Climate Impacts on Water Resources in a High Mountain Catchment: Application of the Open-Source Modeling Workflow MATILDA in the Northern Tian Shan (EGUSPHERE-2025-3462)

Final Response to Reviewer Comments RC1

Opening Remarks

Dear reviewer,

Thank you very much for investing the time and effort to review our manuscript. We are confident that the work will benefit significantly from the suggested revisions.

Before we respond to your comments in detail, we would like to make a few general remarks. Our study is designed as a double publication: Part 1 comprises the model description, sensitivity analysis, code and documentation, and has been submitted to Geoscientific Model Development (GMD). Part 2 is a case study with a calibration approach (HESS). We opted for this approach to ensure the model setup and technical implementation including code and documentation were properly reviewed, while also seeking feedback from the hydrological community on the model's application. The editorial support team suggested the procedure to mark both submissions as companion papers and that they would coordinate the review process. However, due to a lack of topic editors, GMD took five months to start the review process, and the public discussion began just after you submitted your review for part 2. It appears that there was no coordination in terms of aligning the review phases of the two manuscripts.

Nevertheless, we acknowledge the importance of both parts to stand for themselves, and we are grateful for your comments. Please see our detailed suggestions addressing your concerns below. Your original comments are underlaid in blue.

1) *The biggest issue amongst the others is that Part I of the double publication is still under review in another journal. The authors did not introduce the model components in this manuscript at all, sometimes not even expanding abbreviations. Thus, I can't fully assess the quality of the model chain and its ability to provide accurate information to the stakeholders, as mentioned by the authors as the aim of the model in the Introduction.*

The model description manuscript entered the public discussion Dec 15. We have contacted the editor and editorial support to seek for better coordination going forward.

We will add a paragraph briefly summarizing the modeling chain at the beginning of the Methods section. Most of the components are well established and all required sources referenced in Part 1 will be added here as well. A direct reference to Fig. 1 in Part 1 that illustrates the models core routines will be added. Abbreviations including parameter names will be introduced carefully. Long names of all model parameters will be added to Table 2.

2) *The authors discuss the societal aspects and the value that MATILDA can offer to local stakeholders in the Introduction section. However, the remaining sections of the manuscript focus solely on the calibration procedure and the model's limitations. The objectives of this manuscript are not clear to me.*

Many regions most affected by climate change rarely use physically based modeling in water resource management and policy. That is the key problem that motivated the development of MATILDA. The idea is to address this with a cloud-based tool with comprehensive documentation in the form of an interactive digital book (Matilda-Online). The book is designed in style and language as educational material for a non-scientific audience and was already used for summer schools in Central Asia. An additional manual addresses technical questions for more advanced users.

The research papers however target a scientific audience. So while the societal aspects are the general motivation of the project, the objective of this part of the study is to demonstrate the tool's capabilities to assess actual climate impacts at the local scale, describe the calibration approach currently used in Matilda-Online and discuss the uncertainties.

However, you are absolutely correct to question the storyline. Therefore, we suggest shifting the focus away from the technical demonstration of a model towards the actual questions in the study area. Many of the subsequent responses to your comments will adhere to this goal.

In the introduction, we will move lines 16–20 to the end of the section and reframe it to focus more on the regional challenges and the climate's impact on the water balance and less on the model. We will also provide a brief overview of the specific study site and its particular challenges.

3) *I don't quite understand why the Kyzylsuu valley was chosen as the study site from what is written in Sect. 2. Is it because there are not enough monitoring stations? Does it have anything to do with the risk assessment needed by the local stakeholders? This is also linked to the issue of the study's objective being unclear.*

The following information will be added to Section 2 (or pointed out more clearly):

- MATILDA was developed through a Kyrgyz-German research collaboration, which is why a site in Kyrgyzstan was chosen.

- Long-term data is sparse. Few monitoring stations survived the 1990s, and water-related data is a politically sensitive issue in the region.

- Kyzylsuu is a mesoscale catchment that supplies a small community that relies primarily on agriculture and tourism. Therefore, it represents MATILDA's target areas while offering the rare combination of long hydrological and glaciological records, and a weather station. The gauging station and mass balance record both provide the minimum required data coverage after 2000, a critical period in glaciological terms. Although MATILDA is designed to address data scarcity, its effectiveness can only be evaluated if some in-situ data is available.

- At least one glacio-hydrological projection study has been carried out in the valley, enabling comparison.

4) *I don't understand why the authors use ERA5-Land reanalysis data for calibration. In Sect. 4.1, it is stated that ERA5-Land overestimates precipitation by 108+-62%. No bias correction nor data assimilation was implemented to reduce the overestimation. At this point, I stopped reading the rest of the sections in more detail because we can't be sure the calibration procedure is not over-tuning the parameters. To the best of my knowledge, there are statistically downscaled and bias-corrected ERA5-Land products. Why don't the authors use those products? At least, it has to be justified in the manuscript even if the authors don't want to rerun all the simulations.*

Most of the following context is provided in Part 1, but it will also be added briefly at the beginning of the Methods section, or referenced more clearly.:

- The current study is a case study using a model setup intended for any given study site. One objective is therefore to demonstrate that this setup, which provides global coverage, can be effective for a local water balance study. As MATILDA is

cloud-based, all implemented datasets must be accessible and subsettable online, ensuring minimal downloads via an API or Google Earth Engine.
- Of the datasets that fulfil these criteria, ERA5-Land has the best spatial and temporal coverage and is used for many of the glaciological modeling studies in High Mountain Asia.

  (Note: To the best of our knowledge, CHELSA v2.1 is currently the only comparable dataset with a higher spatial resolution available for Central Asia, but the data platform was under reconstruction until recently and the data has not yet been tested.)

The following information is mentioned in the text, but it will be emphasised by restructuring the Methods and Results sections, as described in the comments below:

- From 3.1.2:
  - The standard HBV model incorporates a simple bias adjustment using correction factors for precipitation, snowfall and evapotranspiration, as well as elevation-dependent lapse rates for temperature and precipitation. Therefore, we do not apply any further bias adjustments during pre-processing for the calibration period.
  - To avoid over-tuning, the precipitation factor is statistically constrained during the first calibration step. The snowfall correction factor is disabled, as the snow routine is calibrated separately in step 2.
- From 2. and 4.1:
  - The precipitation comparison is based solely on nine years of observations in the valley centre and the ERA5-Land grid cell containing the station. Since the exact equipment (shielding, heating, etc.) is unclear, we only used summer precipitation (April–September). This only allows for a rough estimate of the extent to which the overestimation may occur and was not used directly for pre-processing.

5) *The language of the manuscript is highly technical; the article lacks sufficient structure. I often need to go back and forth to find out why certain sections are placed in their current location. If its target readership is local stakeholders, it does not serve its purpose.*

The model's website contains a complete walkthrough designed as educational material. However, the manuscript targets a scientific audience who are familiar with glacio-hydrological modelling. Technical details have been included to enable in-depth discussion of the calibration approach and ensure complete reproducibility.

Nevertheless, we agree that the article would benefit from a clearer structure and less detail in the text. Therefore, we will streamline the 'Methods', 'Results' and 'Discussion' sections, moving most of the technical calibration details to Table 2. The following suggestion addresses several of your comments and is outlined here for better readability. Although the number of subsections will be reduced in the process, the original numbering is referenced. In detail:

Methods:

- A short paragraph titled “Forcing Data” inserted before the Calibration section (3.1) will provide:
  - the context described for comment 4
  - A brief description of the the climate scenario data and the bias adjustment
- The separate methods (3.1.2) and result parts (4.2) for every calibration step will be merged into one consistent calibration walkthrough in the Methods section.

- Values already included in Table 2 that are not essential for the workflow description such as filter criteria will be removed from the text.
- Non-essential values not yet included such as sample sizes will be moved to Table 2.
- The section on validation (3.2) will be merged with the calibration procedure.
- The section on uncertainty analysis (3.3) will both be merged with the respective paragraphs in the Results (4.3 and 4.5). , shortened, and moved to the Discussion.

  → The Methods section will therefore consist of a brief model intro followed by the sections "Forcing data", "Calibration data", "Calibration procedure", and "Uncertainty analysis" while all non-essential details will be provided in Table 2.

Results:

- As described above:
  - 4.1 "Reanalysis data" → Methods
  - 4.2 "Calibration" → Methods
  - 4.3 "Validation" → merged with "Calibration" (Methods)
  - 4.4.1 "Bias Adjustment" → Methods
- 4.5 "Uncertainty" will be moved to the Discussion
- 4.4.2 and 4.4.3 will be slightly extended to cover the recent climate trends as well as more climate impact indicators, especially regarding the flow regime and droughts (on request of RC2)
- The first paragraph of the result comparison to Chevallier et.al. (5.2.3) will be added here to directly contrast the outcomes

  → The Results will focus solely on the recent and projected climate trends and impacts on the study site's water balance. All intermediate results will be moved to other sections.

Discussion:

- The general structure will be adjusted to better link uncertainties with modeled results and focus on the robustness of the assessment:
  1. Deficiencies in ERA5-Land (see comment 4), the implications for the calibration, and how the shortcomings in the precipitation product can be addressed in local studies (e.g. using high-pass filters, orographic precipitation models, and station data if available)
  2. Deficiencies in the NEX-GDDP-CMIP6 dataset and the implications for the projections
  3. Uncertainties from the calibration data and their implications for century-long projections
  4. Parameter uncertainty and how it is addressed in the calibration strategy
  5. The use of a conceptual model for century-long projections especially regarding parameter stationarity
  6. Robustness of the results and use cases in the light of the discussed uncertainties

     → Uncertainty estimates (4.5), validation procedures (5.2.1), and study comparison (5.2.3) will be integrated where they add values rather than in a separate section

6) *Thus, I don’t think this manuscript, in its current form, is suitable for publication even after a major revision because we don't know if its 'sister' paper will be accepted at all.*

According to the editorial support, the review process can not be paused but the publication can be coordinated. While the primary goal is to let the revised manuscript stand for itself, we will do what we can to align the review processes.

7) *I see the value in the study, but perhaps the author could put more effort into the storytelling and the accessibility of the language to non-academic experts, if you really want MATILDA to fulfill its purpose.*

See comment 5.

8) *I suggest that the authors revise the manuscript carefully and resubmit it after the first manuscript is accepted.*

See comment 6.

Section 3 Methods

9) *There is no description of the toolkit and its model components. Some abbreviations are not even expanded. I understand that there is a 'sister' manuscript that provides additional details. However, at least some key elements should be included here so that readers don’t need to read another article to grasp what MATILDA does.*

See opening remarks and comment 1.

10) *In addition, the sister manuscript is still under review in another journal. It means that as a reviewer, I need to review two manuscripts to ensure the accuracy of my judgments.*

We are grateful for the effort and time you invest in this review. While we will improve the manuscript to be read independently, the key element of a double publication is the synergy of both studies. Therefore, we ask for your understanding that a brief look in the companion manuscript will be necessary.

Section 4 Results

11) *The structure of this section is very confusing. ERA5-Land is also presented here, even though the authors did not produce the forcing data.*

See comment 5.

12) *Sec. 4.2 basically only summarizes Table 2. What is the message here?*

See comment 5.

13) *A lot of abbreviations are used, which made it very hard to follow what the authors wanted to convey, especially for non-academic readers, whom they addressed in their introduction. Could we consider starting the sections or paragraphs with a sentence that describes their content or key messages?*

See comments 1 and 2.

Section 5 Discussions

14) *Shouldn’t Sect 5.2.3 be in Sect. 4?*

See comment 5.

15) *And it will also be helpful to have some discussion on how MALTILDA is a better tool for the local stakeholders than the other glacio-hydrological tools, as it seems to be the "innovation" presented in the Introduction.*

The following will be pointed out at the end of the Discussion as closing remarks:

- The benefit of MATILDA is that its basic use has comparably little prerequisites: some basic understanding of hydrological processes, a simple computer with internet connection, proficiency in English, and the motivation to acquire new skills.
- With a located discharge time series as only user input, it can provide glacio-hydrological projections and a climate impact assessment for any given location at a comparably fine spatial scale. It also guides through the entire modeling chain as in an interactive tutorial.
- These benefits, namely accessibility and transferability, come at the cost of uncertainties. These can be reduced by including additional datasets and fine tuning the calibration. However, these measures require a deeper technical understanding and have higher computational costs.
- The high number of model runs required for a comprehensive calibration and the most robust projections limits the applicability in the current version. Therefore, one of the main goals of future versions is to provide pre-calibrated parameter sets for focus regions.

16) *From what I can see on the MATILDA website, it appears to be a model chain that combines a PDD model, a glacier volume/Area rescaling routine, and the openly assessable HBV model, combined with a significant calibration component.*

That is correct.

17) *Perhaps it is discussed in the first paper. However, without reviewing the first paper, I would not know if the model chain is really novel. In particular, from what I have read in the current discussion, MATILDA faces the same challenges as other glacial-hydrological models, i.e., forcing data and the modeling workflow.*

The following is summarized from part 1 but will be added briefly to the model description at the beginning of the Methods section (first bullet) as well as the Conclusion (last two bullets):

- The modelling chain is not novel in the sense that it hasn't yet been applied. On the contrary, we deliberately chose a well-established approach, implementing it as a largely cloud-based application and supplementing it with targeted improvements.
- The novelty lies in the combination of automated acquisition of forcing and calibration data with an open-source modeling chain and high-quality visual outputs in a bundle that can be used in the cloud and locally.
- Thus, despite facing similar challenges to comparable models, the tool requires substantially less effort and user-provided data. Therefore, while reducing technical barriers to common scientific workflows, it can deliver results comparable to those of other studies with less effort and data input, or even improve quality, as demonstrated in the study comparison.

Section 6 Conclusions

18) *Overall, I am not convinced that 'MATILDA offers a solution by supporting, educating, and empowering water management stakeholders in regions affected by climate change.' It is not supported by what was written in the manuscript.*

We agree to use less assertive language here, without claiming to have a 'solution' or suggesting operational use in water resource management. However, we hope that by implementing the suggested changes, we can convince you of MATILDA's ability to reduce the barriers to scientific methods for local climate impact assessments.

The conclusion section will be adapted to align with the revised storyline, focusing more on the impact of climate on the local water balance within the study catchment. The performance of the model and the suitability of the approach will be evaluated in light of local challenges. The manuscript will conclude with a rigorous assessment of what MATILDA can and cannot reliably provide.

Figures

19) ***From Figure 2****, we can see that the interannual variation of the simulated SMB does not match that of the reference study and the observations. In the conclusion, the author wrote, "The temperature index model with the modified Δh routine can provide reasonable long-term estimates of glacial contributions to runoff, including stabilizing effects at higher elevations (Schuster et al., 2025c). However, this setup fails to reproduce observed inter-annual changes and neglects important glaciological factors such as glacier dynamics and debris cover". Does this really show the robustness of the model chain? How important are glacier dynamics and debris cover compared to the ability of the model to reproduce the observations in the time period you are looking at, i.e., 20 years?*

The statement is indeed misleading since inter-annual variability is not a robust evaluation metric for a temperature-index model and the available calibration data. It will be removed and the data discussion will point out the following more clearly:

- For the calibration period there are merely 7 years of annual surface mass balance observations for only a single glacier in the catchment. As discussed in 5.1.3, the observed glacier is likely to have higher melt rates than the catchment average and the available period is short relative to typical glacier response times.
- The inter-annual mass balance series from Barandun et al. represent model-based estimates for the seven largest glaciers, constrained by remotely sensed transient snowline observations. Barandun et al. report a conservative uncertainty of ±0.37 m w.e. yr^{-1} for the absolute annual mass balance values, dominated by temporally correlated (systematic) errors. At the annual scale, this uncertainty is large relative to inter-annual differences, and individual extreme years (e.g. 2012) should therefore be interpreted with caution. The uncertainty range will be added to Figure 2 to reflect this.
- As described in part 1, the delta-H routine for glacier rescaling by Seibert et.al. cannot simulate glacier cover exceeding the initial extent. To ensure mass conservation, surplus accumulation is stored and added once glacier cover falls below the initial state. While this preserves long-term mass balance, it may affect absolute annual values, particularly during the initial years of the simulation.
- Given the combined observational uncertainty and the simplified process representation of degree-day-based glacier models, annual-scale evaluation is uncommon in comparable glaciological studies.
- Neglecting glacier dynamics and debris cover is expected to have limited influence during the calibration period but introduces additional uncertainty in long-term projections. For example, high debris cover has been shown to reduce glacier melt rates and promote the formation of mountain permafrost features that can alter flow regimes due to their longer response times and different seasonality.

Consequently…

→ MATILDA is calibrated to reproduce the long-term climatic mass balance

→ Barandun et al. provide the most regionally consistent dataset for Central Asia. Only long-term means are used for calibration.

→ The simple modeling approach has computational advantages but the uncertainty needs to be considered.