Authors replay to Reviewers' comments on manuscript egusphere-2025-4179

The reviewer comments are in black, while the authors' responses are in **bold blue** text. To avoid repetitions in our replies, we apply order for the comments (1, 2, 3, etc.), to allow for cross-referencing our responses (if needed).

Editor (Patricia Saco) decision: Publish subject to minor revisions

We have now received the comments of two referees on the revised manuscript. After my own assessment of the manuscript, I agree with the reviewers that the revised manuscript has improved during the review process. However, there is still work to be done before the manuscript can be recommended for publication. The referees have provided very useful feedback, with key points focusing on model calibration scope, scientific contribution, and clarity of presentation, as follows:

• Scientific contribution needs to be articulated more clearly
Both referees recognise the study's potential but highlight that the novelty, implications, and relationship to previous modelling work need to be described more clearly.

• Discussion needs improvement (identified by both reviewers): better synthesis, context, and scientific positioning; an explanation of how prior studies relate to the findings (rather than primarily presenting technical results); and greater emphasis on the study's broader contribution rather than repetition of results.

• Clarify calibration and evaluation, with an improved description and justification of the calibration procedure (justify the manual calibration strategy, as requested by Reviewer 1, and address Reviewer 3's comments regarding the temperature parameters).

• The manuscript should be more precise and concise
Both reviewers identify issues that reduce clarity, including overly technical passages, repetition, and insufficiently explained statements or references.

• The presentation of the study's limitations and model scope should be clearer
Beyond these main points, the additional reviewer comments provide constructive and detailed feedback that the authors should address carefully, including suggestions for enhancing the figures and other useful minor revisions to improve readability and flow.

Dear Dr. Patricia Saco,
We would like to sincerely thank you for giving us the opportunity to revise the manuscript and Reviewers, for their constructive comments and appreciation. We have carefully considered all the comments and address them one by one. 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 and also the expectations of the Reviewers.

Sincerely,
Mostafa Gomaa Daoud
On behalf of all co-authors

Reviewer 3

General comments

This manuscript, "Unravelling groundwater's role in soil–plant–atmosphere continuum: Integrated ecohydrological modelling approach using STEMMUS–SCOPE and MODFLOW 6", investigates the influence of groundwater on water and carbon fluxes within the soil–plant–atmosphere continuum (SPAC) by coupling the STEMMUS–SCOPE model with the groundwater model MODFLOW 6.

The topic fits well within the scope of HESS. Understanding the interactions between groundwater and ecosystem functioning, and improving the representation of groundwater–ecosystem coupling, is currently one of the major research directions in the ecohydrology community. The modelling framework is technically sound, the manuscript is generally well organized, the writing is clear, and the figures are of good quality. I also greatly admire the authors' efforts in addressing the comments from the first round of review, and the manuscript has been substantially improved.

Nevertheless, I believe several issues still need to be addressed before the manuscript is suitable for publication. Overall, I recommend moderate revision before the manuscript can be considered for publication.

Dear Reviewer,
We would like to sincerely thank you for your precious time in reviewing our manuscript and for providing valuable comments that led to possible improvements in the current version. The authors have carefully considered the comments and addressed every one of them in the best way we could. 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. Below, you could find our responses to your comments.

1. The scope of the research should be more clearly defined. The manuscript aims to investigate the influence of groundwater on the SPAC. However, groundwater interacts with the SPAC through multiple pathways, and the three study sites presented here do not represent all possible groundwater–vegetation interactions. For example, an increasing number of studies have demonstrated that vegetation may obtain water directly from weathered bedrock or fractured bedrock.
   McCormick, E.L., Dralle, D.N., Hahm, W.J. et al. Widespread woody plant use of water stored in bedrock. Nature 597, 225–229 (2021). https://doi.org/10.1038/s41586-021-03761-3

   In all three study sites used here, plants are able to access groundwater directly. Groundwater is very shallow at Cabauw and Veenkampen, while at Loobos, although groundwater is deeper, the prescribed maximum rooting depth can still reach the groundwater table during certain periods. Therefore, the manuscript essentially investigates the influence of shallow groundwater systems, or more specifically groundwater-dependent ecosystems (GDEs), rather than groundwater systems in general. This distinction should be reflected more clearly throughout the manuscript, including in the title.

We thank the Reviewer for these constructive comments and for drawing our attention to McCormick et al. (2021), which provides valuable insights into plant water uptake from weathered or fractured bedrock storage.

We fully agree with the Reviewer's observation regarding our application: the three study sites evaluated in this paper specifically represent shallow-to-moderate water table conditions (0–5 m depth) where vegetation directly interacts with the capillary fringe or saturated zone (i.e., groundwater-dependent ecosystems, GDEs). However, we consider the underlying physical coupling of the STEMMUS-SCOPE & MODFLOW 6 modelling framework, presented in our manuscript, to be more generic, extending beyond strictly shallow water table conditions or specific GDE settings. The framework simulates the fundamental physical interactions between groundwater and SPAC processes whenever hydraulic and thermal connectivity exists between the soil and/or plants' root zone and the underlying saturated domain — making the developed IEM applicable across shallow, moderate, or deep water table conditions. It is true that we have not tested the developed model in deep groundwater conditions in the present three site evaluations. However, members of the authorship list are currently working on a follow-up study under deeper water table conditions, which we view as a promising future direction for demonstrating the broader applicability of the developed model.

To address the Reviewer's comments while preserving the general scope of the modeling framework:

- **We have added a dedicated sentence in Section 1 (lines 111-112) citing McCormick et al. (2021) to explicitly contextualize bedrock moisture uptake alongside shallow porous-media groundwater systems.**

  Lines 111-112 in revised version:
  "In other upland bedrock ecosystems, woody vegetation can even access moisture held in weathered or fractured bedrock storage (McCormick et al., 2021)."

- **We have highlighted testing the model under deeper groundwater regimes and alternative hydrogeological settings as a key direction in our Discussion section (Section 4.5, lines 920-923).**

  Lines 920-923 in revised version:
  "In addition, while the three presented applications focused on shallow-to-moderate groundwater depths (0 – 5.0 m), applying the developed IEM framework to deeper groundwater systems would present another valuable avenue for future research."

However, we respectfully prefer to maintain the current title as it stands to avoid unnecessarily narrowing the scope of the proposed modeling framework, since we are confident that it will work in deeper water table conditions too.

2. The manuscript could be made more concise. Although the overall structure is logical and easy to follow, some sections are unnecessarily lengthy. In several places, the figures already convey the

main findings, while the text repeats the same information. Reducing redundancy between the figures and the text would improve readability.

We appreciate the Reviewer's perspective on manuscript length and conciseness. We would like to clarify that the detailed text accompanying the figures and statistical tables was intentionally structured to guide the reader through specific features of complex, multi-site hydrographs. Furthermore, some of this detailed process description was expanded in response to specific requests from previous reviewers to provide a complete, transparent assessment. While we acknowledge the Reviewer's point about overall manuscript length, we respectfully prefer to retain these explicit text descriptions alongside the figures to ensure readers can easily interpret the key highlights and physical mechanisms shown in the figures without misinterpretation.

3. The classification presented in the Introduction requires revision. In particular, the categorization of SPAC models and integrated hydrological models (IHMs) is not entirely accurate. The main contribution of this study is the coupling of a SPAC model with a groundwater model. However, recent years have seen rapid developments in land surface models (LSMs), many of which have incorporated plant hydraulic processes, including but not limited to Noah-MP and CLM. These models are capable of representing, or partially representing, groundwater processes and groundwater–vegetation interactions. This important category of models has largely been overlooked in the current review of existing modelling approaches.

We have answered this point in detail in our reply number 7 (see below).

4. The Discussion should better position the proposed modelling framework relative to existing models. A qualitative comparison highlighting the advantages and limitations of STEMMUS–SCOPE–MODFLOW 6 relative to current land surface models that include plant hydraulics would strengthen the manuscript and better demonstrate its scientific contribution.

We thank the Reviewer for this comment. A list of the existing integrated eco-hydrological models (IEMs), highlighting their capabilities is already provided in Table 1 (line 186). Furthermore, we would like to clarify that the Discussion has been extensively expanded after the first review round to highlight the specific mechanisms and impacts of adding the groundwater component in modeling the SPAC processes. Adding more details regarding comparisons with other land surface models would expand the manuscript length, which is already long enough. Therefore, we respectfully prefer not to add further details in the Discussion section.

Detailed comments

5. I suggest revising the title to clarify that the study focuses on shallow groundwater systems or groundwater-dependent ecosystems (GDEs) rather than groundwater systems in general.

We have answered this point in detail in our reply number 1 (see above).

6. Line 43: The abbreviation GPP should be introduced after the first occurrence of gross primary productivity. More generally, the manuscript contains a large number of abbreviations. I

recommend including a list of abbreviations at the beginning or end of the manuscript to improve readability. Although Figure 1 contains some abbreviations, it is just part of them.

Done. The GPP symbol is added in line 42. A list of abbreviations is added at the end of the manuscript (lines 954-955).

7. Lines 87–93: I do not fully agree with the current classification. Models such as CLM5 and TC model should not be classified as SPAC models; rather, they are land surface models (LSMs) that incorporate plant hydraulic processes. As noted above, this important category of LSMs has largely been overlooked. Instead, these models appear to have been implicitly divided into the SPAC model and IHM categories, which I do not consider appropriate.

Done. We thank the Reviewer for this comment. The related text has been revised to separate the SPAC models and LSMs explicitly, and citing examples for both as follows (lines 80-103):

Lines 80-103 in revised version:
"Soil-plant-atmosphere continuum (SPAC) models and land surface models (LSMs) are widely used to simulate most of the aforementioned variables. SPAC models simulate the transfer of thermal energy and mass (water and carbon) between soil, plants (through roots, stems and leaves), and the atmosphere (Guo, 1992). LSMs simulate the movement of energy, water, and carbon between the Earth's land surface and the astrosphere. Both model categories are often used interchangeably in the literature as they share core physical functionalities. Furthermore, both LSMs/SPAC models are capable of assimilating remote sensing (RS) information such as canopy reflectance, vegetation optical depth, and SIF (X. Yang et al., 2021). Integrating the RS data into these models enables upscaling the processes from point (plant) scale to regional scales (Senf, 2022; Van Cleemput et al., 2025), allowing for predicting ecosystem responses to regional environmental changes (Dronova & Taddeo, 2022). Examples of SPAC models are the Soil-Plant-Atmosphere (SPA) model (Williams et al., 1996; Williams et al., 2001), STEMMUS-SCOPE model – the integration of Simultaneous Transfer of Energy, Mass, and Momentum in Unsaturated Soil (STEMMUS) and Soil Canopy Observation, Photochemistry, and Energy Fluxes (SCOPE) models (Wang et al., 2021), and the Soil Water Atmosphere Plant (SWAP) model [(van Dam et al., 2008) – specifically the version coupled to the World Food Studies (WOFOST) model (de Wit et al., 2019)]. Examples of LSMs are the Community Land Surface (CLM5) model (Lawrence et al., 2019), the Tethys-Chloris (T&C) model (Fatichi et al., 2012), and the Noah-MP model (Niu et al., 2011). However, existing LSMs/SPAC models either ignore or oversimplify the underlying groundwater system by neglecting spatio-temporal variability in groundwater variables (e.g., head and temperature), and processes (e.g., recharge, discharge and groundwater evapotranspiration). Thus, the LSMs/SPAC models lack or simplify the groundwater influence on the simulated SPAC processes."

8. Lines 98–115: This section actually defines the study scope as groundwater-dependent ecosystems (GDEs). Again, I suggest emphasizing this point more consistently throughout the manuscript, including in the title.

Please see our detailed response to Comment 1 above. We have dedicated one additional sentence (lines 111-112) sentence within this scope section (Lines 104–123) and also in the Recommendation

section lines 920-923) to explicitly highlight the applicability of the developed model across different ecosystems and different water table conditions, while preserving the broader physical generality of the modeling framework in the manuscript title and scope.

9. Line 121: MODFLOW 6 is generally regarded as a groundwater model rather than an integrated hydrological model (IHM), since it has only limited representation of surface hydrological processes. Models such as ParFlow are more commonly classified as IHMs.

We thank the Reviewer for this comment and appreciate the opportunity to clarify our terminology. While MODFLOW historically originated as a groundwater-only simulator (GWF), MODFLOW 6 has evolved into an integrated framework capable of simulating additional surface and subsurface processes. Specifically, MODFLOW 6 includes packages for unsaturated zone flow, overland flow, surface water body interactions (lakes, streams, drains), and solute transport (Langevin et al., 2017; Langevin et al., 2023). In addition, the MODFLOW 6 generic farmwork allowed for representing other processes, i.e. the re-infiltration process within the simulation (Daoud et., 2022). In the same paper, an IHM classification is introduced the different type of IHMs, fully-coupled IHMs (e.g. CATHY (Camporese, Hydrogeosphere, and PARFLOW) and simplified IHMs (GSFLOW and MODFLOW 6). Furthermore, the flexible architecture of MODFLOW 6 allows for representing additional physical processes, such as re-infiltration and improve the overland flow representation through the Cascade Routing and Re-infiltration (CRR) concept (Daoud et al., 2022). As categorized by Daoud et al. (2022), IHMs can be classified as fully coupled IHMs (e.g., HydroGeoSphere, ParFlow, CATHY) and simplified IHMs (e.g., GSFLOW and MODFLOW 6).

In addition to flow processes, our study leverages the recently added Groundwater Energy Transport (GWE) package (Morway et al., 2025; Langevin et al., 2022) to simulate heat transport processes. Therefore, given its integrated surface–subsurface hydro-thermal capabilities, we consider MODFLOW 6 to function as an IHM. A brief description of MODFLOW 6 was originally included in the text (lines 215-222).

10. Line 141: A reference is needed.

Done. It is added in line 149.

11. Line 333: From a logical perspective, Equation (3) appears to first calculate RWUs, after which RWU = RWUs. Therefore, the equations may be more logically presented as RWUs = ..., followed by RWU = RWUs.

Done. The equation is modified (line 342).

12. Line 414: Please use the abbreviation for STEMMUS–SCOPE consistently. Throughout the manuscript, both the full model's name and the abbreviation are used inconsistently.

Thank you for pointing this out. We have clarified and standardized our terminology throughout the manuscript. The abbreviations ST-SC (standalone simulation) and ST-SC-MF6 (coupled simulation) are used specifically when comparing the two model runs. The terms STEMMUS-SCOPE is reserved for referring to the underlying model structure or parameter settings that is shared by

both simulations. We have clarified these details at the start of Section 2.4 (lines 415-418):

Lines 415-418 in revised version:
"For brevity, from herein, when comparing simulations, the standalone STEMMUS-SCOPE simulation is referred to as ST-SC and the coupled STEMMUS-SCOPE-MODFLOW 6 simulation is referred to as ST-SC-MF6. The term "STEMMUS-SCOPE" is reserved for referencing the underlying model structure or parameter settings that is shared by both standalone and coupled simulations."

13. Line 415 and Lines 451–455: The information on rooting depths (5 m for Loobos, 1.5 m for Cabauw, and 1.2 m for Veenkampen) is already provided in the corresponding table. Repeating it in the text is unnecessary and could be removed to improve readability.

Done. We have removed the numbers on line 415 and referred to Table 2 instead. The updated sentence in line 415 is as follows:

Lines 424-427 in revised version:
"The unsaturated zone in STEMMUS-SCOPE was modelled with site-specific total thicknesses – each equals or exceeds the site's maximum observed water table depth (Table 2)."

Regarding Lines 451–455, we could not locate any relevant text on this issue, so those lines remain unchanged.

14. Line 469: The evaluation metrics currently used do not account for differences in model complexity. Since additional parameters are introduced, I recommend including an information criterion such as the Akaike Information Criterion (AIC) to evaluate whether the increased model complexity is justified.

We thank the Reviewer for this valuable recommendation. We computed the Akaike Information Criterion (AIC) across all observation types [soil moisture (θ), soil temperature (T_s), evapotranspiration (ET), gross primary productivity (GPP), and net ecosystem exchange (NEE)] at the three study sites using the following equations:

$$AIC = n \ln\left(\frac{RSS}{n}\right) + 2k$$

$$\Delta_{AIC} = AIC_{ST-SC-MF6} - AIC_{ST-SC}$$

where n is number of observations, RSS is residual sum of squares between observed and simulated values, and k is number of estimated parameters.

The calculated AIC values of both models (AIC_{ST-SC} and $AIC_{ST-SC-MF6}$) and their differences (Δ_{AIC}) are summarized in the table below.

Site	Variable	n	AIC_{ST-SC}	$AIC_{ST-SC-MF6}$	Δ_{AIC}
Loobos	θ (Avg)	7.00E+05	-4.51E+06	-5.48E+06	-9.71E+05
	T_s (Avg)	8.26E+05	1.51E+06	7.76E+05	-7.37E+05
	ET	2.89E+03	1.84E+03	-3.22E+01	-1.87E+03

	GPP	1.89E+03	3.77E+03	3.92E+03	+1.46E+02
Cabauw	θ (Avg)	4.75E+05	-1.28E+06	-2.40E+06	-1.12E+06
	T_s (Avg)	1.37E+06	1.62E+06	2.62E+05	-1.35E+06
	ET	2.44E+03	-2.32E+02	-1.31E+03	-1.08E+03
	NEE	1.49E+04	5.81E+04	4.56E+04	-1.25E+04
Veenkampen	θ (Avg)	2.95E+06	-8.94E+06	-1.42E+07	-5.28E+06
	T_s (Avg)	2.95E+06	5.41E+06	4.92E+06	-4.92E+05
	ET	2.83E+03	6.60E+02	-7.63E+02	-1.42E+03
Overall	All Observations	9.31E+06	1.06E+07	7.01E+07	-3.57E+06

As shown by the substantial negative Δ_{AIC} values across almost all observations, the increased model complexity of ST-SC-MF6 yields a significant improvement in predictive performance, strongly justifying the additional parameterization. The only exception is the GPP at Loobos (positive Δ_{AIC}), where groundwater dynamics exert minimal control, as discussed in the Results Section (lines 635-638) and in the Discussion Sections (lines 890-903).

We respectfully prefer not to add this analysis in the manuscript to avoid adding more details and unnecessary length to an already detailed long one. Instead, we have provided three sentences (see below) in the revised version (in the Methods, Results and Discussion Sections) to briefly refer to this evaluation to demonstrate that the added complexity is fully justified.

Sentence in Methods (lines 491-493):
"Additionally, to evaluate whether the increased complexity of ST-SC-MF6 is justified relative to parameter addition, the Akaike Information Criterion (AIC; Akaike, 1973) was computed across all observed variables."

Sentences in Results (lines 502-506):
"Furthermore, the AIC evaluation (details omitted here for brevity) yielded substantially lower values for ST-SC-MF6 across almost all observations as compared to the ST-SC, confirming that the enhanced predictive performance strongly justifies the added model parameterization, with the sole exception of GPP at Loobos where groundwater dynamics exert minimal control."

Sentences in Discussion (lines 664-666):
"The AIC evaluation confirmed that the added model complexity of ST-SC-MF6 is strongly justified by its enhanced predictive accuracy across all study sites."

15. Line 535: How were the temperature-related parameters calibrated? It appears that no temperature observations were available during the calibration period.

We thank the Reviewer for pointing this out. We have added one sentence for clarification (lines 548-552) as follows:

Lines 548-552 in revised version:
"Although, observed GT data in Loobos were available for only one year (1 April 2023 – 31 March 2024) during the validation period, the seasonal GT fluctuations from that period were assumed to be representative of the site's thermal regime and were exceptionally used to calibrate the

thermal parameters (K_T and c_s; Error! Reference source not found.**) during the calibration period (1 April 2016 – 31 March 2020).”**

16. Line 604: Is it appropriate to describe the model performance as "well" when the KGE is below 0.5?

Done. Changed to “showed satisfactory agreement” in the updated line 637.

17. Figure 9: Please keep the figure style consistent with the previous figures. The line width in this figure appears thicker.

Done. The figure is updated with thinner line width (line 801).

18. Conclusion: The current conclusion reads more like a summary of the results. I suggest restructuring it to better emphasize the broader implications, scientific contributions, limitations, and future research directions.

We thank the Reviewer for this suggestion. We would like to clarify that the manuscript already includes a dedicated Recommendations section immediately preceding the Conclusion (Section 4.5, lines 915–932), which explicitly details model limitations and future research directions (e.g., 2D/3D extension, automated calibration, boundary layer schemes). To avoid redundant repetition, we respectfully prefer to maintain the limitations and future research discussion in that dedicated section.

Regarding the Conclusion itself, we have revised the text (lines 934–954) to better highlight our main scientific contributions and broader implications. However, we have retained few site-specific details because the groundwater impact on the SPAC processes is strongly controlled by local site conditions (e.g., water table depth and soil properties). Over-generalizing these findings across all settings could be scientifically misleading.

Lines 934-954 in revised version:
“This study demonstrated the added value of incorporating the groundwater component in modelling the SPAC processes. To achieve this, an IEM framework was developed by coupling the STEMMUS-SCOPE SPAC model with the MODFLOW 6 IHM, resulting in the ST-SC-MF6 IEM. Results from three site applications in the Netherlands (Loobos, Cabauw, and Veenkampen) demonstrate that explicitly including groundwater (mass and energy) in SPAC modelling improves the simulation of key ecohydrological variables, i.e. the θ dynamics, particularly at sites with shallow groundwater tables (Cabauw and Veenkampen) and during wet periods. This improved θ representation propagated to other processes, i.e. Ts and Tc dynamics, ET components, carbon uptake (NEE) and SIF, particularly during dry conditions across all three sites. Remarkably, groundwater impacts on GPP were limited at Loobos, due to its limited specific retention capacity and relatively moderate groundwater depths, emphasizing that groundwater – vegetation interactions are strongly site dependent and controlled by groundwater depth, soil properties, and rooting characteristics. Overall, this study demonstrates that explicitly including groundwater dynamics with SPAC modelling is essential for realistically simulating ecosystem water, energy, and carbon cycles.”

Minor comments

19. Line 84: RS appears for the first time and should therefore be written in full before introducing the abbreviation.

Done. It is added in line 86.

20. Line 134: Consider starting a new paragraph after "Additionally".

Done. A new paragraph starts from line 143.

21. Line 361: Should this be "the annual average air temperature"?

Yes, it is modified in line 370.

22. Table 3: Under the source of leaf area index, placing "MODIS-LAI" in the middle of the table may give the impression that it only applies to the Cabauw site. Please revise the table layout for clarity.

Done. The word "MODIS-LAI" is placed under the three columns (Loobos, Cabauw and Veenkampen) separately instead of one word located in the middle of the corresponding row.

23. Line 467: Insert a line break before "Table".

Done. Added in lines 493-494.

24. Line 552: Perhaps "improved model structure" would be more appropriate?

Done. Added in lines 585.

25. Line 592: ET does not need to be italicized.

We apologized for this typo mistake. Some abbreviations have been shifted to Italic after saving the revision file in PDF format. It is corrected in the revised version (lines 625).

26. Lines 609–610: GPP and NEE do not need to be italicized.

We apologized for this typo mistake. Some abbreviations have been shifted to Italic after saving the revision file in PDF format. They are corrected in the revised version (lines 642-643).

27. Line 639: Use e.g.,.

Done. Changed (lines 674).

28. Line 773: The number "6" does not need to be italicized.

We apologized for this typo mistake. Some numbers have been shifted to Italic after saving the revision file in PDF format. It is corrected in the revised version (line 710).

Reviewer 4

The authors have provided detailed responses to all of the questions raised during the first round of review. Although they have addressed most of the issues and comments raised during the first round, a few points remain to be clarified.

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.

1. On the content: Subsection 2.4.3: The explanation of the calibration strategy needs more information, particularly given the absence of an objective, automated calibration procedure. How can readers be confident that the STEMMUS-SCOPE model has not converged to a local maximum? This possibility could, in itself, explain the lower optimization performance compared with the STEMMUS-SCOPE-MODFLOW 6 model and could therefore have implications for the subsequent analysis presented in the study.

We thank the Reviewer for raising this point. We have added further details in the revised text to clarify that:

- **Both ST-SC and ST-SC-MF6 setups across all three study sites were initialized using identical baseline parameter sets derived from site-specific measurements and relevant literature.**
- **The manual trial-and-error calibration for both ST-SC and ST-SC-MF6 setups followed the exact same step-by-step procedure, adjusting sensitive parameters within identical, physically realistic parameter bounds to maximize KGE and minimize RMSE across all observed state variables.**
- **The enhanced performance of ST-SC-MF6 is fundamentally driven from the physical representation of dynamic surface–subsurface hydro-thermal interactions in the coupled ST-SC-MF6, which are structurally missing in the standalone ST-SC model, rather than discrepancies in parameter search effort.**

To address the Reviewer's comments:

- **We updated Section 2.4.3 (lines 480-484) to detail the identical baseline parameterization and calibration search bounds used for both model setups to guarantee an unbiased baseline setup.**

  Lines 480-484 in revised version:
  "To ensure an unbiased comparison, both ST-SC and ST-SC-MF6 setups across the three sites were initialized using identical baseline parameter values derived from each site measurements and literature. The manual calibration for both setups followed the same

iterative step-by-step procedure within identical, physically realistic parameter bounds (Table 4).”

- **We added an explicit statement at the opening of Section 3.1 (Results, lines 497-502) clarifying that the observed performance gain stems from structural model capabilities rather than differential calibration effort.**

  Lines 497-502 in revised version:
  “At each of the three study sites, the coupled ST-SC-MF6 framework consistently outperformed the standalone ST-SC model across most of the simulated variables (Tables 5, 6 and Figures 4-9). This superior performance stems from better physical representation of dynamic surface–subsurface hydro-thermal interactions in the coupled ST-SC-MF6 (as discussed in detail in subsequent sections)—which are structurally missing in the standalone ST-SC model—rather than discrepancies in parameter search effort.”

- **As noted in our Discussion section (Section 4.4, Lines 923–925), we explicitly highlight implementing automated, multi-objective calibration as a key target for future work .**

  Lines 923-925 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.”.

Additional comments on presentation:

2. L525-531 sub-section 3.2 - These lines may be somewhat problematic because pumping was not explicitly accounted for in the model, unless it was not presented in the methodology section. Acknowledging that the model does not account for lateral flows (runoff and groundwater flow) would be sufficient to explain why better calibration metrics might be obtained with a fully three-dimensional model, which is beyond the scope of the present study.

We thank the Reviewer for this comment. As suggested, we have revised the text to simplify it and removed the confusion part (“the implicit incorporation of the lateral inflows, drainage, irrigation and pumping activates”). The updated text (lines 556-564) is as follows:

Lines 556-564 in revised version:
“Although lateral groundwater flow was not simulated in the present 1D applications, calibrating (i.e., adjusting the aquifer parameters) the ST-SC-MF6 models against the observed GWL allowed the models to capture the local water table state. Nevertheless, incorporating lateral groundwater flow in future 2D/3D applications would likely improve process realism and predictive capability, particularly when assessing groundwater as a managed water resource and its broader role in ecosystem functioning.”

3. Discussion section: Most of the references cited in the discussion section are introduced without sufficiently describing the cited studies or discussing how their findings relate to the material being discussed. This requires readers to have prior knowledge of the cited studies in order to

understand the connections being made with the results of the present study, which makes the overall line of argument more difficult to follow.

We understand the Reviewer's concern regarding the citation context in the Discussion. However, adding more details regarding the cited studies would expand the manuscript length which is already long enough after so many other requests from the Reviewers in this round and last review round. Therefore, we respectfully prefer not to add further details in the discussion section.