Enhancing physically based and distributed hydrological model calibration through internal state variable constraints

Frédéric Talbot[1], Jean-Daniel Sylvain[2], Guillaume Drolet[2], Annie Poulin[1], Richard Arsenault[1]

[1] Hydrology, Climate and Climate Change Laboratory, École de technologie supérieure, Université du Québec, Montréal, H3C 1K3, Canada

[2] Direction de la recherche forestière, Ministère des Ressources naturelles et des Forêts, Québec, G1P 3W8, Canada

Correspondence to: Frédéric Talbot (frederic.talbot.2@ens.etsmtl.ca)

RC1 (https://doi.org/10.5194/egusphere-2024-3353-RC1)

Review of "Enhancing Physically Based and Distributed Hydrological Model Calibration through Internal State Variable Constraints"

General Comments

The authors have presented a comprehensive study that investigates the impact of different calibration approaches on hydrological models using the WaSiM model. The paper explores three distinct configurations: Baseline (BL), Physical Groundwater Model (GW), and Physical Groundwater with Recharge Calibration (GW-RC) to evaluate their effectiveness in representing various hydrological variables. This research addresses an important topic in hydrological modeling by highlighting the significance of integrating internal state variables into the calibration process.

The paper is well-structured, and the authors have made a significant effort to present detailed analyses across multiple catchments. The inclusion of groundwater recharge as a calibration variable is an important approach that aligns with the growing need for multi-variable calibration frameworks in hydrological modeling. The findings makes effort to underscore the importance of considering both streamflow and internal hydrological processes for robust model performance.

We thank the reviewer for the thoughtful comments and suggestions regarding our manuscript. The feedback will certainly help improve the clarity and impact of our research.

However, I have a major concern regarding the primary objective of the study, which requires clarification. The current presentation leaves the reader uncertain about whether the study aims to compare calibration strategies or assess the impact of model complexity on hydrological process representation. Addressing this ambiguity will make clear the paper's overall contribution and impact.

We acknowledge the importance of distinctly defining our research focus. To address the highlighted ambiguity, we propose to revise the text at L90-95 of the introduction to explicitly clarify that the primary aim of our study is a hybrid approach, aiming to improve the representation of hydrological variables through various model complexities (BL and GW) and constrained calibration focusing on recharge (GW and GW-RC).

Major Comments

1. Unclear Research Focus: The primary research question of the paper is not clearly defined. It remains ambiguous whether the authors aim to compare calibration strategies or demonstrate the added value of increasing model complexity.

- If the goal is to compare calibration strategies, the authors should focus on showing how the constrained recharge parameter improves the realism of the results when compared to both streamflow and PACES data.
- If the goal is to assess model complexity, the paper should clearly outline what unique complexities are introduced in each configuration and how they enhance the model's capability to represent hydrological processes.

We indeed addressed both aspects: comparing calibration strategies and assessing model complexity. Specifically, the manuscript compares the calibration strategies by contrasting the GW and GW-RC configurations and evaluates the model complexity by comparing the BL and GW setups. We acknowledge that this dual focus may not have been articulated clearly enough in the initial draft.

To rectify this, as stated above, we will revise the text at L90-95 of the introduction to clearly define and differentiate these two intertwined aims. This clarification will ensure that readers fully understand the scope and the dual objectives of our study.

2. Simplifying the Experimental Setup: The current experimental design includes three configurations (BL, GW, GW-RC), but most of the observed differences in results seem to be attributed to the choice of model complexity rather than calibration strategies.

- To demonstrate the impact of the constrained recharge parameter, the authors could simplify their experimental setup by comparing two GW-RC experiments: one calibrated solely to streamflow using the KGE metric and another using a modified objective function that accounts for both the mean and variability of recharge. This would directly illustrate the benefit of including internal state variables in the calibration process.

We understand the suggestion to simplify the experimental setup to more distinctly demonstrate the impact of the constrained recharge parameter. However, our study design, which includes the GW and GW-RC configurations, is specifically crafted to assess the effects of introducing recharge in addition to streamflow in model calibration.

To clarify, both GW and GW-RC configurations operate under the same model complexity, with GW calibrated on streamflow only and GW-RC utilizing both streamflow and recharge during calibration. This setup is intended to explicitly isolate and compare the impact of including recharge alongside streamflow in the calibration process.

In response, we will revise the manuscript to ensure this distinction is clear. We propose adding the following sentence after the first sentence of section 2.4.3 where the configuration GW-RC is introduced. "Importantly, GW-RC uses the same model complexity as GW, with the goal of isolating the effect of adding groundwater recharge in calibration."

3. Clarifying the Role of GW-RC: The GW-RC configuration is described as integrating recharge calibration into the model. However, the results suggest that most of the observed improvements are due to the activation of more physically based processes rather than the calibration strategy itself. If the authors wish to emphasize the importance of incorporating internal state variables, they should isolate and highlight the specific impact of the recharge constraint.

We acknowledge the importance of clearly distinguishing the impacts attributable to the introduction of recharge in the calibration process.

In the manuscript, the GW and GW-RC configurations operate under the same model complexity, ensuring that any observed improvements can be associated with the calibration strategy rather than model complexity enhancements. The GW configuration is calibrated solely on streamflow metrics, while the GW-RC setup integrates recharge, allowing us to directly compare the effects of adding recharge calibration on model performance.

To address your feedback, as stated in comment no. 2, we propose to add a sentence in the beginning of section 2.4.3.

Specific Suggestions

1. **Objective Statement**: In the introduction, clearly state whether the study aims to evaluate calibration strategies or model complexity.

   As stated above, we propose to revise the text at L90-95 of the introduction to clearly define and differentiate these two intertwined aims.

2. **Experimental Design**: Consider restructuring the experimental setup to compare GW-RC configurations with and without recharge constraints. This would make the study's focus more precise.

   We appreciate the constructive feedback and would like to clarify some points. As stated above, our experimental design already includes the GW and GW-RC configurations, which are fundamentally identical in terms of model complexity but differ in their calibration targets. The GW configuration is calibrated solely using streamflow data, whereas the GW-RC configuration additionally incorporates recharge data into the calibration process. This setup precisely allows us to isolate and compare the effects of integrating recharge in calibration. In response to your suggestion, we will enhance the clarity of this aspect in the methodology section. As stated in comment no. 2, we propose to add a sentence in the beginning of section 2.4.3 to address this concern.

3. **Discussion Section**: Emphasize the role of model complexity in the observed differences in results. If the paper aims to introduce new complexities in the model, demonstrate their unique contribution to the model's performance.

   We appreciate the input on the need to better highlighting the distinct contributions of each model configuration to the observed results. To address the comment, we will revise the text at L286-293 to better articulate how the baseline (BL) configuration, utilizing a conceptual method for groundwater flow, contrasts with the GW configuration which employs a physically based approach to groundwater dynamics. This will ensure that model complexity introduced by the GW configuration will be better understand by the reader

4. **Results Interpretation**: Clearly distinguish between the impact of calibration strategies and model complexity in the results section. For example, highlight how much of the improvement in GW-RC is due to the activation of groundwater dynamics versus the recharge calibration constraint.

We acknowledge this concern. However, we have already addressed this distinction in our article. Our study was designed to differentiate between model complexity and calibration methods using three configurations. When comparing configuration BL and GW, the differences are due to model complexity, as the BL configuration uses a conceptual method for groundwater

flow, while GW employs physically based equations. Conversely, when comparing configuration GW and GW-RC, the improvements are due to the integration of groundwater recharge into the calibration process, not model complexity. Both GW and GW-RC configurations employ the same level of physical process modelling, ensuring that any performance differences are due to calibration strategies alone.

Conclusion

The paper presents valuable insights into the role of multi-variable calibration in hydrological modeling. However, the main research target needs to be more clearly defined. The authors should clarify whether their focus is on calibration strategy comparison or model complexity assessment. Additionally, simplifying the experimental setup to directly compare the impact of recharge constraints would make the study more impactful.

We believe the proposed changes will address the concerns and enhance the manuscript. We thank the reviewer once again for the constructive comments.

Best regards,

Frédéric Talbot on behalf of all authors