Review of the revised paper egusphere-2025-2157 (R1)

Assessing the Impact of Earth Observation Data-Driven Calibration of the Melting Coefficient on the LISFLOOD Snow Module

By Premier et al.

General Comments

The revised version represents progress: structure and clarity have improved, and some figures are better presented. Nevertheless, substantial revision is still required before this paper can meet HESS publication standards.

The manuscript continues to lack clear scientific positioning, conceptual depth, and interpretative discussion. While the topic (linking Earth Observation (EO) data with large-scale hydrological modelling) is relevant, the paper does not yet demonstrate a well-defined scientific contribution or clear novelty relative to existing work on EO-driven or multi-objective calibration of degree-day snow models.

The Introduction remains too brief and descriptive. It lists references but does not sufficiently articulate the research gap, justify the approach, or explain how it advances the state of knowledge. This section needs a fundamental rewrite to strengthen conceptual framing and situate the study within the broader literature.

Methodologically, several aspects remain ambiguous or insufficiently justified: the rationale for the MODIS downscaling/reaggregation workflow, the temporal averaging of snowmelt coefficients, the calibration–control procedure, and the regionalisation approach applied to some basins.

While figures have improved, some still add limited insight and require better readability.

At a deeper level, the core scientific question is weakly posed. The study tests whether a spatially variable snowmelt coefficient improves snowmelt representation relative to a fixed, discharge-calibrated coefficient. However, since the spatial coefficients are derived from EO snow data, this result is almost self-evident. Reported improvements are small and within model uncertainty, with no clear evidence of enhanced discharge skill or reduced equifinality.

The discussion remains mostly descriptive, lacking physical interpretation. The authors should explain more clearly why spatial variability in the snowmelt coefficient is physically justified (e.g. slope, aspect, solar radiation) and to what extent the observed patterns reflect structural limitations of LISFLOOD, notably its limited treatment of elevation-dependent precipitation.

Finally, while writing has improved, the manuscript still requires careful language editing for conciseness and fluency.

In summary, the paper addresses an important topic but still needs major revision to clarify its contribution, justify its methods, and deepen its discussion.

Specific Comments

L6–7: Melt factors in degree-day models vary spatially and temporally with topography and energy fluxes. Some models use radiation-indexed melt factors. Please consider mentioning in the introduction section.

L10–11: Round numerical values; excessive decimal precision implies false accuracy.

L12: “Highlighted an effect” is vague—please specify the nature of the effect.
L13–14: Clarify whether the reported change in snow dynamics affected discharge simulations.
L16–25: Modern calibration rarely relies on discharge alone. Multi-objective calibration (snow + discharge) should be acknowledged consistently in the introduction and discussion.
L32–34: Clarify your positioning relative to the wide range of multi-objective calibration approaches (SWE, snow depth, SCA; sequential vs simultaneous; optimisation strategies). The introduction should reflect this diversity and specify your scientific contribution.
L36–40: What specific efforts are referred to here? Please elaborate rather than listing references without explanation.
L38–40: This appears to be the central question of the manuscript. Calibrating snow parameters (including degree-day factors) based on snow-cover products independently of runoff parameters has already been examined (see suggested references). How does the present work differ from those? What is the scientific gap it aims to fill?
L43: The “pixel-by-pixel” basis might represent an original contribution compared with previous studies. This should be emphasised by better contextualisation.
L46–47: Avoid forward references to results in the introduction.
Figure 1: The downscaling–reaggregation workflow (MODIS 500 m → Sentinel 50 m → 1 arcmin) is not clearly justified. Why not directly aggregate daily cloud-free MODIS images from 2000 onward to model scale?
L65: Clarify why a “pixel-wise average over time” is used and which season is excluded from calibration.
L76–77: Given the large number of studies using MODIS snow products, the statement “limiting their utility for hydrological applications” seems overstated. Methods exist to fill cloud gaps, and cloud-free products are available. Regarding spatial resolution, while 500 m may be coarse for some applications, it is adequate for others (e.g. models operating at similar resolution or elevation bands). In your study, since the model runs at 1 arcmin resolution, the MODIS data are degraded anyway: therefore, the limitation is not clearly justified.
L210: Unless I have missed it, the simulation periods were not previously defined. Hence, the term “final period” is difficult to interpret here.
L239: What is meant by “monthly basis”? If the model runs daily, why not compute KGE on a daily basis? What is the rationale for aggregating to monthly? If the intention is to assess the mean seasonal regime, then computing KGE on mean daily values would seem more consistent. Please clarify.
L261–262: Specify which anthropogenic activities (dams, abstractions, transfers) are modelled and how; otherwise, exclude these basins.
L264: I could not find any mention of the regionalisation approach in Section 2.2.1. However, you refer to using several gauging stations per basin for calibration. How many stations were used for each basin? Why do the Guadalfeo and Adige basins require regionalisation? Do they lack discharge observations? If so, why were they retained for model evaluation? These choices are unclear and could confuse readers.
L268–272: Should it be understood that the evaluation of snow dynamics against remotely-sensed products is limited to 2017–2023, whereas discharge simulation with LISFLOOD covers 1992–2023? What is the calibration period? Are there independent evaluation periods? This is presented in a very confuse manner and should be more rigorous. Please clarify.
Figure 2: Please include scale bars on each map to enable comparison of basin areas.
Figures 3–4: The legend for Cm is too small and difficult to read. Since it appears in each panel, please enlarge it and place it outside the Adige panel, at the bottom of the figure. The black-to-yellow colour scheme is not intuitive and poorly represents gradual Cm variations. A red gradient (light to dark, 0–10 mm $°C^{-1}$ day^{-1}) would be potentially clearer. Also, please add basin

outlines to aid spatial interpretation (especially where white areas occur, e.g. Guadalfeo in Fig. 4).
L282: The statement "compensate for erroneous precipitation inputs" raises the question of whether elevation-dependence of precipitation should instead be explicitly treated, as done in other snow and glacier models. Discuss whether Cm compensates for underestimation of high-elevation precipitation rather than true melt variability.
L287: This now clarifies the use of a "temporal average" but very late (see earlier remark). The temporal stability of Cm probably reflects systematic precipitation-forcing errors with elevation. In particular, poorly represented orographic gradients and snow under-catch (due to coarse model resolution and lack of dedicated parameters) likely explain the spatial variability in Cm, which compensates these errors. The influence of slope and aspect appears secondary here.
L295–296: Highlight that this reflects model structural limits, not physical variation in melt.
L296–297: The correlation with elevation is positive for flat basins, but is it statistically significant? Report significance (p-values) for correlations.
L302: If I understand correctly, this L–Cm configuration is not entirely lumped, since it may depend on sub-catchments (see Fig. 3). Please state this clearly.
L386–387: Dashed lines add little information and make the figure harder to read.
Figure 12: This figure contains too many curves and colours, reducing legibility. Avoid dashed lines and retain only three curves per panel (observed; L–Cm; EO–Cm).
L404–409: Including this basin seems questionable given its strong influence from dam operations, which invalidates interpretation of snow-process sensitivity. Consider removing it from the analysis.
Table 6: It is unnecessary to include the subcomponents of KGE here. They are numerous and do not enhance the comparison between L–Cm and EO–Cm.
L428–438: I recommend deleting this section (and the corresponding Figure 13). It adds little to the comparison between L–Cm and EO–Cm and presents low KGE values indicating poor discharge performance.
L439–452: Likewise, I suggest deleting this section (and Figure 14). It is overly technical and does not provide convincing or insightful results for the comparison.
L454: The study aims to test whether spatially variable snowmelt coefficients improve snowmelt representation compared with fixed coefficients calibrated against runoff. Since the spatial coefficients seem preliminary calibrated against remotely-sensed snow cover, the research question seems somewhat trivial, and the answer self-evident.
L455: "Enhances realism of what?" Please specify.
L456: The procedure appears lengthy and computationally demanding (downscaling MODIS with Sentinel, upscaling to 1 arcmin, multi-year spatial calibration of Cm) for an improvement that is debatable. There is no clear evidence that snow cover, SWE, or discharge are better simulated. The added spatial variability in Cm introduces more parameters without clear gains in predictive skill, internal consistency, or uncertainty reduction.
L458–461: It seems that this statement is not supported by the presented results.
L465–471: Other operational models can be applied at similar scales (1,000–10,000 km^2) with greater realism in mountain environments, as they explicitly account for elevation-dependence in meteorological forcings. Compare explicitly with elevation-band models that better treat orographic precipitation.
L473–474: Cite studies showing that SCA-based calibration alone provides limited constraint.
L475: What is meant by "required for a more robust calibration of accumulation processes"? Please clarify and suggest how this could be achieved.
L481–485: Mention other models incorporating solar radiation for variable melt factors.

L488: Clarify whether timing and magnitude changes impacted discharge; results suggest otherwise.
L492–495: Please support this statement with appropriate references.
L497–499: You have not demonstrated improved discharge simulations (higher KGE) or reduced equifinality.
L500: I am also not convinced that internal model consistency has improved. The presented results do not substantiate this claim.
L513–516: Discuss correlation with elevation and relate findings to existing EO-constrained modelling studies.
L539–540: If snow accumulation and melt timing or magnitude were truly affected, one would expect a corresponding influence on runoff timing. How do you explain that discharge performance remains similar across approaches? Are snow dynamics genuinely impacted?
L540–541: This formulation comes from Kirshner (2006) (please cite). Beyond this reference, do you consider that your model achieves the "right answer" (discharge simulations) for the "right reasons" (internally consistent snow dynamics)? If so, why? If not, why not?
L542–543: "… our findings suggest that a sequential calibration approach—first adjusting the snowmelt coefficient using EO-derived SCA, followed by post-calibration of streamflow—can be a viable and potentially more efficient alternative." More efficient compared with what?
L544–545: What concrete improvement does this approach bring? Is the added complexity justified by demonstrable gains in model performance, internal consistency, or uncertainty reduction?
L554–618: Appendix A is rather long and highly technical. The products differ in many subtle ways (spatial resolution, temporal depth, revisit frequency, etc.), which makes a direct comparison difficult to present within an appendix. Given the spatial resolution used in LISFLOOD (≈ 1.4 km), it is unclear why snow cover area data derived from the daily 500 m MOD10A1F MODIS snow product was not used. MOD10A1F is a cloud-gap-filled version of MOD10A1 that provides daily, highly accurate snow-cover estimates from 2000 onward (Hall et al., 2019). I suggest using these products rather than downscaling MODIS with Sentinel data and then upscaling to 1.4 km. Your approach is difficult to justify because it reduces the available temporal depth due to the limited historical record of Sentinel, casts doubt on the ability to gap-fill cloud-contaminated products given Sentinel-2's limited temporal frequency, and does not add a spatial-resolution benefit since snow cover is ultimately aggregated to 1.4 km for your study.
L620–636: This comparison is interesting. The formulation of Zaitchik and Rodell (2009) is clearer and simpler than that of Swenson and Lawrence (2012). It does not require tuning free parameters because τ and SWE_max are fixed values. Given that both formulations produced comparable performance in your tests, you should rather favour the Zaitchik and Rodell (2009) formulation.
L627: change SCFmax to SWEmax.
L640–661: This section is central to the manuscript and should be included in the main paper rather than relegated to an appendix. Equation C2 shows that the computation of EO-Cm depends strongly on the precipitation estimate. Orographic precipitation errors largely explain the spatial variability of the EO-Cm that is ultimately computed.
L646: Please provide the full derivation and all details of the equation to help readers understand the contribution of each component.
Appendices C-D: Appendices C and D are highly technical and are presented in a somewhat confusing manner; they are at times unnecessarily redundant with the main text. As currently presented they create more confusion than clarity about the manuscript's aims. Please remove, simplify and/or clarify these appendices.

L668–671, L672-678, L679–698: Author contributions, competing interests, acknowledgements, and references are misplaced within the appendices. Please move them to their correct locations.

Suggested bibliography on the topic

https://doi.org/10.1029/2010WR009824
https://doi.org/10.5194/hess-21-3325-2017
https://doi.org/10.5194/hess-18-4773-2014
https://doi.org/10.1016/j.jhydrol.2023.129867
http://dx.doi.org/10.1659/MRD-JOURNAL-D-11-00092.1
https://doi.org/10.1029/2011WR010559
https://doi.org/10.1016/j.jhydrol.2024.130820