# A Summary of Revisions and Responses on "A spatially hybrid hydrological modeling approach based on spatial heterogeneity of watershed characteristics" (Ref: HYDROL63641)

**With regards to comments from the Editor:**

*Although your manuscript falls within the aim and scope of this journal, it is being declined due to lack of sufficient novelty. We receive a much larger number of papers than we are able to accept.*
*Although the problems being addressed are potentially of interest to our readership, your manuscript does not meet the required quality standards to be considered for publication. The research results reported are too premature for publication. More work is needed to substantiate the conclusions in your manuscript.*

We sincerely appreciate the editor's consideration and the reviewers' constructive comments, which significantly improved the quality and clarity of our manuscript. All reviewers acknowledged the core contribution of our study—proposing a flexible hydrological modeling approach—while also raising important concerns regarding the clarification of its novelty and the interpretation of experimental results. We have now substantially revised the manuscript and addressed all comments in detail.

First, regarding the novelty of our approach: our primary contribution lies in proposing a flexible modeling approach that enables spatially hybrid hydrological modeling—a capability that is notably limited in existing modeling frameworks—not in claiming an overall superior accuracy of the spatially hybrid model structure compared to spatially consistent model structures. Although the case study was conducted in a simplified application context, it successfully demonstrated both the integration feasibility (i.e., the ability to combine and execute different subbasin-level model structures within a unified watershed model) and simulation performance of the spatially hybrid model structure. Moreover, the case study highlighted the importance of careful model structure construction and the essential needs for multi-site and long-term observation data to constrain calibration, as discussed in the revised manuscript.

Second, regarding the definition of "physically-based model". We now clearly distinguish between conceptual, semi-physically-based, and fully-physically-based models, and explicitly state that this study focuses on the first two categories. While our approach is theoretically capable of incorporating subbasin-level fully-physically-based simulation algorithms, as discussed in section 4.3, this remains a future work. We provided detailed explanations of the revised definition and classification in Response 2 to Reviewer 1.

Third, concerning the model performance of the proposed approach, we have refined the interpretation of the experimental results. We now emphasize that while the spatially hybrid model achieved generally better performance in terms of NSE and RSR, it also inherited some limitations, such as simulation bias from its conceptual components. We carefully analyzed performance differences observed between calibration and validation periods, and among different model structures. This analysis clarifies that the spatially hybrid model structure provides a promising compromise between integration flexibility and simulation accuracy.

In addition, we added a new discussion section (Section 4) that discusses on the broader implications and limitations of the proposed approach and the exemplified implementation, including the significance of the spatially hybrid modeling approach, the advantages of SEIMS in implementing the proposed approach, and limitations and future directions.

We believe the revised manuscript more clearly demonstrates its scientific contribution and meets the journal's quality standards. We would be grateful for the opportunity to have it reconsidered for publication.

**With regards to comments from the Reviewer #1**

*1. This paper by Wang et al. is a revision of manuscript HYDROL55981. The authors have made significant revisions to the manuscript and have extensively responded to my previous comments. I appreciate their effort; however, I do not feel that the major concerns I previously raised have been adequately addressed.*

We sincerely thank the reviewer for acknowledging the efforts we have made in revising the previous version of our manuscript. In this revised version, we have refined the definition and classification of hydrological process representations, particularly by distinguishing between semi-physically-based and physically-based representations. We have also expanded the experimental results and deepened the analysis and discussion, aiming to more clearly demonstrate both the integration feasibility (i.e., the ability to combine and execute different subbasin-level model structure within a unified watershed model) and simulation effectiveness (i.e., the modeling performance of the spatially hybrid model structure). We hope these revisions have now adequately addressed the reviewer's concerns.

*2. One of the central issues remains the definition of "physically-based models." In their response, the authors assert that physically-based models are those that "incorporate solving algorithms simplified from physical principles and adhere to mass and energy balance," whose "many parameters (especially terrain and soil attributes) are derived from field measurements and have clear physical meanings." By this definition, however, virtually all computational models today could be classified as physically-based.*
*The authors cite TOPMODEL as an example of a physically-based model. While this may have been true when TOPMODEL was introduced 45 years ago, the understanding of "physically-based" modeling has evolved. What was once state-of-the-art is now often regarded as a lumped conceptual model. Beven et al. (2021) in their paper, The history of TOPMODEL, describe it as "a structurally, parametrically, and computationally parsimonious model which gave it advantages over the full implementation of the physically-based model blueprint set out by Freeze and Harlan (1969)." This suggests that even the original authors, Drs. Beven and Kirby, might not categorize TOPMODEL as a physically-based model by today's standards.*
*I acknowledge that definitions of "physically-based models" can vary, and while I respectfully disagree with the authors' interpretation, I will defer to the editor's judgment on this matter.*

We sincerely thank the reviewer for the critical insights regarding the definition of physically-based model. We acknowledge that our previous definition was too broad. We agree that the understanding and classification of hydrological models have involved over time, and that a precise and consistent definition is essential for clarity. In the revised manuscript, we have substantial revised the third paragraph of the Introduction to provide a clearer classification of hydrological process representations. We now distinguish three categories:

- **Fully-physically-based models** (e.g., MIKE SHE) apply fundamental physical laws to simulate the conservation of mass, momentum, and energy in both vertical and lateral directions. This is commonly achieved by solving coupled partial differential equations, such as the Richards' equation for unsaturated flow and Saint-Venant equations for surface flow.
- **Semi-physically-based models** (e.g., TOPMODEL, SWAT, and RHESSys) employ simplified physical equations for individual hydrological processes that are generally simulated independently, avoiding full numerical coupling. Although empirical components may be included, most model parameters retain clear physical interpretations and often can be estimated from field observations.
- **Conceptual models** (e.g., HBV and GR4J) simplify hydrological processes within a watershed into a few lumped components, typically including water storage, water losses, and flow routing, and are often implemented using conceptual reservoirs (or buckets).

Regarding to the specific example of TOPMODEL, we fully acknowledge the reviewer's point and the updated interpretation by Beven et al. (2021). Accordingly, in the revised manuscript, we now classify TOPMODEL as a semi-distributed and semi-physically-based model. We recognize that its classification may remain debatable. To avoid distracting from the core contribution of this manuscript, we are open to remove this example if preferred by the reviewer.

We hope that the revised definition and classification better align with current understanding in the hydrological modeling community and adequately address the reviewer's concerns.

*3. The authors' claim that the hybrid model "demonstrated an overall better performance" compared to the physically-based and conceptual models is also questionable. While the hybrid model performed better during the calibration period, its performance during the validation period was not superior to the "physically-based" model. Specifically, the hybrid model and the "physically-based" model yielded the same NSE and RSR values in the validation period, while the hybrid model exhibited a much higher bias, despite a slightly better R2. I would like to note that R2 is widely recognized as an unreliable metric for assessing model-data agreement in discharge simulations.*

We appreciate the reviewer's insightful comments regarding the performance evaluation of the spatially hybrid model structure.

We agree that the $R^2$ is not a reliable indicator to evaluate the agreement of simulated data and observed data, as it primarily reflects the correlation and can be misleading in hydrological analysis. Accordingly, we have removed $R^2$ from the revised manuscript.

We also acknowledge that the previous claim of "overall better performance" was too broad and did not adequately reflect the trade-offs observed in the PBIAS. In the revised manuscript, we have refined this statement to: "Spatially hybrid model improved general performance but also inherited limitations" (e.g., see revised Highlight 5). The results and discussion of the spatially hybrid model structure (HybM) were revised in the third paragraph in the section "3.3.1 Model performance metrics of calibrated model structures": "The spatially hybrid model structure (HybM) demonstrated a balanced model performance by leveraging the strengths of both PhyM and ConM, achieving the highest NSE (0.72/0.60) and lowest RSR (0.53/0.63) among all experimental cases. However, its PBIAS (10.02%/-18.28%) was lower than (i.e., better overall bias) that of both ConM cases but higher than (i.e., worse overall bias) both PhyM cases. This reflects that the spatially hybrid model structure provides a promising compromise between modeling flexibility and simulation accuracy. That is, while HybM combined the calibration responsiveness of ConM and the process realism of PhyM, it also inherited some of their respective limitations, such as bias associated with the conceptual representation."

In addition, we recognize that the primary aim of the current case study is to demonstrate the integration feasibility and simulation effectiveness of the proposed approach under a simplified application context. To better assess the robustness, generalizability, and uncertainty implications of the proposed approach, further studies in more diverse watersheds (e.g., with varying climatic, topographic, and data conditions) are warranted. We have added this discussion in section "4.3 Limitations and future directions".

*4. An intriguing observation is the disparity in performance between the calibration and validation periods for the hybrid model. The "physically-based" model displayed consistent performance across both periods, whereas the hybrid model's performance noticeably degraded during validation. This raises the possibility that the hybrid model may have been overfit during calibration. The hybrid model's inclusion of both physically-based and conceptual components likely increases its degrees of freedom, enabling it to better match observations during calibration. However, this flexibility might result in overfitting, as evidenced by its diminished performance during validation.*

We appreciate the reviewer's insightful observation regarding the disparity in model performance between the calibration and validation periods for the hybrid model. We revised the interpretation of the experimental results to emphasize that the spatially hybrid model structure provides a promising compromise between integration flexibility and simulation accuracy, that is, while HybM combined the calibration responsiveness of ConM and the process realism of PhyM, it also inherited some of their respective limitations, such as bias associated with the conceptual representation.

To clarify this issue, we substantially revised the experimental results and discussion in section 3.3.1. As noted, performance degradation across periods is also observed in the lumped conceptual models (ConM). The ConM using the universal calibration strategy (ConM1) has six parameters and showed a significant drop in NSE (0.41 to 0.32), an increase in RSR (0.77 to 0.82), and a sharp rise in PBIAS (2.89% to -42.87%). In contrast, the ConM using regional calibration strategy (ConM2) has 12 parameters (six for each type of region) and achieved more consistent NSE (0.45 to 0.50) and RSR (0.74 to 0.71) across periods, but still exhibited a notable increase in PBIAS (10.96 to -21.03). Conceptual models like ConM, with simple model structure and few parameters, are often considered easier to calibrate and potentially perform better under regional calibration. However, in this study, both ConM1 and ConM2 exhibited relatively low NSE and high RSR values in both periods, indicating that the GR4J-based conceptual model structure has limited ability to represent the hydrological processes in this watershed. The noticeable decline in validation performance for both ConM1 and ConM2 (e.g., the substantial shift in PBIAS from overestimation to underestimation) further reflects the limited generalization of ConM. Therefore, we believe that the observed performances of ConM and HybM more likely reflect the structural limitations of the GR4J-based conceptual model in representing the hydrological processes of this study area, rather than providing clear evidence of overfitting or underfitting.

The PhyM maintained stable performance and lower bias between calibration and validation periods. This suggests that PhyM captures spatial heterogeneity directly through spatially distributed input data and parameters, implying a low dependency on calibration strategies to represent spatial variations in hydrological processes

The spatially hybrid model structure (HybM) achieved the highest NSE and lowest RSR among all experimental cases, but its PBIAS was higher than that of PhyM. This highlights that although HybM balances integration flexibility and simulation accuracy, it is still susceptible to limitations inherited from its constituent components, such as the simulation bias from ConM.

These findings also highlighted the importance of careful model structure construction and the essential needs for multi-site and long-term observation data to constrain calibration. We have expanded the experimental results and deepened the analysis and discussion based on the above analysis in section 3.3.1 of the revised manuscript.

*5. Another issue pertains to the model calibration process, which remains unclear. The authors state that "The objective function for parameter calibration is maximum NSE, minimum absolute values of RSR and PBIAS, with each component given equal weight." Does this imply that the model was calibrated three separate times using distinct objective functions? Or was a composite objective function, such as $a * NSE - b * RSR - c * PBIAS$, employed? Additional clarification would enhance the transparency of the calibration process.*

Thank you for your insightful comment regarding the clarity of the model calibration process. We apologize for the ambiguity in the previous description. To clarify, three performance metrics, i.e., NSE, RSR, and the absolute value of PBIAS, were treated as independent objective functions within the NSGA-II algorithm. The model was not calibrated three times with distinct single objective functions, nor did we use a composite objective function (e.g., $a * \mathrm{NSE} - b * \mathrm{RSR} - c * \mathrm{PBIAS}$). Instead, we applied the NSGA-II algorithm to optimize these three objectives simultaneously to identify a Pareto front of non-dominated solutions.

The previous phrase “each component having the same weight” was meant to indicate that all three metrics were considered equally important during optimization. In the context of NSGA-II, this equal importance is typically reflected in mechanisms such as the calculation of crowding distance, which ensures a well-distributed set of solutions along the Pareto front, rather than through explicit weighting in a single aggregated function.

We have revised the manuscript to clarify the calibration process as follows: "The multi-objective optimization aimed to maximize NSE and minimize RSR and absolute value of PBIAS."

*6. In conclusion, the modeling framework proposed by the authors is intriguing and offers the flexibility to support models of different categories and spatial discretizations within the same simulation. However, based on the results presented in the manuscript, I do not find sufficient evidence to conclude that the hybrid model/approach surpasses the "physically-based" model. Consequently, I feel that the authors' conclusions are not strongly supported by their results.*

We sincerely appreciate the reviewer's approval of the innovation of our manuscript and the constructive evaluation. As responded to the earlier comments of the reviewer, we have substantially revised the experimental results and discussion to clarify that the spatially hybrid model structure represents a trade-off between integration flexibility and simulation accuracy, rather than surpassing the physically-based model structure.

More importantly, the core innovation of our study is to propose a flexible modeling approach that enables spatially hybrid hydrological modeling - a capability that is notably limited in existing modeling frameworks.

Furthermore, we have added section "4.3 Limitations and future directions" to explicitly discuss key limitations and outline future research directions, including: (1) expanding the SEIMS module library, (2) improving inter-subbasin connectivity mechanisms under the subbasin-independent architecture, (3) developing domain knowledge-driven methods that incorporate various knowledge types to reduce the substantial modeling burdens on users (particularly non-expert users) when constructing appropriate model structures for specific application contexts, and (4) conducting additional case studies in diverse watershed application contexts to comprehensively evaluate the robustness, generalizability, and uncertainty of the proposed approach.

We hope that these revisions better align the conclusion with the presented results and adequately addresses the reviewer's concerns.

**With regards to comments from the Reviewer #2**

*1.This manuscript proposes an extension for the Spatially Explicit Integrated Modeling System (SEIMS). While it is well-structured study. But I doubted it as a "novel" method as stated by the authors. In hydrological modelling, based on Beven (2012) identifying the goal of modelling and understanding the situation and the process of our watershed is crucial (as perceptual model). And this is the reason that there are some flexible modelling structure frameworks, as discussed in the manuscript.*

We sincerely thank the reviewer for the thoughtful comment and for referencing Beven (2012), which highlights the importance of perceptual models and flexibility in hydrological modeling. We fully agree that perceptual understanding is crucial to model design and has inspired the development of several the existing flexible modeling frameworks, like SUPERFLEX.

While we acknowledge that flexible modeling is not a new concept, our work contributes to this concept by introducing a new and practically significant modeling approach aimed at addressing specific limitations in state-of-the-art hydrological modeling frameworks. Specifically, existing frameworks like SUPERFLEX primarily focus on enabling spatially varying configurations of lumped conceptual model structures within a watershed. In contrast, our proposed approach advances the spatial flexibility by allowing the combination of compatible simulation units (lumped, semi-distributed, and fully distributed) with multiple types of simulation algorithms (e.g., conceptual and semi-physically-based in the current implementation), and then integrating subbasin-level model structure into a single watershed model.

The novelty lies not in the general idea of flexibility itself, but in the specific dimensions of flexibility and their combined implementation within a modeling framework. To better reflect this positioning, we now describe the proposed approach as “a new and practically significant spatially hybrid modeling approach” in several places in the revised manuscript, such as the last paragraph of the Introduction section.

We hope this clarification clearly articulates the contribution of our work and how it fits into the broader landscape of flexible hydrological modeling.

*2.The authors mentioned two limitations for their study (in the conclusion section). I think this extension can have more limitations especially in adding much uncertainty from conceptual model structure part to the physical model part and the importance of connectivity of hydrological components also needs to be discussed, maybe in future applications.*
*I agree that this idea can provide more flexibility in modelling, but how can we balance this flexibility with other new arise issues including increasing various sources of uncertainty and finding optimal spatial constitutions for new case studies.*

We appreciate the reviewer’s insightful suggestions regarding potential limitations and challenges of the proposed approach. We agree that increased model flexibility may introduce additional uncertainty sources. As responded to the Reviewer 1, we have refined the experimental results to acknowledge this trade-off: “Spatially hybrid model improved general performance but also inherited limitations." This statement emphasizes that the spatially hybrid model structure provides a promising compromise between integration flexibility and simulation accuracy.

We also agree with the reviewer that careful design of model structure is crucial to better apply the proposed approach. We have added a paragraph to discuss the future research on developing domain knowledge-driven methods to construct appropriate model structures for specific application contexts (see the fourth paragraph of section 4.3 in the revised manuscript).

Regarding the connectivity of hydrological components, in the current implementation of the proposed approach based on SEIMS, we simplified the hydrological connectivity between subbasin-level models by assuming that mass and energy exchange between subbasins occurs only at the outlets through channel network (as described in the last paragraph of section 2.1). This assumption may not fully reflect complex hydrological interactions and representations. Therefore, more flexible connectivity mechanisms could be proposed, such as accounting for interactions between basic simulation units near subbasin boundaries.

Overall, to address these valuable points and enhance the transparency of our work, we have added section "4.3 Limitations and future directions" in the revised manuscript. This section explicitly discusses key limitations and outlines future research directions, including: (1) expanding the SEIMS module library, (2) improving inter-subbasin connectivity mechanisms under the subbasin-independent architecture, (3) developing domain knowledge-driven methods that incorporate various knowledge types to reduce the substantial modeling burdens on users (particularly non-expert users) when constructing appropriate model structures for specific application contexts, and (4) conducting additional case studies in diverse watershed application contexts to comprehensively evaluate the robustness, generalizability, and uncertainty of the proposed approach.

*3.Beven (2012). Rainfall-Runoff Modelling: The Primer. (In one of the chapters, it introduces the 5 steps (or the modelling cycle) in hydrological model development. The first step is called as perceptual model developments, which is some time neglected in hydrological model studies)*

We sincerely appreciate this insightful reference to Beven's (2012) modeling cycle and the crucial role of developing perceptual models in hydrological studies. We fully agree that perceptual understanding of the application contexts of hydrological modeling (e.g., watershed characteristics) is essential for guiding the selection or construction the appropriate model structure.

In fact, the proposed approach is inspired by this idea, that is, representing spatial heterogeneity in watershed processes through differentiated model structures across subbasins and integrating them into a unified watershed model. We have also emphasized that users are responsible for ensuring the rationality of the constructed model structure based on their hydrological knowledge and the specific application context (e.g., as noted in the last sentence of the third paragraph in section 2.1 and the last sentence of the second paragraph in section 4.2). This principle aligns with the idea of perceptual modeling development.

We thank the reviewer for highlighting this important aspect of hydrological modeling.

**With regards to comments from the Reviewer #3**

*This manuscript presents a novel spatially hybrid hydrological modeling approach, combining lumped conceptual and distributed physically-based models to address the challenges posed by spatial heterogeneity in watersheds. The authors implemented their method using the SEIMS framework and validated it through a case study of the Heihe River Basin, highlighting the model's ability to outperform consistent modeling approaches. In general, the paper makes a significant contribution to hydrological modeling by introducing a flexible and scalable framework for hybrid modeling.*

*As the additional reviewer for the second round, I find that the authors have thoroughly and adequately addressed the comments raised by the reviewers, including those that I would have highlighted in my own review process. Therefore, I would like to recommend acceptance of the manuscript, subject to a few minor revisions and suggestions.*

We sincerely appreciate the reviewer's positive evaluation of our work.

*Suggestions:*

*1. Please consider adding a discussion on the potential of adapting the spatially hybrid hydrological modeling approach to advance integrated modeling, particularly given its ability to adapt and incorporate existing structure-varying models for large-scale hydrological applications. Its design enables the reuse of pre-existing hydrological models, facilitating their seamless integration into a unified framework. This adaptability enhances the ability to address large-scale hydrological challenges by leveraging prior modeling efforts, thereby improving efficiency and ensuring consistency across various scales and regions.*

We sincerely appreciate the reviewer's thoughtful suggestion regarding the potential of the spatially hybrid modeling approach to support integrated modeling applications and advance large-scale hydrological modeling.

We have expanded the related discussion in the revised manuscript by adding a new section "4.2 Advantages of SEIMS in implementing the proposed approach" to emphasize this aspect. Specifically, we highlighted that the SEIMS framework supports module-level reuse and integration of existing models through its modular and subbasin-independent architecture. This design facilitates the construction of spatially hybrid watershed models by assembling existing SEIMS modules and newly added modules. Such flexibility provides a promising groundwork for future integrated modeling studies across diverse hydrological application contexts.

We thank the reviewer for highlighting this important application direction, which we believe will significantly enhance the practical relevance and broader impact of the proposed approach.

*2. Consider adding an explanation or some examples to the fourth paragraph to show why fixed model structures have trouble with diverse and complex applications.*

We have enhanced the fourth paragraph of the Introduction with a specific example: "For example, models that rely solely on the infiltration-excess runoff mechanism are unsuitable for watersheds where saturation-excess runoff dominates."

*Minor comments:*

*1. Change "primarily" to "usually" and "may" to "can" in line 79.*

We have modified the original sentence "Parameters used in such physically-based models primarily have clear physical meanings and may be derived from field measurements" to "Although empirical components may be included, most model parameters retain clear physical interpretations and often can be estimated from field observations."

*2. It is not clear what categories the listed models belong to in line 89-95. Please rewrite this sentence to link models to their corresponding categories.*

We have revised this paragraph to explicitly categorize each example model to its corresponding category: "They vary from lumped to distributed spatial discretization schemes, and from conceptual to fully-physically-based representation of hydrological processes. Based on the most typical configuration or usage of each model, examples include: the lumped and conceptual HBV (Hydrologiska Byråns Vattenbalansavdelning; Lindström et al., 1997), GR4J (modèle du Génie Rural à 4 paramètres Journalier; Perrin et al., 2003); the semi-distributed and semi-physically-based TOPMODEL (TOPography based hydrological MODEL; Beven and Kirkby, 1979), SWAT, and RHESSys (Regional Hydro-Ecologic Simulation System; Tague and Band, 2004); the fully-distributed and semi-physically-based DHSVM (Distributed Hydrology Soil Vegetation Model; Wigmosta et al., 1994); and the fully-distributed and fully-physically-based MIKE SHE."

*3. Change "was" to "is" in line 134.*

We have modified this sentence to use the present tense: "To further enhance the flexibility of spatially varying model structures, another idea involves internally constructing and integrating multiple model structures within a single modular hydrological modeling framework such as FLEX-Topo."