ecosystems with shallow GWL, such as in Cabauw, shallow groundwater can enhance net carbon uptake by suppressing soil respiration under near-saturated soil conditions, thereby increasing the net ecosystem carbon sink (Sun et al., 2021). Their analysis, in two meadow sites in the Tibetan Plateau in China, revealed that the influence of high  $\theta$ , maintained by shallow GWL, played a more significant role in controlling the carbon fluxes rather than low temperature. Moreover, the pronounced improvement of simulated NEE during the dry periods (Figs. 7e and f) is consistent with the evidence that there is a significant positive correlation between groundwater depth and carbon fluxes under certain conditions, i.e. when vegetation activity is minimum (Koirala et al., 2017).”*

### Lines 1057-1062 in revised version:

*“Meanwhile, at the Loobos site, there was negligible GWL effect on the simulated  $\theta$  and consequently on the simulated GPP (Figs. 7b and c). This is because: a) the simulated GWL was relatively moderate (3 – 5 m), limiting its hydraulic connection to the unsaturated zone, consistent*

*with the observation that deeper GWL reduces carbon fluxes (Glanville et al., 2023), and b) the maximum observed  $\theta$  was generally low ( $0.16 \text{ m}^3 \text{ m}^{-3}$ ), indicating low water availability for vegetation and consequently poor carbon assimilation.”*

**22. Model Applicability Concern:** The study sites at Cabauw and Veenkampen are situated in agricultural settings. A key limitation is that the employed 1D model framework cannot account for lateral water fluxes, which are likely significant due to irrigation and pumping practices. Please address the potential impact of this simplification on the results and discuss its validity for these sites.

We thank the Reviewer for this constructive comment. To address this comment, we have added relevant text in the new Results subsection “3.2. Saturated zone dynamics” (lines 584-590)

Lines 584-590 in revised version:

*“Although the lateral groundwater flow was not simulated in the present 1D applications, its potential effect was indirectly accounted for calibrating the ST-SC-MF6 models against the observed GWL. By adjusting aquifer parameters so that simulated GWL closely matched observed ones, the 1D ST-SC-MF6 models implicitly incorporated the net outcome of lateral inflows, drainage, irrigation, and pumping on the GWL dynamics. Nevertheless, incorporating lateral groundwater flow in future 2D/3D applications would likely improve process realism, particularly when assessing groundwater as a managed water resource and its broader role in ecosystem functioning.”*

**23.** The section would benefit from a more substantive and critical analysis. Currently, it tends to restate results rather than interpret their significance. For instance, the claim that comparing ST-SC and ST-SC-MF6 results “highlights the benefits” (Lines 586-588) is vague. The discussion should explicitly articulate what these specific benefits are—quantitative improvements, mechanistic insights gained, or limitations addressed—and explain why they matter in the broader context of SPAC modelling.

Okay, we have rephrased this sentence (lines 819-820) to:

*“The benefits of representing the groundwater in models of soil-plant-atmosphere continuum are described in the following subsections.”*

In addition, we have significantly expanded the Discussion Section to explain in detail those benefits and the differences across the three sites. The structure of the revised Discussion is explained in our reply to major comment 3.

**24.** The discussion appears somewhat isolated from the wider body of literature. To strengthen its scholarly contribution, findings should be actively compared and contrasted with existing studies.

We thank the Reviewer for this constructive comment. To address this, the Discussion section has been significantly expanded and reorganized to better highlight the main contributions of the study and to position the findings within the broader state of knowledge. The new structure of the revised Discussion is explained in our reply to major comment 3.

25. The current recommendations are overly general. To be more useful, this section should transition from stating ambitions (e.g., expanding to 2D/3D) to providing actionable insights based on this study. Please address: 1) What was the actual simulation quality at the three sites? Under what conditions did the coupled model perform well or poorly? 2)

Section 4.4 (renumbered as Section 4.5 in the revised manuscript) is intended to summarize the limitations of the current study and provide recommendations for future research. We believed that the mentioned recommendations would improve the quality of future simulations based on the experience gained during this work.

To further enhance the actionability of the recommendations, we emphasize the need for a multi-objective automatic calibration framework in future studies to improve the quality of the simulations (lines 1089-1093). Regarding what conditions did the coupled model perform well or poor, this is summarized in the Abstract (lines 49-54), where the improvements achieved at each site for specific variables are briefly reported.

Lines 1089-1093 in revised version:

*“Furthermore, although calibration was performed manually in this proof-of-concept study, future work should incorporate a multi-objective automatic calibration framework to systematically reduce parameter uncertainty and enhance model robustness. Implementing such an approach in future 2D/3D applications will increase confidence in predictive simulations and strengthen the broader applicability of the developed IEM.”*

Lines 49-54 in revised version:

*“Notably, benefits of ST-SC-MF6 simulation were particularly prominent during both wet and dry periods, i.e. improving the representation of: a) the  $\theta$  dynamics during wet conditions, when hydraulic connectivity between the phreatic surface and the unsaturated zone is strong, b) the root water uptake and carbon assimilation during dry periods, when shallow groundwater mitigated vegetative water stress. and c) the  $T_s$  dynamics during both wet and dry periods, influenced by both GT and GWL, and further propagated to other  $T_s$  dependent processes.”*

26. Based on your experience, what are the main practical or theoretical constraints (e.g., computational cost, data requirements, specific site conditions) that currently limit the model's application? 3) Is the coupled framework readily operable and transferable to other regions? What are the key prerequisites (data, expertise) and potential pitfalls for potential users?

We thank the Reviewer for this comment.

Regarding the main practical or theoretical constraints

- Based on our experience, the applicability of the coupled framework depends on site conditions. In particular, the model is most beneficial at sites where there is a hydraulic and/or thermal connection between the groundwater system and the unsaturated zone or land surface. When groundwater is relatively deep (e.g., > 10 m below land surface), the influence of groundwater on soil–plant–atmosphere processes becomes limited, and therefore the

coupled model may not produce significantly different results compared to the standalone STEMMUS-SCOPE model.

- In addition, based on our experience, calibration of the coupled framework can also be challenging due to the relatively large number of parameters and interacting processes. In this study, calibration was particularly challenging at the Loobos site, where achieving a balanced representation of water, energy, and carbon fluxes required iterative adjustments. Therefore, future studies would benefit from applying automatic or multi-objective calibration approaches as we highlighted in the revised version (lines 1089-1093) to improve parameter optimization and model robustness.

Regarding if the coupled framework readily operable and transferable to other regions

- Yes, the coupled framework is generally operable and transferable to other regions. To apply the framework, users need to prepare the required input datasets and ensure the availability of some key observations to support model calibration and validation. These required inputs and observations are summarized in Table 3 of the revised manuscript.

Regarding the key prerequisites (data, expertise) and potential pitfalls for potential users

- The main prerequisites relate to both data availability and modelling expertise. As summarized in Table 3 of the revised manuscript, the model requires meteorological forcing, soil and vegetation parameters, and hydrogeological information as inputs, as well as observational datasets for calibration and evaluation. To comprehensively evaluate the role of groundwater in the water, energy, and carbon cycles, multiple types of observations are desirable (e.g., soil moisture, soil temperature, evapotranspiration, carbon fluxes, groundwater level, groundwater temperature, and vegetation-related variables such as SIF). However, such comprehensive datasets are not always available at many sites, which may limit the depth of model evaluation. In addition, users need to be familiar with the modelling environments of both STEMMUS-SCOPE and MODFLOW 6, which may require some time and effort for new users.

#### Comments on the figures:

1. Figures 3-9: Please reconsider the use of red-green color combinations, as they are not accessible to readers with color vision deficiencies. Employing a colorblind-friendly palette (e.g., blue-orange) or differentiating lines with symbols/patterns is strongly recommended.

Agreed, we have revised all figures to use colorblind-accessible colors (orange and blue).

2. Figures 4, 5, 6, and 9: The key comparisons between models are difficult to discern due to high data density and overlapping traces in the main panels. To improve clarity, consider 1) plotting the difference between models directly, 2) using faceted subplots for major variables, or 3) highlighting specific time periods of interest.

Agreed, we have revised Figures 3 to 9 to improve their readability.