Hydrological modelling in complex-terrain areas is highly challenging due to uncertainties in model structure, parameters, and input data, further complicated by the issue of equifinality, which often obstructs accurate interpretation of hydrological processes. This study utilizes multi-source data to calibrate a semi-distributed hydrological model, thereby providing better constraints on key hydrological processes. The findings offer valuable insights and contribute to advancing the understanding of water cycle in complex terrains. Main concerns are listed as follows.

We thank the reviewer for the constructive comments and suggestions, which gave us the opportunity to strengthen our manuscript. We believe that we have addressed all concerns and incorporated all suggestions as detailed in the point-by-point response below.

1) Regarding the hydrological and calibration strategy, more details are needed. Key information, such as the inputs and outputs of the model, model integration time step, temporal resolution of reference data used for model calibration, and the number of iterations for each calibration, needs to be clearly specified.

We will further specify the requested information in the revised manuscript where necessary, as illustrated below:

- inputs and outputs of the model: the hydrological model was widely used (L166-168) and, in order not to further lengthen the main text, it is described in detail in Section S1 of the Supplement (L180-182). However, we agree that it is appropriate to make the required information explicit by inserting the following sentence in Section 2.5 at L174: "It uses precipitation, air temperature and PET (here derived from ET_0) as forcing data, and reproduces various internal states and hydrological fluxes, including outgoing streamflow and AET";
- model integration time step: as stated in L174, the model was applied at hourly time-step;
- temporal resolution of reference data used for model calibration: the sentences in L285-288 will be slightly modified to better specify this information, as reported below. "KGE is the Kling-Gupta efficiency index (Gupta et al., 2009), computed between observed and simulated streamflow <u>time series</u>, while APFB is the annual peak flow bias proposed by Mizukami et al. (2019). AET_{sim} and AET_{ref} indicate the simulated and reference <u>cumulative</u> AET volume <u>in the period</u> at basin scale, <u>respectively,</u> while <u>similarly</u> $IMV_{j,sim}$ and $IMV_{j,ref}$ are the simulated and reference glacier-sourced meltwater volume for the *j*-th sub-basin aggregate". Slight changes will also be made to the material sections to better clarify the resolution of the available data. Specifically, streamflow data was collected once a day at a fixed known time (L116), while for glacier water loss, the mass change attributed to the period 2000-2014 was available (L158, see also in AC2 the updated version of the sentence in the response to comment on L156-165);
- the number of iterations for each calibration: technical details on the calibration algorithm configuration will be included in Section S1 of the Supplement, as reported below. "For model calibration, a single-objective global optimization algorithm was applied, specifically the Covariance Matrix Adaptation Evolution Strategy, CMA-ES (Hansen et al., 2003). Several CMA-ES control settings were not altered from their defaults, except for three termination criteria. Specifically, "TolFun" and "TolX", which are mainly associated with the objective function values and the standard deviation of the normal distribution used to sample the parameter values, respectively, were set at 10^{-4}. Furthermore, the maximal number of function evaluations ("MaxFunEvals") was fixed at 10^5. The model parameters were encoded in the range [0; 10] so that they would presumably have similar sensitivity. For this purpose, a linear or logarithmic transformation was used, depending on the original parameter range. For each parameter, the initial solution point was generally chosen equal to the intermediate value (i.e., 5), and the initial

standard deviation for the sampling distributions was set to one-third of the parameter range (i.e., ≈ 3.33)".

2) In terms of validation, the comparison of water balance components across different scenarios should be strengthened, particularly with respect to AET and glacier meltwater loss.

We will conduct a more systematic analysis of the water balance in the different simulation scenarios in Section 3.3, also summarizing the points of interest already introduced in Section 3.1 ("Calibration and validation analysis"). See also the response to comment 15, where an indicative figure for comparison is provided. With reference to the AET and glacier melting components, it is worth noting that the differences between scenarios are mainly explained by the inclusion or not of these variables in the calibration, rather than from specific model configurations. This results in some scenarios being very similar to each other in terms of water volumes, which is why some analyses were illustrated with reference only to the baseline, having previously highlighted how specific scenarios differed from it.

3) Given that runoff observations are rare in complex terrain areas, while satellite-observed evapotranspiration and glacier changes are more readily available, it is recommended to consider another scenario that explores the model performance when calibrated without using observed runoff (i.e., utilizing only AET and glacier meltwater loss), which may have broader applicability in complex-terrain areas.

In the revised version we will integrate the scenario recommended by the reviewer (Scenario 5). Here we anticipate that the simulation has returned some interesting results, with the AET and glacier meltwater loss data proving effective in correcting for precipitation and thus representing a water balance which does not deviate excessively from the other scenarios, despite significant limitations in simulating streamflow without specific constraints. The table below shows the updated results in terms of efficiency indices, integrating Scenario 5.

Period	φ	Scenario							
		1	2	3	4	5	1B	1C	1D
CAL	KGE	**0.882**	**0.894**	**0.907**	**0.911**	0.656	**0.904**	**0.930**	**0.878**
	Eff_{APFB}	**1.000**	**1.000**	**1.000**	0.940	0.131	**1.000**	**0.999**	**1.000**
	Eff_{AET}	**1.000**	**1.000**	0.651	0.642	**1.000**	**1.000**	**0.999**	**1.000**
	Eff_{IMV}	**0.993**	0	0	0	**0.994**	**0.995**	**0.982**	**0.993**
VAL	KGE	0.828	0.823	0.856	0.860	0.582	0.873	0.918	0.831
	Eff_{APFB}	0.966	0.955	0.982	0.935	0.301	0.952	0.951	0.968
	Eff_{AET}	0.998	0.998	0.648	0.640	0.995	0.999	1.000	0.998

Table 3. Efficiency indices considered for the objective function, during the calibration (CAL) and validation (VAL) periods, for the different scenarios. In bold the values of the metrics optimized during the calibration.

Some minor points:

4) The title focuses on flood modelling, which is inconsistent with the paper's limited coverage of this topic. It is suggested that the title be revised to more accurately reflect the actual content of the paper.

We agree with the reviewer on the lack of in-depth investigation at the scale of flood events, due to limitations in the representation of extreme and localized rainfall in the coarse meteorological dataset. Hydrological modelling is in any case oriented towards the high flow regime, with satisfactory streamflow performances offering insights into flood response dynamics and indicative of the potential as a tool for flood simulation. Accepting the suggestion, we will change the title to the following:

"Harnessing multi-source hydro-meteorological data for high flows modelling in a partially glacierized Himalayan basin".

5) Line 15, "the model … precipitation input", observed streamflow was also used for model calibration.

We will modify the sentence as follows: "The model was calibrated using multi-variable data, including satellite-based glacier water loss and actual evapotranspiration in addition to streamflow, also to address bias in the precipitation input".

6) Line 18, "Multi-variable calibration improved…", Multi-variable calibration not always improved the simulation of hydrological fluxes, as evidenced by the poorer streamflow simulation in Scenario 1 compared to Scenario 4. However, multi-source data calibration can provide a more plausible representation of hydrological processes.

We will modify the sentence as follows: "Multi-variable calibration provided a more plausible representation of hydrological processes and highlighted the value of using complementary satellite-based information in data-poor mountain regions".

7) Line 184, the method for rainfall-snowfall partitioning should be provided.

The partition method is applied in a basic version, with precipitation classified as snow if the air temperature is equal or below the threshold, and as rain otherwise. We will go into more detail by modifying the part of interest as follows: "A common threshold value was applied to air temperature both to classify simultaneous precipitation as rain or snow and to reproduce melting processes within the well-known degree-day method, with the snowpack acting as temporary water storage. Specifically, melting over snow- and ice-covered areas was simulated based only on hourly air temperatures as: …".

8) Does DDF in Equation 3 represent the degree-day factor for snow, ice, or both?

As stated in L204, Equation 3 is referred to snow, but due to the linearity in Equation 2 a relationship of the same type describes the variation of DDF_{ice} with altitude (with only the rescaling of the minimum and maximum DDF values).

9) Line 227-229, it would be appreciated to provide a figure or table related to these results.

We will add a figure in the Supplement (Fig. S4), which we also provide here.

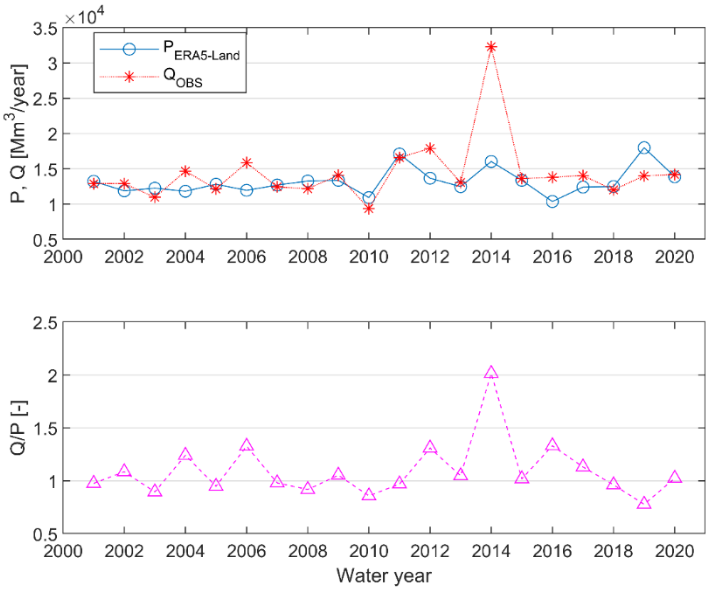


Figure S4. Annual distribution of ERA5-Land precipitation and gauge streamflow volume at basin scale (on the top) and corresponding values of the runoff-to-precipitation ratio (on the bottom).

10) Why are glacier mass losses not simulated in Scenarios 2-4?

As stated in L290-293, glacier mass losses are not simulated in the absence of specific constraints (i.e., Φ_4 in the objective function), mainly due to underestimation of precipitation which would tend to be compensated by overestimation of modelled ice melt, considering also that the process that is thus neglected has a limited impact on the water budget at basin scale.

11) Line 289, the basis for determining these weights requires clarification.

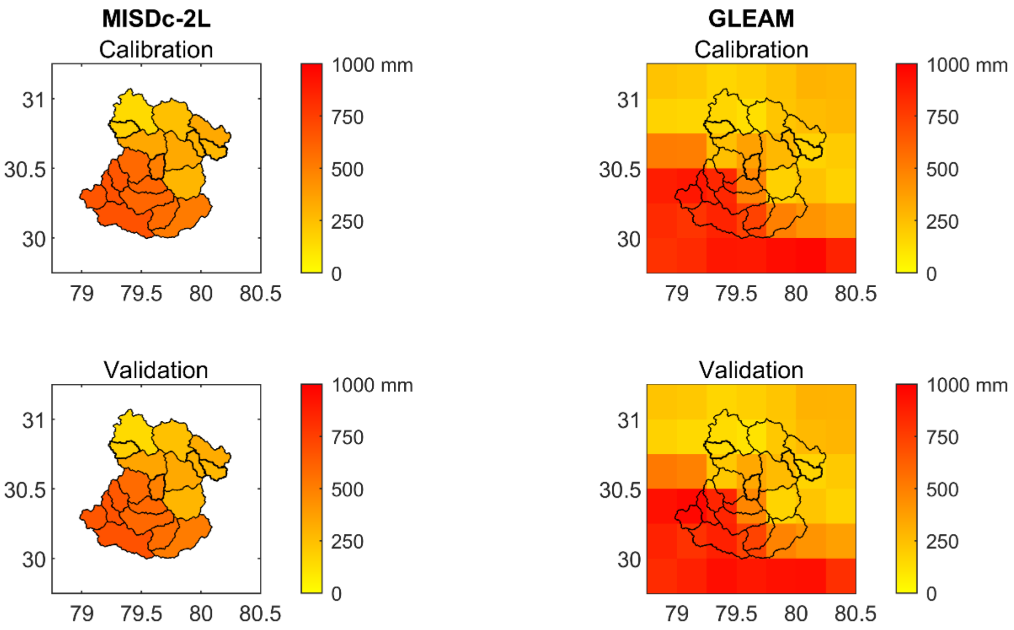
We will restructure the sentence in L259-263 by providing additional details to clarify the basis for determining the weight values: "Despite some disadvantages (Efstratiadis & Koutsoyiannis, 2010), such an embedded multi-criteria calibration approach is widely used (e.g., Gao et al., 2017; van Tiel et al., 2018; Mei et al., 2023), adopting suitable weights for an acceptable trade-off between the different objectives based on the model purpose. The weights allow the user to give different priorities to certain objectives and, with respect to these, they should reflect factors such as relative importance, reference data reliability, and actual differences in scale. Here, the practice of seeking a balanced solution by performing prior tests with variable weights was followed (e.g., Viviroli et al., 2009; Tarasova et al., 2016; Sleziak et al., 2020; Ruelland, 2024; Wagner et al., 2025), considering how the individual components were simulated".

12) Line 423-424, it is suggested to provide the root mean square errors for calibration and validation periods, respectively.

We will provide the metric values as requested. It should be noted that in the updated version this assessment will be done with reference to version 4 of GLEAM. The latter will be considered due to the comments subsequently provided by the second referee. The sentence will be changed as follows: "Figure 4 presents monthly timeseries of basin-scale averaged AET, provided by GLEAM v4 and the hydrological model for Scenario 1. A good agreement was obtained, with a root mean square error <u>of about 3.1 and 2.8 mm month^{-1} during calibration and validation periods, respectively</u>".

13) Figure 5, it is recommended that the mean annual AET for both the calibration and validation periods be presented separately.

We thank the reviewer for pointing this out. AET (both simulated and reference) was found to be rather stationary, with no significant differences in the calibration and validation sub-periods (see also Fig. 4). This also applies to spatial patterns, as highlighted by the figure below, which aligned with the reviewer's request but was then replaced with Fig. 5 present in the manuscript and referred to the entire study period. Agreeing that it is appropriate to provide this information, we will modify the related sentence to: "Figures 5a and 5b illustrate spatial distributions of mean annual AET at sub-basin and pixel scales for MISDc-2L and GLEAM v4 data, respectively, <u>over the entire study period since there are no significant differences between the calibration and validation years</u>". The final version of Fig. 5, which now integrates three panels and returns AET differences at the sub-basin scale, is provided in AC2. Even using GLEAM v4, no significant differences in spatial patterns were observed between the two periods.



14) Figure 6, it is recommended to incorporate reference values for glacier mass loss into this figure.

We will insert an updated composite figure (see below) with a second subplot representing the cumulative distribution of simulated glacier water losses, where the reference data for comparison are also reported.

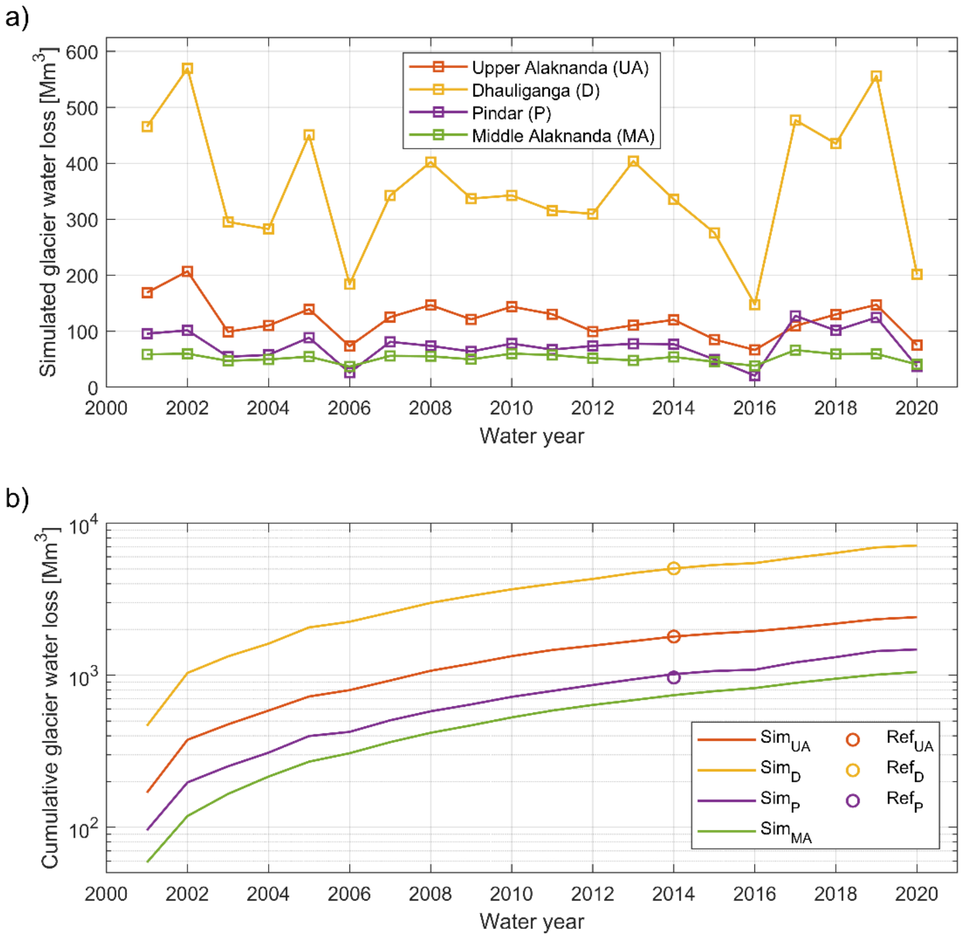


Figure 6: Simulated glacier water loss in the study period, for different sub-basin aggregates according to Scenario 1: annually distributed (a) and cumulative (b) values, the latter compared with the geodetic reference for the calibration period.

15) It is recommended to compare the mean annual water balance across different scenarios, following the format of Figure 8.

We will integrate the suggested analysis, also through the indicated figure, which we report below.

"Compared to the baseline, the other simulation scenarios show deviations in terms of mean annual water balance (Fig. 9) which can be explained by various elements previously highlighted. The precipitation volume is mainly governed by which terms (I_{melt}, AET, and Q) are simulated and considered in the multi-variable calibration. The distribution of total precipitation between rainfall and snowfall is obviously analogous in the scenarios that adopt the adjustment described by Eq. (4), while Scenarios 1B and 1C are distinguished by a higher fraction of solid precipitation. Scenario 1C, which has a more flexible precipitation adjustment, best captures the observed streamflow volumes (see the term $\mu_{\mathrm{sim}}/\mu_{\mathrm{obs}}$ close to 1 both in calibration and validation) resulting in a lower precipitation increase among all the baseline variants. The consistency across scenarios of snow and glacier modules parameters translates in very similar ice melting and snowpack accumulation volumes, depending on whether the contribution of the glaciers is considered or not. Similarly, the AET component remains at approximately the same values depending on whether it was included in the objective function or not. The configurations of the individual scenarios have a significant impact on the distribution between surface and subsurface runoff, with the latter being reliably represented as the major contributor to the total streamflow (except in Scenario 5 where the overestimation of infiltration excess was previously highlighted)".

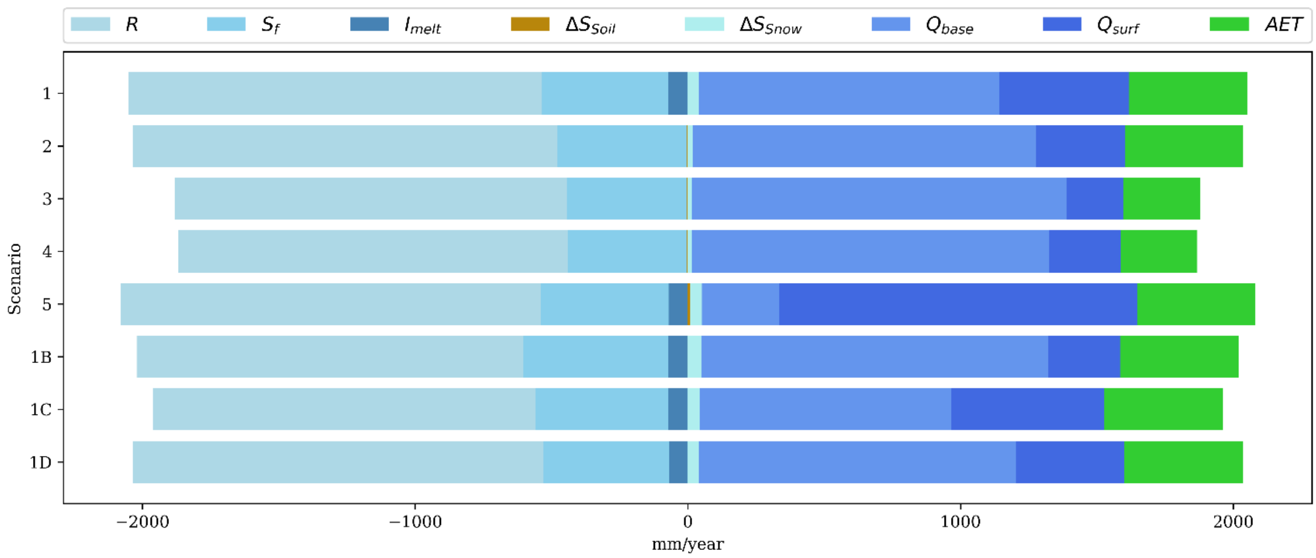


Figure 9: Simulated mean annual water balance for the study basin, according to Eq. (11), for the different scenarios.