Author's response to editor (April 13th, 2026)

Impact of seasonal snow on the recharge of a mountain karst aquifer under climate change: the Dévoluy case study (Southern Alps, France).

EGUSPHERE-2025-5654

Nathan Rispal[1], Bruno Arfib[2], Philippe Audra[1], Pierre Henry[2], Benoît Viguier[3], Alexandre Zappelli[2], Ludovic Mocochain[4], Marianna Jagercikova[4], Christine Vallet-Coulomb[2], Hélène Miche[2], Laurent Cadilhac[5], Vianney Sivelle[6], Olivier Delestre[1]

[1] Univ. Côte d'Azur, Polytech'Lab- UPR 7498, Sophia- Antipolis, France.
[2]Aix Marseille Univ., CNRS, IRD, INRAE, CEREGE, Aix-en-Provence, France.
[3]Géoazur, Univ. Côte d'Azur, CNRS, Observatoire de la Côte d'Azur, IRD, Valbonne, France.
[4]Independent researcher.
[5]Agence de l'Eau Rhône Méditerranée et Corse, Lyon, France.
[6]HSM, Université de Montpellier, IRD, INSU,CNRS, Montpellier, France.

Correspondence to : Nathan Rispal (nathan.rispal@univ-cotedazur.fr) and Bruno Arfib (arfib@cerege.fr).

Dear Dr Ackerer,
We appreciate the invitation to revise our manuscript based on the referees' comments.
We have carefully considered the comments from Referee 1, who suggested strengthening the methodological approach of our study. The main recommendations concern adding uncertainty assessment to our temperature variation scenarios and clarifying the presentation of the model structure and validation process.
Referee 2's main comments focused on the need for better quantitative evaluation of model performance using additional metrics such as RMSE, MAE, and bias. The Referee also proposed improving result visualization by modifying some figures.
Finally, Referee 3's main concerns were about better framing the topic and scope of the study, clarifying the use of a first-order approach to test the sensitivity of simulated discharge to temperature change, and providing more thorough discussion of evapotranspiration response to warming temperature
To address the reviewers' comments, we applied the corrections we had proposed in response to the reviewers at the end of the HESS open discussion.

Revisions made on the manuscript are explained in the following pages. Referees' comments are written in italic. Changes are referenced by line number of the reviewed manuscript.

We believe these revisions effectively address the referees' concerns and significantly strengthen the manuscript.

Sincerely yours,

Nathan Rispal

Referee 1 :

Authors: We thank the Referee for the detailed assessment of our study. Following the comments and recommendations from Referee 1, we have revised the way the models are presented: we now present Model 1 without a snow routine, followed by Model 2 with a snow routine. This validated Model 2 with a snow routine is then used to test two scenarios involving an increase in air temperature. The two scenarios enable us to conduct sensitivity tests on the response of spring discharge to temperature variations in a karst aquifer with seasonal snow. In the revised version of the article, we have modified the results presented for the temperature increase scenarios: we now show 5,000 simulation results, and can thus provide an uncertainty assessment of the results (as suggested by Referee 1). We therefore believe we have improved the study and addressed the main weakness identified by Reviewer 1. The article now appears to us to be more robust from a methodological standpoint, which in turn lends robustness to the reported results on the impact of air temperature variations on the discharge of the karst springs studied.
We would also like to take this opportunity to add Dr. Vianney Sivelle as a co-author (CNRS-France). His contribution was essential in addressing the points raised by the referee.

R1.1: As already mentioned, the current study design seems to be a bit shallow. While I agree that the study of temperature changes and their impacts on catchment dynamics is relevant, this seems to be done a bit superficial right now. The use of a very conceptual model does not really allow to expand the study to analyse model system states or related aspects. However, what I would think could be interesting in this context is an uncertainty assessment of the model (results). As described in line 286, 10,000 parameter sets yield simulation results that satisfy the objective function criteria. These could be used to conduct the analysis regarding the effect of temperature changes. This would result in a range of potential model reaction behaviour under changing temperatures, letting us make more robust interpretations and deductions from the results. I think this would greatly enhance the study. In the same way, different model structures and their impact on simulated discharge under temperature shift could be included in the analysis — as you mentioned that there are different structures that show equally good model performance.

Authors: As noted by the reviewer, the model optimizes all parameters using a quasi-Monte Carlo approach (n parameter sets), which gives access to n results. In the initial version of the paper, we used the parameter set giving the highest objective function value as the calibrated model (W_{obj_max}). Following reviewer's suggestion, we have revised the paper to incorporate the full ensemble of n simulation results for the temperature change scenarios, providing an uncertainty assessment of the model results and giving more robustness to the predictive discharge, including the range of results between 5-95th percentile.

Using the KarstMod software, we simulated n parameter sets for calibration and validation periods, benefiting from a new beta version of the software that allows the use of n sets of calibrated parameters to run simulations with modified input conditions (e.g., a change in temperature). We retain the 5,000 best parameter sets, for which the objective function exceeds 0.85.

Manuscript has been changed in method section line 232-238 and line 290-297 to explain how KarstMod allows retrieving the 5,000 best parameter set and how this have been used and parametrized in this study.

Results are presented in time series graphs and violin plots (Matplotlib-Python), and figures have been revised accordingly, also incorporating the visualization improvements suggested by Referee 2. Modified figures are : Fig. 4; Fig. 6, Fig. 7; Fig. 8; Fig. 9; Fig. 10; and Table 1.

To illustrate the consistency between both approaches (W_{obj_max} and 5,000 parameter set), we focus on August, the critical low-flow month when water demand is highest. In the original manuscript, monthly mean August discharge was reported as 1.19 $m^3 \cdot s^{-1}$ for Model 2, decreasing to 0.86 $m^3 \cdot s^{-1}$ under T+2°C (-28%) and 0.67 $m^3 \cdot s^{-1}$ under T+4°C (-44%). In the revised analysis, the ensemble median values for August are 1.47 ± 0.24 $m^3 \cdot s^{-1}$ for Model 2, 1.01 ± 0.21 $m^3 \cdot s^{-1}$ under T+2°C (−26%) and 0.85 ± 0.19 $m^3 \cdot s^{-1}$ under T+4°C (-42%), where uncertainty is expressed as the Median Absolute Deviation (MAD) across the 5,000 parameter set. The 5-95th percentile ranges are 0.91-2.01, 0.63-1.55 and 0.46-1.29 $m^3 \cdot s^{-1}$ respectively.

The relative decreases between scenarios are consistent between the two approaches W_{obj_max} and 5,000 parameter set (-28% and -44% vs. -26% and -42%). The projected reduction in August discharge under both warming scenarios is therefore a robust signal across the full parameter space.

These modified results are displayed in "results" and "discussion" sections where we also discuss the benefit of taking into accounts the distribution of the results from the 5,000 calibrated parameter sets line 565-571.

Regarding the suggestion to compare different model structures, we agree this is an interesting perspective. However, testing and comparing numerous model structures requires extensive analysis that goes beyond the scope of the present study. The chosen model structure (based on preliminary recession curve analysis) is a standard karst model that adequately represents the fast and slow flow components typical of the Dévoluy system. The primary objective of this study is to accurately reproduce the observed discharge at this specific catchment, and to assess the effect of seasonal snow under warming temperature scenarios. The revised manuscript now includes 5,000 simulation results for the temperature increase scenarios, providing a robust uncertainty assessment of the projected hydrological changes. The influence of model structure on the results could be addressed in future work.

We now better explain why and how the model structure is defined. We present the fact that the model's structure is constrained by hydrograph observation in the introduction (line 77-78) and then line 315-322 in method section to ensure that the reader correctly understand why we choose to focus on one specific model structure.

R1.2: Also, I dont really see the sense behind the split in model 1 and model 2 in the current study design. If model 1 is without snow routine, meaning that all precipitation occurs as rainfall, then why is a model 2 necessary with snow routine deactivated? From my understanding, the study only needs one model without snow routine and one model with snow routine (both including potential parameter sets / model structures that show satisfactory model performance) to assess the impact of snow inclusion on model performance. This would enhance clarity regarding the methodology.

Authors: We agree the recommendation of the reviewer. We believe that the paper will be clearer by presenting only one model without snow routine and one model with snow routine, and then 2 scenarios of air temperature shift. We have removed Model_2_without snow routine. The two scenarios (Model 2 T+2°C and Model 2 T+4°C) are enough to show the case with a large decreasing amount of snowfall. This led to the following modification in the manuscript :

- Suppression of section "3.4.3. Scenario simulations : altered temperature and snow regimes" (previously line 348)
- Suppression of section "4.2.1. All precipitation forced as rainfall" (previously line 440-453)
- Suppression of comparison between Model_2 and Model_2_SRoff in section 5.1 (previously line 493-501)
- Fig 8 (previously Fig 6) have been modified with suppression of Fig. 6e.
- Discussion : section 5.1 has been modified with no longer comparison between Model 2 and Model 2_SRoff from now line 495-517.

R1.3: Lines 74 to 77: I think this is more methodology rather than being relevant for the introduction.

Authors: These sentences have been removed as suggested by referee 1.

R1.4: Line 180: I dont think this figure is necessary in the current form. Subfigures a and c really dont offer anything that needs to be depicted as a figures. Subfigures b and d are debatable, if they are really necessary, as the snow depth is shown later within comparisons as well as the discharge. The only unique information here is the discharge for the Souloise river, which is not used later if I am correct?

Authors: We agree that this figure could be simplified by deleting subfigure (c). We chose to keep (a) precipitation to illustrate the range of daily values in this case study in France, discharge at Gillardes Spring and Souloise river because comparison between river and spring is usefull to show the predominance of the karst discharge, (b) snow observation demonstrates that this mid-altitude karstic catchment is subject to seasonal snow. We can also put this figure in supplementary material if the reviewer and editor want to shorten the length of the paper.

R1.5: Lines 204 to 208: Are [L], [L/T], or [/T] supposed to be the units for the corresponding values? Are those something like m³, litres, km²?

Authors: This corresponds to the physical dimensions of the various variables to allow kind of generalization of the model regarding the input data provided by the user. To avoid potential confusion, we have replaced L by Length, T by Time (now line 213-216)

R1.6: Line 222: I would suggest writing things like Wobj or anything like that as some sort of variable: W_{obj}, $Wobj_{min}$, E_{min}, P_{sr}.

Authors: We have replaced Wobj as W_{obj}, Emin as E_{min} and other variables as you suggested in the whole manuscript.

R1.7: Line 234: I think to say that you only using one single evaluation crtieria is similar to a multi-objective criteria calibration is difficult, because it's not true. You still only rely on the discharge, so no other aspects like snow height or ETa are explicitly evaluated, which would be a real multi-criteria evaluation.

Authors: We agree that referring to a single composite performance criterion as being comparable to a true multi-objective calibration may be misleading. Although KGEnp integrates several statistical components (correlation, bias, and variability), the calibration in our study is indeed based solely on discharge observations and does not include additional state variables such as snow storage or actual evapotranspiration.
We have revised the manuscript accordingly to clarify that KGEnp is a composite metric applied to discharge only, and we changed wording suggesting equivalence with a true multi-criteria or multi-variable calibration : "KGEnp provides a robust evaluation of the agreement between observed and simulated discharge across a wide range of hydrological conditions, combining correlation, bias, and variability within a single composite metric. The calibration is performed using discharge observations only, and KGEnp is therefore applied as a composite discharge-based criterion" Line 262-265.

R1.8: Lines 237 to 242: These "soft-criteria" mentioned here are surely important, but the calibration (and evaluation) of model parameterisations is only done on the discharge by the KGEnp. Why are those aspects mentioned here not incorporated into the calibration/evaluation?

Authors: The spring discharge is correctly modeled (using the KGEnp objective function) by automatically adjusting the parameters that influence flow dynamics (k coefficients) and also the parameters for snow accumulation and melt (snow routine parameters). The soft criteria we list are proposed by the KarstMod software and we mention them as part of a "good practice" approach to modeling.

R1.9: Lines 247 to 265: Why is the snow routine introduced after model optimization? I think it would make more sense if the order is following a more logic way. First the study area, then the data, then the model and how it works, then how its is calibrated and then how the further study is conducted.

Authors: We agree with this suggestion and we have switch part 3.2 (now line 239) and 3.3 (now line 258) to present first the snow routine and then the model optimization.

R1.10: Lines 269: I feel like the figure would benefit from a legend that shortly points out what each abbreviation stands for. Currently it is in the caption, but that makes it less easy to grasp.

Authors: We choose a similar representation to the one regularly used in the literature dealing with KarstMod rainfall discharge model and its applications (Çallı et al., 2022; Mazzilli et al., 2019; Sivelle et al., 2019) . To keep the figure clear, we suggest keeping the abbreviations explained in the caption.

R1.11: Lines 271 to 275: I think this paragraph can be deleted. Lines 276 to 277: This is more suited for the introduction. No need here.

Authors: We agree with the reviewer's comment. Lines 271–275 have been removed. Regarding lines 276–277, we have simplified by removing the lines mentioned and start the section directly with the modelling strategy.

R1.12: Lines 286 to 288: Why is only the "best" model retained if there are several parameterizations that seem to be fulfilling the criterion — and therefore show sufficiently good model behaviour?

Authors: As answered in the introductory comment R1.1, we have revised the paper by adding the results of the n=5,000 sets of calibrated parameters (all the results satisfy Wobj (KGEnp) > 0.85) applied to the 2 warming scenarios. We have run the 5,000 simulations with modified input conditions (i.e., temperature shift T+2°C or T+4°C)..
These results are now displayed in section "4. Results" and "5.Discussion". The range of probable discharge simulation for every model is highlighted by specific curves in figures (e.g. median, percentile 5 and 95). The previously "best model" (for Wobj max) is now only one of the possible results. When specific values are given, such as monthly or seasonal discharge, an interval is given around this value (±) and correspond to the Median

Absolute Deviation (MAD). Previous "best model" objective function result (KGEnp) is now replaced by median value.
We are now able to assess the uncertainty of the modelling results, and give more robustness to the predictive discharge.

R1.13: Also, what does "Cross calibration validation tests did not improve performance" mean here? From my understanding and experience, the sense of a cross-calibration would here be to assess, if model performance is highly influenced (dominated) by a single year, for example.

Authors: Concerning the cross-calibration tests, we agree that our initial sentence was unclear. We performed a split-sample test by exchanging the calibration and validation periods to assess whether model performance was strongly dependent on a specific time interval (validation from September 2015 to September 2016 and calibration from September 2016 to September 2019). The resulting Wobj values remained comparable between calibration and validation configurations (median of the 5,000 parameter set higher than 0.74), indicating that model performance is not dominated by a particular year or sub-period and supporting the temporal robustness of the calibrated parameter sets. We have modified the text line 294-297.

R1.14: Lines 320 to 324: kinda a repetition of what is mentioned in lines 310 to 313.

Authors: We agree that this section contained some redundancy. In the revised manuscript, we have condensed paragraph 1 to avoid repetition and to present the information only once in a clearer and more concise manner line 340-343.

R1.15: Lines 324 to 325: Expalantion of positive and negative shift makes no sense in this formulation. It should be formulated more as a relative value that is negative for heights with greater values than the reference and positive when heights are lower than the reference.

Authors: We have changed the formulation of the sentence according to your recommendation : "It is negative with altitude higher than the reference and positive with when altitude is lower than the reference" now line 337-338.

R1.16: Lines 332 to 338: This can be combined with the explanations before to shorten on this, as it seems trivial.

Authors: We have shorten the explanation on the Ts parameter, highlighting its role and how it should be parametrized : "This regional gradient is a rough estimate, close to the one proposed by Ruelland (2023), but the model will compensate for this approximation by adjusting the T_s parameter (temperature threshold for melting), to improve the distribution of snow and rain over the catchment." Now line 355-358.

R1.17: Lines 339 to 342: You said those alternative structures did not improve model performance. Does that mean, those are equally feasible? Why are those then not incorporated into evaluation? Model structure uncertainty could be well assessed here and would help increase the robustness of the model result's interpretation.

Authors: The current structure of the model is based on the identification of the main recession coefficients in the spring discharge time series before the rainfall-snow-discharge modeling. Once the main flow components (very fast flow, fast flow and slow flow) have been identified and treated as compartments, the purpose of the study was not to test several structures. The sentence line 339 was unclear. We decided to simplify the text here as we now already explain in introduction (now line 77-78) that we use only one model structure. Line 315-322 we explain how this structure is determined :
"The model structure is determined on the basis of a preliminary analysis of the hydrograph, which allows dividing the recession curve into successive phases. On a semi-logarithmic plot of discharge versus time, these recession phases appear as distinct linear segments with different slopes, each representing a specific drainage component of the karst system. Three linear segments are recognized and their regression slopes give three recession coefficients used as an initial estimate of the characteristic time for the three conceptual flows through compartments discharging to the spring in KarstMod. For the automatic calibration, these three values define a pre-constrained parameter range based on expert knowledge, within which the optimization is run to search for the best-fit values. This ensures parameters remain within physically plausible bounds while allowing sufficient exploration of the parameter space."

R1.18: Line 356: Isn't the comparison of the model with and without snow module the relevant method to assess the impact of snow incorporation on model performance, rather than the inclusion of different temperature changes?

Authors: We refer to our previous response regarding the removal of Model 2_SRoff on comment #R1.2 As explained, this configuration corresponds to a boundary case in which temperature changes would prevent snowfall entirely. To simplify the manuscript and clarify the presentation of the simulation workflow, we have removed this model from the study and deleted the sentence line 356.

R1.19: Lines 369 to 373: This is more suited for model description.

Authors: We have deleted the model 2_SRoff accordingly to your recommendation so we deleted theses sentences.

R1.20: Lines 374 to 377: Delete.

Authors : We have removed this paragraph.

R1.21: Line 380: The figure should be updated accordingly with the changes made to the methodology.

Authors: As we decided to remove the use of the model_2SRoff we have modified the figure accordingly line 304.

R1.22: Line 389: The table can probably be moved to the supplementary.

Authors : We chose to keep this table in the main text, as it provides essential information about the models parameters, their calibration ranges, and optimized values. This information is key for the reader to directly assess the calibration results and to evaluate the physical consistency of the optimized parameter sets. Moving it to the supplementary material would reduce the readability of the results section. However, values of the table have been modified according to modifications made to implement the 5,000 best parameter set. The table now display the median value of each parameter.

R1.23: Line 416: Figure numbers are wrong through the whole manuscript from here. Please check.

Authors: There was a problem with figure numeration from there, we have now fixed it, figures are numbered from 1 to 10. Fig. 11 is in supplementary material.

R1.24: Line 425: Please check sentence syntax.

Authors: We have modified the sentence to make it clearer ". Compartment M acts for baseflow, with large inertia, discharging to the Spring for months after the recharge period feed by rainfall and snowmelt. Compartment C, by its high storage (high water level) and high dynamic response (high recession coefficient); accounts for discharge during high-water flow in spring season influenced by snowmelt but also during peak events from rainfall or snowmelt during the year." Now line 433-436.

R1.25: Line 428: The model can not provide a more robust performance criteria. Semantically not logic.

Authors: The sentence was not clear. We compare results and performance criteria from Model 2 to Model 1 (initial model without snow routine). This is why we said that the value of the performance criteria is more robust, higher. We modified the sentence to make is clear and logic : "By integrating snow storage and melting, the model not only aligns more closely with observed hydrological patterns but also provides a higher value of performance criteria, which strengthens its reliability for further scenario analysis" now line 439.

R1.26: Line 443: What sense has a model here that is not validated?

Authors: In section 4.1.1 we demonstrate that Model 1 is not valid for the Dévoluy catchment, as it fails to correctly reproduce discharge during periods influenced by seasonal snow. Presenting temperature change projections with an invalid model would therefore be meaningless. Additionally, as discussed in our response to R1.2, we have removed Model_2SRoff and section 4.2.1, as the two warming scenarios (Model 2 T+2°C and Model 2 T+4°C) are sufficient to illustrate the expected decrease in snow-driven discharge under future climate conditions.

R1.27: Lines 512 to 513: If these temperature variant simulations are not to be understood as predictive forecasts, why is it framed like that through the manuscript?

Authors: We thank the reviewer for raising this point. We acknowledge that the framing of the temperature-variant simulations may have been ambiguous throughout the manuscript. These simulations are not predictive forecasts based on projected climate data, but sensitivity tests assessing the impact of temperature increases on karst aquifer dynamics using observed inputs modified by a uniform temperature offset. We have revised :

- Abstract line 21-22 :

“We then tested the sensitivity of the karst spring discharge to temperature perturbations”

- Introduction line 88-94 :

“In this paper we aim to (1) simulate the daily discharge of a karstic spring in a mid-altitude Alpine environment using the KarstMod V3 platform (Sivelle et al., 2025), (2) quantify the contribution of seasonal snow to spring discharge, and (3) test the sensitivity of a karst spring discharge to temperature perturbations consistent with TRACC-type (climate change adaptation trajectory) warming trajectories (French TRACC guideline, Soubeyroux et al., 2024). The originality of this study is based on (i) a mid-altitude Alpine karst system with a strongly seasonal snowpack, (ii) the use of a parsimonious rainfall-snow-discharge conceptual model, (iii) a scenario design based on simple temperature perturbations, and (iv) a quantitative focus on seasonal metrics, focusing on the summer period”

- Discussion line 566-568 :

“These scenarios are not intended as predictive forecasts, but rather as a temperature-sensitivity test to explore how the system responds to temperature forcing in a hypothetical case where precipitation remains as observed between 2015 and 2019

These modifications explicitly frame these scenarios as sensitivity analyses, as also requested by Referee 3, and clarify that this represents a first-step approach that could be further improved by incorporating fully predicted input datasets.

R1.28: Lines 590 to 592: Does the fact, that the scaling parameter for the catchment had to be expanded to 191 km² instead of the actual 150 km² not imply that underground exchanges or a differing aquifer catchment size are very likely possible ?

Authors: We do not consider that a larger catchment size or significant exchanges with other aquifers are plausible. That is why we consider that this overestimation of the catchment size is due to a bias in precipitation in input data. However, this point was not sufficiently explained in the initial version of the manuscript, particularly in the Case Study presentation section. The Dévoluy system is a perched karst aquifer that is clearly disconnected from other potential aquifers due to its topographic position and geological setting, as shown in Fig. 1. The size of the catchment is therefore a well-defined parameter. We have modified the manuscript line 143-148 :

“The catchment area drawn on Fig. 1 extends over 176 km²; it includes the watershed of the Souloise and Ribière rivers and the recharge area of the Gillardes Spring. The Bure Plateau is excluded because it is drained southward. The Dévoluy massif is a perched karst aquifer forming an isolated syncline (Fig. 1), topographically and geologically disconnected from adjacent massifs and aquifers. This specific setting ensures that the catchment boundary is well constrained with minimal uncertainty, and makes significant underground water exchanges with surrounding systems unlikely.”

Minor comments:

R1.29: Line 126: main instead of "mains"

Authors : This has been corrected.

R1.30: Line 171: are rare and make

Authors : This has been corrected.

R1.31: Line 193: hydrological

Authors : This has been corrected.

R1.32: Line 207: the reference unit length is fixed Lines 249 to 251: Units in []-brackets, multiply dots wrongly formatted.

Authors : This has been corrected.

R1.33: Line 268: Input liquid precipitation stands for rainfall?

Authors: Liquid input precipitation stand for effective rainfall at the outlet of the snow routine : it depends on the water released by the snow routine and can be different to the input daily precipitation.

R1.34: Lines 308 to 310: Formal language.

Authors: We have modified the sentence " However, the uniform distribution of snow across the Dévoluy massif being too rough an approximation, sub-catchments are introduced to better discretize snow storage and melting zones"

R1.35: Line 330: Range

Authors : This has been corrected.

R1.36: Line 331: unit formatting

Authors : This have been corrected.

R1.37: Line 359: "climate models already show warming"

Authors : This have been corrected.

R1.38: Line 364: ambiguous formulation regarding the temperature definition. Do you mean you use the temperature recommendations to define the simulated temperature increases?

Authors: As also suggested by Referee 3, we clarify that the French "TRACC" dataset consists of a set of regional climate trajectories. From these trajectories, we extract the temperature perturbations applied in our simulations. In other words, the temperature shifts imposed in the model follow the guidance provided by TRACC for plausible warming scenarios. The manuscript has been revised in the introduction line 86-87 and in method section line 373-376.

R1.39: Line 486: comma missing

Authors : This have been corrected.

<u>Referee 2 :</u>

Authors : We thank the Referee for the careful reading of our manuscript and for the positive overall assessment of the modelling approach and its clarity. We appreciate the constructive suggestions provided, particularly regarding the quantitative evaluation of model performance and the improvement of result visualization. We agree that including additional performance metrics and adapting some figures to facilitate comparison between simulations will strengthen the robustness and readability of the manuscript. We have revised the manuscript accordingly and respond to each point in detail below.

 # R2.1: Overall, I found the models and methods were well described. The authors did a good job justifying the inclusion of snow in the karst discharge simulation under multiple climate-change temperature scenarios. Overall, I found the manuscript well written and clear. Thank you for letting me review this research.

Overall, there are a few major requests I have:

1. *Greatest recommendation: the research needs to include RMSE, MAE, bias and/or other metrics between the simulations and in-situ data to calculate and quantify the amount of model error. These will also need associated visualization for the reader to ensure that the model statistically performed well.*
2. *I recommend aggregating the time series (e.g. figure 4-6) into bar charts and/or box and whisker, etc. so the reader can easily compare the simulation results.*

Authors: These recommendations are taken into account and explained in responses to comments in the following document.

Please see my comments below:

Abstract:

 # R2.2: Line 18: At the end of this sentence, "Results show that accounting for snow processes is essential to reproduce the observed discharge dynamics" - I think it would be clearer to add "in karst environments"

Authors: We have added "in karst environment" line 21 to be more precise as you suggested.

Introduction:

 # R2.3: Line 57: The paragraph on hydrologic modelling as an efficient tool and the models that incorporate snow as a parameter come before the next paragraph that explains the importance of incorporating snow into karst models. I think these two paragraphs should be swapped – e.g. paragraph at line 57 should go at line 46.

Authors: Following your recommendation, we have switched these two paragraphs to provide a more logical progression in the introduction.

Methods:

Data

 # R2.4 : 137: It is unclear how "precipitation (mm/day) and mean temperature (°C)" factor into the sentence. Are you trying to say that precipitation (mm/day) and mean temperature (°C) also have the four grid meshes with 8x8 spatial resolution, etc.? If so – then this needs its own sentence – the format of this sentence leaves the reader confused. Please clarify.

Authors: We have split the sentence in two to make it clearer that both temperature and precipitation are computed from the average value of the four selected meshes.
"The average value of the 4 selected meshes was computed and used as daily input to the KarstMod model from 2015 to 2019. The meteorological inputs consist of precipitation (mm/day) and mean air temperature (°C)" now line 144-146.

Accuracy Assessment:

 # R2.5 Please include an RMSE, MAE, bias for evaluating the accuracy of the simulated results compared to the in-situ results. The KGEnp mentioned in Line 411 is one of the few mentions of model efficiency. I would like to see more on model accuracy and bias. Please include figures to visually demonstrate these results.

Authors: KarstMod provides several tools to assess and visualize model performance. In response to your comment, we have expanded the evaluation of model performance by adding NSE, VE MAE and PBIAIS and explain this choice line 298-304 :

"To evaluate model performance, five complementary criteria were selected: Nash-Sutcliffe Efficiency (NSE) for overall fit, Volume Error (VE) for temporal distribution of errors without compensation bewteen positive and negative values, Mean Absolute Error (MAE) for average error magnitude, Percent Bias (PBIAS) for systematic bias direction, and non-parametric Kling-Gupta Efficiency (KGEnp) for integrated assessment of correlation, variability, and bias. This combination allows comprehensive evaluation of both global performance and specific aspects critical for snowmelt simulation in karst systems."

The results of these metrics are discussed line 396-400 and line 418-426. Values are calculated year by year and by yearly snowmelt influenced season from April to august. They are shown in Table S4 in supplementary.

Also we have updated figures in the article to show a range of results rather than only the result of the calibration with the maximum objective function. This was also a request from referee 1. We used KarstMod to obtain 5,000 satisfactory results (for Wobj = KGEnp > 0.85), and we added to the figures the median value and the 5th and 95th percentile values, this allows a better uncertainties assessment of the modeled discharge over the different models tested.

Results:

 # R2.6 Figure 4: the pink/blue in Fig 4A and yellow/gray are hard to discern (e.g. the colors overlap). Can you try more high contrast line colors?

Authors: We have slightly changed the contrast in these colors to make the figure more discernable. Note that the correct number of this figure is now Fig. 6.

 # R2.7 The graphics don't line up with the text. For example, line 470 mentions Figure 8e, then the next figure is Figure 6. Please put the figures after the sections where they are first mentioned.

Authors: There was an error in figure numbering starting from line 430. We apologize for this typing error. This has now been corrected. The figure citations in the text were already correct, but the figure numbering itself was incorrect, which caused the apparent mismatch. We have fixed this issue and ensured that all figures now appear in the correct order and after the sections where they are first mentioned.

 # R2.8 I find the time series difficult to interpret. I think it would be easier if you included aggregated final results for the readers such as:

Figure 4: bar chart showing the average discharge for the simulated and discharge for the warmup, calibration, etc. to easily compare if the simulated had a similar amount, or more, or less predicted snow.

Authors: This figure, in its current form, allows the visual comparison of the Model 1 and Model 2 curves, showing the fact that Model 1 does not correctly represent discharge during key periods influenced by seasonal snow, and that Model 2, with its snow routine, allows a better result to be obtained during these periods. Using box plots based on the calibration/validation period and mixing all discharge values together would not show this mismatch observed on the curves, as the model nonetheless achieves a good overall Wobj criterion, meaning that over the four-year periods simulated, volumes are quite close to observed ones. We therefore choose to keep Figure 4 (now Figure 6) with the curves, adding the modifications suggested by Referee 1 regarding the use of the 5,000 best parameter set . However, following your recommendation to include a figure summarizing the results, we have used a violin chart to illustrate the results by month now numerated Fig.10.

 # R2.9 Figure 5: I recommend a similar aggregated visual that makes it easy for the readers to compare. Also, make the caption clearer to define was E-S, M-S, and C-S mean (e.g. explain it – the reader needs more than a copy/paste from the visual).

Figure 5 is not intended to compare different models but to illustrate internal variables of the model 2, such as water levels and discharge within each compartment. Therefore, an aggregated representation similar to Figure 7 is not straightforward or necessarily meaningful for these time series. However, we propose to clarify the caption and now explicitly define the meaning of E-S, M-S, and C-S to improve readability :

" Fig. 7: (a) Median of 5,000 simulations of seven-day averaged water levels in compartments E, M, and C (Model 2). (b) Median of 5,000 simulations of seven-day averaged internal discharges (QES: discharge from E to S, QMS: discharge from M to S, QCS: discharge from C to S). Seven-day cumulative precipitation is shown on a secondary y-axis (inverted) ."

Please note that the corrected number of this figure will now be Figure 7.

 # R2.10 Figure 6 and 8 – same comment – aggerate so the reader can easily compare.

Authors: The curves shown in Fig. 8 (previously Fig. 6) have been aggregated on a monthly basis in Fig 10 (previously Fig. 8). The corresponding results have been updated to address Referee 1's suggestion regarding the assessment of model results uncertainty. We chose to keep Figure 8 because it allows a visual inspection of discharge evolution across the four modeled years for the different scenarios. Following Referee 1's recommendation, we decided to remove model2_SRoff, the version without the snow routine, from these comparisons.

For example, Figure 7 was easily digestible to me because I could make side-by-side comparisons.

 # R2.11 I think you should reference Figure 4 in the Figure 8 caption so readers can easily reference the different model simulations.

Authors: We propose to update the caption of Fig. 10 (previously Fig. 8) to reference Fig. 4, allowing readers to easily relate the different model simulations as suggested.

 # R2.12 Figure 6A: I find this unimportant – a reader could just as easily find that in the text – or you could put it in the caption, but the bar at the top seems out of place and unnecessary.

Authors: We agree that Fig. 6a is not essential, we have removed the top bar to simplify the figure, and made the same in Figure 2.

Discussion:

 # R2.13 Line 485: Please include your accuracy assessment values behind this sentence to justify and support that snow improved modelling results.

Authors: We have added the accuracy assessment values to support the conclusion that including snow improves the model results. Line 527.

<u>Referee 3 :</u>

We thank the referee for the constructive assessment of our manuscript. From our understanding, the reviewer was convinced by the use of the case study and the KarstMod lumped rain-snow-discharge model to test the sensitivity of karst spring discharge to the expected rise in air temperature under various scenarios. Referee 3's comments helped us to amend the text of the article on the points identified.

 # R3.1: The manuscript addresses an important topic, but I recommend that the authors clarify and nuance their novelty claim regarding the effects of temperature increase on karst spring discharge. There is now a small but growing body of work that has already used modelling approaches to quantify climate-change or warming impacts on karst recharge and spring behaviour (e.g. Doummar et al., 2018; Fan et al., 2023; and related numerical-modelling studies on karst groundwater under climate scenarios). Rather than presenting the study as the first to model the impact of temperature increase on karst discharge, it would be more accurate and convincing to position the contribution in terms of its specific combination of features: (i) a mid-altitude Alpine karst system with a strongly seasonal snowpack, (ii) the use of a parsimonious rainfall–snow–discharge conceptual model (KarstMod with a degree-day snow routine), (iii) a scenario design based on simple temperature perturbations consistent with TRACC-type warming trajectories, and (iv) a quantitative focus on seasonal metrics, particularly changes in summer/August low flows. Framing the contribution in this way would acknowledge prior work while clearly highlighting what is genuinely new and distinctive in the present case study and modelling set-up.

Authors: The Referee suggests slightly restructuring the topic announcement. We have followed these recommendations, based on the proposed text and modified the introduction line 84-90:

"In this paper we aim to (1) simulate the daily discharge of a karstic spring in a mid-altitude Alpine environment using the KarstMod V3 platform (Sivelle et al., 2025), (2) quantify the contribution of seasonal snow to spring discharge, and (3) test the sensitivity of a karst spring discharge to temperature perturbations consistent with TRACC-type (climate change adaptation trajectory) warming trajectories (French TRACC guideline, Soubeyroux et al., 2024). The originality of this study is based on (i) a mid-altitude Alpine karst system with a strongly seasonal snowpack, (ii) the use of a parsimonious rainfall-snow-discharge conceptual model, (iii) a scenario design based on simple temperature perturbations, and (iv) a quantitative focus on seasonal metrics, focusing on the summer period".

 # R 3.2: I suggest adding a short clarification on the climate scenarios and the choice of perturbations. At present, the manuscript may give the impression that the French TRACC storyline is essentially a temperature-only scenario, whereas in reality regional climate projections provide concurrent changes in temperature, precipitation and other variables. In the modelling set-up, you chose to hold precipitation constant and perturb only temperature (+2 °C, +4 °C), which effectively turns your experiments into temperature-sensitivity tests rather than full climate-scenario runs. This is a reasonable and useful first-order approach to isolate the role of warming on snow processes, recharge timing and evapotranspiration, but it should be clearly framed as such. I recommend that you (i) explicitly acknowledge that realistic future scenarios involve both temperature and precipitation changes, (ii) justify why precipitation is kept unchanged here (e.g. to avoid adding precipitation-scenario uncertainty and to focus on snow-related mechanisms), and (iii) state that using full regional climate projections (with modified precipitation) would be a logical next step for more comprehensive impact assessments.

Authors: We agree that we are applying a first-order approach on temperature sensitivity based on the TRACC storyline. Precipitation is kept unchanged as the study focuses specifically on the hydrological response to temperature change and snow processes. We already specified those points in the first version of the manuscript: in the introduction (lines 82-83), method section (lines 349-352; 364), discussion (lines 511-513) and conclusion (line 621). In the conclusion (lines 630-635) we also discussed the fact that a further step would be to use full regional projections for predictive modelling. While these elements are already present in the manuscript, we have reframed and clarified them in the revised version following Referee 3's suggestions :

- In the abstract line 21-22 : "We then tested the sensitivity of the karst spring discharge to temperature perturbations."
- In the introduction line 84-87: " "In this paper we aim to (1) simulate the daily discharge of a karstic spring in a mid-altitude Alpine environment using the KarstMod V3 platform (Sivelle et al., 2025), (2) quantify the contribution of seasonal snow to spring discharge, and (3) test the sensitivity of a karst spring discharge to temperature perturbations consistent with TRACC-type (climate change adaptation trajectory) warming trajectories (French TRACC guideline, Soubeyroux et al., 2024)."
- In the method line 473-476 :" the TRACC provides a reference warming trajectory for the 21st century, including climate indicators such as temperature, precipitation, and extremes, to guide adaptation strategies across multiple sectors (Soubeyroux et al., 2024). In this study we use the TRACC temperature recommendations to perform a temperature sensitivity analysis, keeping precipitation unchanged".
- In the discussion line 573-577.

 # R 3.3: In the discussion of evapotranspiration responses to warming, the manuscript rightly mentions the role of increasing temperature (via the Oudin PET formulation) and briefly of rising atmospheric CO_2, but it would be important to acknowledge that changes in phenology (longer growing seasons, earlier leaf-out, shifts in vegetation activity) are also expected to significantly affect both potential and actual ET. Earlier onset and extended duration of transpiration can partly offset or amplify temperature-driven PET changes and will influence the seasonal partitioning between ET and recharge. I recommend adding a short paragraph noting that future ET trajectories will not depend on temperature and CO_2 alone, but also on phenological shifts and vegetation dynamics, which are not explicitly represented in the present modelling framework. This will better situate your experiments within the wider set of ecohydrological feedbacks expected under climate change.

Authors : We agree that evapotranspiration responses to warming also depend on vegetation dynamics and phenological shifts, which can modify both the magnitude and seasonal distribution of water losses and are not explicitly represented in our modelling framework. Climate warming is well documented to advance spring leaf-out and extend the growing season ((Menzel and Fabian, 1999; Richardson et al., 2013) with direct consequences for seasonal ET and recharge partitioning. In the revised manuscript, we added a short paragraph in the Discussion acknowledging these ecohydrological feedbacks and situating our temperature perturbation experiments within this broader context line 636-643 :

“In addition, climate warming is expected to induce significant phenological changes, such as earlier leaf-out and a lengthening of the growing season (Menzel and Fabian, 1999; Richardson et al., 2013), which can alter both the magnitude and seasonal distribution of evapotranspiration. An earlier onset and extended duration of vegetation activity may enhance transpiration over longer periods, thereby modifying the seasonal partitioning between evapotranspiration and recharge. These ecohydrological processes are not explicitly represented in the present modelling framework, and their effects may either amplify or offset temperature-driven changes in evapotranspiration. Thus, future evapotranspiration trajectories are likely to depend not only on temperature and atmospheric CO_2, but also on vegetation dynamics and phenological shifts.”

Specific comments

 # R 3.4 : Section 3.1 is missing information about the spatial distribution and discretization of model parameters and the subcatchments . The subcatchments should be additionally delineated in Fig 1.

Authors: KarstMod is a lumped model, not a spatialized model. The specificity added by the snow routine, with non-spatialized sub-catchments, is explained in part 3.3. We have explained more clearly in section 3.1 the fact that KarstMod is a lumped model : line 195; 201and 205.

 # R 3.5 L174 Groundwater recharge: It should be mentioned how this was estimated

Authors: We changed the sentence to better explain how groundwater recharge is estimated by comparing surface discharge to groundwater discharge at the outlet of the massif: " Groundwater discharge predominates in the volumes of water measured at the outlet of the massif. On average, groundwater (i.e. Gillardes Spring) accounts for 86% of the total discharge, while surface flow (i.e. Souloise river upstream the karst springs) accounts for only 14%. From an annual water budget point of view, the previous percentage gives a rough estimate of the area of the watershed contributing to groundwater discharge. Considering that 86% of the watershed recharges the groundwater outflowing at the Gillardes Spring, the surface of the karst recharge area is deduced to be around 150 km².” now line 182-187.

 # R 3.6 L256 This information should be provided earlier in the model description (Section 3.1), where it is currently missing.

Authors: See response to comment 3.4.

 # R 3.7 L257 compartment E – add reference to Figure or explain what this is

Authors: There is no reference to compartment E line 257 or in this part of the manuscript.

 # R 3.8 L331 either a notation table should be provided, or the parameters should be explained everytime when used. Same applies for Table 1- I’d suggest either adding a column explaining the variables or referring to a notation table

Authors: All parameters are presented and explained in Table 1 line. We have a reference to Table 1 at line 361(previously line 331).

Technical corrections

 # R 3.9 L171 precipitation is – there is no plural form

Authors: We have corrected this.

 # R 3.10: L187 add reference or link to the used version of Karstmod

Authors: The link to the used version of KarstMod is provided line 676 :

“The KarstMod modelling software used in this study is freely available online (https://hal.science/hal-02071006).”

The number of the used version is given line 196 “KarstMod V.3.1.0 has been used in this study”