

March 7th, 2022

Dear Reviewer #2,

We thank you very much for all your valuable comments on our manuscript. We are aware that this review has been a significant time investment and therefore especially appreciate your feedback and commitment. We have addressed all the comments, and the vast majority of them will lead to additions or clarifications in the text. Please find below a detailed answer to all the raised points. All the comments were relevant and well placed and will certainly contribute to increasing the clarity and overall quality of our work.

Best regards,

Francesca Carletti, on behalf of the author team.

- Reviewer comments are repeated in *grey italic*
- Replies from the authors are reported in black regular font
- Quotes from the paper text are reported in *black italic*
- Issues that have been clarified with Reviewer #1 are quoted in *blue italic*
- Proposed additions are reported in *green italic*, unless stated otherwise

This paper addresses an important and interesting topic-comparison of the hydrological models with different level of complexity in the high Alpine catchments. The authors compared two degree-day models and one full energy-balance models in the context of climate change. Overall the paper is well organized and the presentation is good. I suggest a major revision, and there are several issues to be further improved. The comments are as follows:

On Section 1:

Are there any other similar models that could also reach the goals? Why do you decide to select these two models for comparison? I suggest some literature review and explanation could be given in section 1.

Thank you for this suggestion. We agree: Section 1 is not well balanced, as it dwells a lot on the object of study - the model comparison - but hardly at all on the specific choice of these two models.

We propose a new Introduction which takes into account this and the further comments on Section 1. Additions addressing specifically this comment are highlighted in *red italics*.

This section listed many references that are mainly related to the comparisons of the Alpine3D model and the degree-day model. However, there is a lack of the summary of the relation and innovation of this research which differs from the previous studies. Some discussion in more detail on the relevance of the references to the present research are needed. The innovation of this study should be highlighted.

Again, we agree. Section 1 does not sufficiently highlight the innovation of our work compared to previous ones.

In the new Introduction, additions addressing specifically this comment are highlighted in *orange italics*.

One of the aims of the study is “getting a better understanding of the conditions under which one kind of melt scheme and/or hydrological model outperforms the other”. The study only considered two catchments, thus I regard it as a case study. We don’t know the how do the models perform in other cases. I’m concerned that the cases in the research are not strong enough to support the generalization.

The Reviewer is right: the cited sentence generalizes the case study too much and creates ambiguity about the purpose of the paper. However, our impression is that throughout the Discussion, the Climate Change dissertation and the Conclusion (Section 4.2.4, Section 4.3 and Section 5, respectively), no particular concept is generalised, but everything is clearly related to the case study specifically.

Thus, this sentence has been modified in the new Introduction and it is highlighted in *purple italics*.

We propose the following modified Introduction.

The hydrology of high Alpine catchments is dominated by the melt of seasonal snow cover and glaciers, and thus particularly sensitive to climate change¹. The amount of runoff and its seasonal pattern is likely to be heavily modified in the future, impacting ecology, water resources management and the overall quality of life in inhabited areas^{2,3}. Change in summer discharge in Alpine areas will also increase the sensitivity to air temperature, enhancing the warming of Alpine rivers with climate change⁴. Therefore, the development of models reproducing reliable predictions of the response of Alpine catchments discharge to climate change is a crucial step.

Previously, both Degree-Day and Energy-Balance melt models have been implemented to simulate runoff in Alpine catchments^{5,6,7,8,9,10}. Even if these two types of models are different with respect to how the physics is represented, they have proven to give similar results when considering present climatic conditions^{11,12,13,14}. Degree-Day models might be preferred

¹ Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, 438, 303–309, <https://doi.org/10.1038/nature04141>, <https://doi.org/10.1038/nature04141>, 2005.

² Yvon-Durocher, G., Allen, A. P., Montoya, J. M., Trimmer, M., and Woodward, G.: The Temperature Dependence of the Carbon Cycle in Aquatic Ecosystems, 43, 267–313, <https://doi.org/https://doi.org/10.1016/B978-0-12-385005-8.00007-1>, <https://www.sciencedirect.com/science/article/pii/B9780123850058000071>, 2010.

³ Schaeffli, B., Hingray, B., and Musy, A.: Climate change and hydropower production in the Swiss Alps: quantification of potential impacts and related modelling uncertainties, *Hydrology and Earth System Sciences*, 11, 1191–1205, <https://doi.org/10.5194/hess-11-1191-2007>, <https://hess.copernicus.org/articles/11/1191/2007/>, 2007.

⁴ Michel, A., Schaeffli, B., Wever, N., Zekollari, H., Lehning, M., and Huwald, H.: Future water temperature of rivers in Switzerland under climate change investigated with physics-based models, *Hydrology and Earth System Sciences Discussions*, 2021, 1–45, <https://doi.org/10.5194/hess-2021-194>, <https://hess.copernicus.org/preprints/hess-2021-194/>, 2021a.

⁵ Huss, M., Farinotti, D., Bauder, A., and Funk, M.: Modelling runoff from highly glacierized alpine drainage basins in a changing climate, *Hydrological Processes*, 22, 3888 – 3902, <https://doi.org/10.1002/hyp.7055>, 2008.

⁶ Bavay, M., Lehning, M., Jonas, T., and Löwe, H.: Simulations of future snow cover and discharge in Alpine headwater catchments, *Hydrological Processes*, 23, 95–108, <https://doi.org/https://doi.org/10.1002/hyp.7195>, <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.7195>, 2009.

⁷ Magnusson, J., Farinotti, D., Jonas, T., and Bavay, M.: Quantitative evaluation of different hydrological modelling approaches in a partly glacierized Swiss watershed, *Hydrological Processes*, 25, 2071–2084, <https://doi.org/10.1002/hyp.7958>, 2011.

⁸ Zhang, S., Ye, B., Liu, S., Zhang, X., and Hagemann, S.: A modified monthly degree-day model for evaluating glacier runoff changes in China. Part I: model development, *Hydrological Processes*, 26, 1686–1696, <https://doi.org/https://doi.org/10.1002/hyp.8286>, <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.8286>, 2012.

⁹ Farinotti, D., Usselman, S., Huss, M., Bauder, A., and Funk, M.: Runoff evolution in the Swiss Alps: Projections for selected high-alpine catchments based on ENSEMBLES scenarios, *Hydrological Processes*, 26, 1909–1924, <https://doi.org/10.1002/hyp.8276>, 2012.

¹⁰ Gallice, A., Bavay, M., Brauchli, T., Comola, F., Lehning, M., and Huwald, H.: StreamFlow 1.0: an extension to the spatially distributed snow model Alpine3D for hydrological modelling and deterministic stream temperature prediction, *Geoscientific Model Development*, 9, 4491–4519, <https://doi.org/10.5194/gmd-9-4491-2016>, <https://www.geosci-model-dev.net/9/4491/2016/>, 2016.

¹¹ Zappa, M., Pos, F., Strasser, U., Warmerdam, P., and Gurtz, J.: Seasonal water balance of an Alpine Catchment as Evaluated by different methods for spatially distributed snowmelt modelling, *Nordic Hydrology*, 34, 179–202, <https://doi.org/10.2166/nh.2003.0003>, 2003.

¹² Magnusson, J., Farinotti, D., Jonas, T., and Bavay, M.: Quantitative evaluation of different hydrological modelling approaches in a partly glacierized Swiss watershed, *Hydrological Processes*, 25, 2071–2084, <https://doi.org/10.1002/hyp.7958>, 2011.

¹³ Kobierska, F., Jonas, T., Zappa, M., Bavay, M., Magnusson, J., and Bernasconi, S.: Future runoff from a partly glacierized watershed in Central Switzerland: A two-model approach, *Advances in Water Resources*, 55, 204–214, <https://doi.org/10.1016/j.advwatres.2012.07.024>, 2013.

¹⁴ Bavera, D., Bavay, M., Jonas, T., Lehning, M., and De Michele, C.: A comparison between two statistical and a physically-based model in snow water equivalent mapping, *Advances in Water Resources*, 63, 167 – 178, <https://doi.org/https://doi.org/10.1016/j.advwatres.2013.11.011>, <http://www.sciencedirect.com/science/article/pii/S030917081300242X>, 2014.

because they reduce the computational load and require simpler, commonly-available input data¹⁵. However, when considering climate change, the use of such models may be disputable since the value of the calibrated parameters required may change under different climatic conditions^{16,17}. This is particularly relevant for (partly) glacierized catchments, as models have to deal with snow and ice melt under global warming and therefore varying glacier surface. ~~Additionally, land use and weather conditions are highly diverse within any Alpine context and may as well experience future evolution as a consequence of rising temperatures.~~

In this study, three different models are compared: the Degree-Day model Poli-Hydro (PH hereafter) and the process-based model chain Alpine3D+StreamFlow (A3D+SF hereafter), in its full Energy-Balance configuration and with a new hybrid Degree-Day mode. Both models have been used recently to perform climate change studies^{18,19}.

A3D is a good example of a physically-based model that precisely describes many alpine surface processes. As it has been designed from the start for avalanche warning applications²⁰ it must describe the snow metamorphism and microstructure, the snow density, temperature and liquid water content²¹, the liquid water transport in snow²², the liquid water preferential flow²³, the turbulent kinetic energy exchanges at the surface²⁴, and of course, the snow stability²⁵. Besides, in view of its use for avalanche risk forecasting²⁶, it is

¹⁵ Zappa, M., Pos, F., Strasser, U., Warmerdam, P., and Gurtz, J.: Seasonal water balance of an Alpine Catchment as Evaluated by different methods for spatially distributed snowmelt modelling, *Nordic Hydrology*, 34, 179–202, <https://doi.org/10.2166/nh.2003.0003>, 2003.

¹⁶ Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, *Journal of Glaciology*, 45, 101–111, <https://doi.org/10.3189/S0022143000003087>, 1999.

¹⁷ Magnusson, J., Jonas, T., López-Moreno, I., and Lehning, M.: Snow cover response to climate change in a high alpine and half-glacierized basin in Switzerland, *Hydrology Research*, 41, 230–240, <https://doi.org/10.2166/nh.2010.115>, 2010

¹⁸ Michel A, Schaefli B, Wever N, Zekollari H, Lehning M, Huwald H. Future water temperature of rivers in Switzerland under climate change investigated with physics-based models. *Hydrol Earth Syst Sci*. 2022;26(4):1063-1087. doi:10.5194/hess-26-1063-2022

¹⁹ Fuso, F., Casale, F., Giudici, F., and Bocchiola, D.: Future Hydrology of the Cryospheric Driven Lake Como Catchment in Italy under Climate Change Scenarios, *Climate*, 9, <https://doi.org/10.3390/cli9010008>, <https://www.mdpi.com/2225-1154/9/1/8>, 2021.

²⁰ Lehning, M., Völsch, I., Gustafsson, D., Nguyen, T.A., Stähli, M. and Zappa, M. (2006), ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrol. Process.*, 20: 2111-2128. <https://doi.org/10.1002/hyp.6204>

²¹ Köhler, Anselm, Jan-Thomas Fischer, Riccardo Scandroglio, Mathias Bavay, Jim McElwaine, and Betty Sovilla. "Cold-to-warm flow regime transition in snow avalanches." *The Cryosphere* 12, no. 12 (2018): 3759-3774.

²² Wever, N., Comola, F., Bavay, M., and Lehning, M.: Simulating the influence of snow surface processes on soil moisture dynamics and streamflow generation in an alpine catchment, *Hydrol. Earth Syst. Sci.*, 21, 4053–4071, <https://doi.org/10.5194/hess-21-4053-2017>, 2017.

²³ Würzer, S., Wever, N., Juras, R., Lehning, M., and Jonas, T.: Modelling liquid water transport in snow under rain-on-snow conditions – considering preferential flow, *Hydrol. Earth Syst. Sci.*, 21, 1741–1756, <https://doi.org/10.5194/hess-21-1741-2017>, 2017.

²⁴ Schlögl S, Lehning M, Mott R. How Are Turbulent Sensible Heat Fluxes and Snow Melt Rates Affected by a Changing Snow Cover Fraction? *Front Earth Sci* . 2018;6. <https://www.frontiersin.org/article/10.3389/feart.2018.00154>.

²⁵ Richter B, Schweizer J, Rotach MW, van Herwijnen A (2021). Modeling spatially distributed snow instability at a regional scale using Alpine3D. *Journal of Glaciology* 67(266), 1147–1162. <https://doi.org/10.1017/jog.2021.61>

²⁶ Morin, S., S. Horton, F. Techel, M. Bavay, C. Coléou, C. Fierz, A. Gobiet, P. Hagenmuller, M. Lafaysse, M. Ližar, C. Mitterer, F. Monti, K. Müller, M. Olef, J. S. Snook, A. van Herwijnen and V. Vionnet, Application of physical snowpack models in support of operational avalanche hazard forecasting : a status report on current implementations and prospects for the future, *Cold. Reg. Sci. Technol.*, 170, 102910, <https://doi.org/10.1016/j.coldregions.2019.102910>, 2020

constantly being tested during the snow season. The StreamFlow²⁷ distributed hydrological model based on A3D has specifically been designed for alpine catchments with the ability to simulate discharge and streamflow temperatures^{28,29}. Moreover, A3D does not require any calibration and is used "as is" on any new catchment. It has been used in various conditions, from the European Alps to Canada^{30,31}, Antarctica³², Finland³³, Japan^{34,35}, central Asia³⁶. Moreover, the influence of the configuration parameters has been examined³⁷.

SNOWPACK³⁸, the snow physics model running for each cell within A3D, has participated in the ESM-SnowMIP³⁹, the most data-rich Model Intercomparison Project entirely dedicated to snow modelling. In the context of this MIP, a total of twenty-seven models are compared in terms of simulations at five mountain sites, one urban-maritime site and one Arctic site. Among all experiment sites SNOWPACK showed a slightly negative bias for SWE and snow surface temperature, a slightly positive bias for albedo and almost no bias for soil temperature, as representative for the family of multi-layer snow physics models.

On the other hand, PH is a well-assessed model that has been used over a large array of conditions from high-altitude, heavily cryospheric conditions, to low-altitude, arid or semi-arid areas, with or without snow/ice contributions and over catchments of largely varying size

²⁷ Gallice, A., Bavay, M., Brauchli, T., Comola, F., Lehning, M., and Huwald, H.: StreamFlow 1.0: an extension to the spatially distributed snow model Alpine3D for hydrological modelling and deterministic stream temperature prediction, *Geoscientific Model Development*, 9, 4491–4519, <https://doi.org/10.5194/gmd-9-4491-2016>, <https://www.geosci-model-dev.net/9/4491/2016/>, 2016.

²⁸ Gallice, A., Bavay, M., Brauchli, T., Comola, F., Lehning, M., and Huwald, H.: StreamFlow 1.0: an extension to the spatially distributed snow model Alpine3D for hydrological modelling and deterministic stream temperature prediction, *Geosci. Model Dev.*, 9, 4491–4519, <https://doi.org/10.5194/gmd-9-4491-2016>, 2016.

²⁹ Michel A, Schaefli B, Wever N, Zekollari H, Lehning M, Huwald H. Future water temperature of rivers in Switzerland under climate change investigated with physics-based models. *Hydrol Earth Syst Sci.* 2022;26(4):1063-1087. doi:10.5194/hess-26-1063-2022

³⁰ Côté, Kevin & Madore, Jean-Benoit & Langlois, Alex. (2014). EVALUATING THE POTENTIAL OF USING SNOWPACK AND ALPINE3D SIMULATIONS IN THREE CANADIAN MOUNTAIN CLIMATES. 10.13140/2.1.3463.9363.

³¹ Morteza pour, Marzieh & Menounos, Brian & Jackson, Peter & Erler, Andre & Pelto, Ben. (2020). The role of meteorological forcing and snow model complexity in winter glacier mass balance estimation, Columbia River basin, Canada. *Hydrological Processes*. 34. 10.1002/hyp.13929.

³² Wever N, Maksym T, White S, Leonard KC. Ice mass balance data PS81/517 from Weddell Sea, Antarctica, 2013. July 2021. doi:10.1594/PANGAEA.933424

³³ Rasmus, Sirpa & Räisänen, Jouni & Lehning, Michael. (2004). Estimating snow conditions in Finland in the late 21st century using the SNOWPACK model with regional climate scenario data as input. *Annals of Glaciology*. 38. 238-244. 10.3189/172756404781814843.

³⁴ Sato, A., Ishizaka, M., Shimizu, M., Kobayashi, T., Nishimura, K., Nakai, S., . . . Sato, A.(2004). Construction of snow disaster forecasting system in Japan. Fifth international conference on Snow engineering, Vol. 5, p.235-238.

³⁵ Hirashima, Hiroyuki & Nishimura, K. & Baba, Emiko & Hachikubo, Akihiro & Lehning, Michael. (2004). SNOWPACK model simulations for snow in Hokkaido, Japan. *Annals of Glaciology*. 38. 123-129. 10.3189/172756404781815121.

³⁶ Bair, E. H., Rittger, K., Ahmad, J. A., and Chabot, D.: Comparison of modeled snow properties in Afghanistan, Pakistan, and Tajikistan, *The Cryosphere*, 14, 331–347, <https://doi.org/10.5194/tc-14-331-2020>, 2020.

³⁷ Schlögl S, Marty C, Bavay M, Lehning M. Sensitivity of Alpine3D modeled snow cover to modifications in DEM resolution, station coverage and meteorological input quantities. *Environ Model Softw.* 2016;83:387-396. doi:<https://doi.org/10.1016/j.envsoft.2016.02.017>

³⁸ Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., & Zimmerli, M. (1999). SNOWPACK model calculations for avalanche warning based upon a new network of weather and snow stations. *Cold Regions Science and Technology*, 30(1-3), 145-157. [https://doi.org/10.1016/S0165-232x\(99\)00022-1](https://doi.org/10.1016/S0165-232x(99)00022-1)

³⁹ Menard CB, Essery R, Krinner G, et al. Scientific and Human Errors in a Snow Model Intercomparison. *Bull Am Meteorol Soc.* 2021;102(1):E61-E79. doi:10.1175/BAMS-D-19-0329.1

(from ~10 to ~10000 km²) in Italy⁴⁰, Ethiopia⁴¹ and Nepal⁴², with satisfactory accuracy in reproducing stream flows, snow/ice dynamics and cryospheric contributions.

No direct comparisons between PH and other models have yet been pursued. However, the choice of this model as representative for the Degree-Day model family applied to cryospheric studies is motivated by the aforementioned wide range of applications with acceptable results. Moreover, a recent study⁴³ discusses the gain in information obtained when using properly tuned Degree-Day models for snow/ice melt, demonstrating satisfactory results for the meltwater amount and the streamflow tuning.

Model comparisons were performed before on partly glacierized catchments. The work of Magnusson et al.⁴⁴ showed accurate runoff simulations provided by the Energy-Balance model A3D during the snowmelt season and reduced performances during the glacier ice ablation phase. On the other hand, the Degree-Day model (based on an approach proposed by Hock et al.⁴⁵) showed poor performance in reproducing snowmelt, and the simulated total runoff was considerably overestimated during the snowmelt phase. Runoff was accurately reproduced in the ice melt season. However, due to the fact that the study relied upon data from temporary stations in the catchment of maximum 2 years, no long term comparisons were possible.

Kobierska et al.⁴⁶ compared full Energy-Balance Alpine3D runoff predictions with those obtained with the degree-day model PREVAH⁴⁷. Their results showed a lower sensibility of PREVAH to climate change, which was accentuated in summer when glacierized parts of the basin show the highest contribution to runoff. The authors explained this behaviour by considering that Degree-Day models might not perceive the faster seasonal albedo change due to the earlier exposure of glacier ice to solar radiation. For this reason, the absorbed shortwave radiation in the Energy-Balance might be underestimated. However, two completely different model frameworks were used in this study, with differences involving not

⁴⁰ Casale, Francesca & Fuso, Flavia & Giuliani, Matteo & Castelletti, Andrea & Bocchiola, Daniele. (2021). Exploring future vulnerabilities of subalpine Italian regulated lakes under different climate scenarios: bottom-up vs top-down and CMIP5 vs CMIP6. *Journal of Hydrology: Regional Studies*. 38. 100973. 10.1016/j.ejrh.2021.100973.

⁴¹ Bombelli, G.M., Tomiet, S., Bianchi, A., Bocchiola, D., Impact of prospective climate change scenarios upon hydropower potential of Ethiopia in GERD and GIBE Dams. *Water*, 13(5), 716; <https://doi.org/10.3390/w13050716>, 2021.

⁴² Soncini A., Bocchiola, D., Confortola, G., Minora, U., Vuillermoz, E., Salerno, F., Viviano, G., Shrestha, D., Senese, A., Smiraglia, C., Diolaiuti, G., Future hydrological regimes and glacier cover in the Everest region: the case study of the Dudh Koshi basin, *Science of the Total Environment* STOTEN, 565, 1084-1101, 2016. <http://www.sciencedirect.com/science/article/pii/S0048969716310683>

⁴³ Soncini, Andrea & Bocchiola, Daniele & Azzoni, Roberto & Diolaiuti, Guglielmina. (2017). A methodology for monitoring and modeling of high altitude Alpine catchments. *Progress in Physical Geography*. 41. 030913331771083. 10.1177/0309133317710832.

⁴⁴ Magnusson, J., Farinotti, D., Jonas, T., and Bavay, M.: Quantitative evaluation of different hydrological modelling approaches in a partly glacierized Swiss watershed, *Hydrological Processes*, 25, 2071–2084, <https://doi.org/10.1002/hyp.7958>, 2011.

⁴⁵ Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, *Journal of Glaciology*, 45, 101–111, <https://doi.org/10.3189/S0022143000003087>, 1999.

⁴⁶ Kobierska, F., Jonas, T., Zappa, M., Bavay, M., Magnusson, J., and Bernasconi, S.: Future runoff from a partly glacierized watershed in Central Switzerland: A two-model approach, *Advances in Water Resources*, 55, 204–214, <https://doi.org/10.1016/j.advwatres.2012.07.024>, 2013.

⁴⁷ Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R.: An introduction to the hydrological modelling system PREVAH and its pre- and post5 processing-tools, *Environmental Modelling & Software*, 24, 1209 – 1222, <https://doi.org/https://doi.org/10.1016/j.envsoft.2009.04.001>, <http://www.sciencedirect.com/science/article/pii/S1364815209000875>, 2009.

only the melt model. Thereby, it was difficult to ascribe the deviations uniquely to the models' melting scheme.

With this in mind, Shakoor et al.⁴⁸ used A3D to simulate both Energy-Balance and Degree-Day melt schemes on high-altitude, snow-covered Alpine catchments. These experiments allowed to identify uncertainties associated with each melt model and to exclude that differences in reproduced meteorological variables might arise from the use of different data interpolation methods or different set-up of snow vertical profiles (single versus multi-layer). This study showed that an Energy-Balance melt scheme can outperform a Degree-Day approach in the representation of the correct melt dynamics if the former is carefully fed with solid input data sets which are truly representative of the catchment. On the other hand, the Energy-Balance melt scheme showed less accurate performance compared to the Degree-Day one in catchments where data coverage was rather poor and unrepresentative. By distributing surrounding meteorological input data to the catchment, the model generated a few variables (wind speed and long-wave radiation) that were not representative of the catchment's weather, and as a consequence discharge was significantly overestimated.

In this paper, we build upon the work of Shakoor et al. in the sense that A3D will be used to simulate both Energy-Balance and Degree-Day melt schemes, coupled with the hydrological model StreamFlow for discharge computation. Additionally, a spatially semi-distributed Degree-Day model called Poli-Hydro^{49,50} will be used. In the context of this case study, we want to assess how one kind of melt scheme and/or hydrological model outperforms the other, in order to gain a better understanding of the limitations and potential of certain models applied to different river basin types and to corroborate the aforementioned findings. The most important aim that this paper follows compared to previous works is to assess which kind of model might be more appropriate to represent future discharge changes induced by climate change. In fact, the development and identification of suitable models to predict the response of Alpine catchments to climate change is a crucial challenge nowadays. However, it would be simplistic to focus on the melting schemes alone in order to assess models' suitability for climate change studies based on their performances in the present climatic conditions. Thus, a key point of this paper is to assess the relative weight that has to be given to the melt scheme and to the calibration process itself, which might force the model to give realistic results in the current condition but prevent further application under a changing climate.

This study presents the results of hydrological discharge simulation and the major runoff components, i.e. precipitation and snow and glacial melt, with a focus on the melt dynamics. It is performed over two Alpine catchments which differ in size, exploitation and quality of data coverage. The first one is the small, almost-natural Dischma catchment, where many

⁴⁸ Shakoor, A., Burri, A., Bavay, M., Ejaz, N., Ghumman, A. R., Comola, F., and Lehning, M.: Hydrological response of two high altitude Swiss 25 catchments to energy balance and temperature index melt schemes, *Polar Science*, 17, 1–12, <https://doi.org/10.1016/j.polar.2018.06.007>, <https://doi.org/10.1016/j.polar.2018.06.007>, 2018.

⁴⁹ Bocchiola, D., Soncini, A., Senese, A., and Diolaiuti, G.: Modelling Hydrological Components of the Rio Maipo of Chile, and Their Prospective Evolution under Climate Change, *Climate*, 6, <https://doi.org/10.3390/cli6030057>, <https://www.mdpi.com/2225-1154/6/3/57>, 2018.

⁵⁰ Casale, F., Bombelli, G. M., Monti, R., and Bocchiola, D.: Hydropower potential in the Kabul River under climate change scenarios in 35 the XXI century, *Theoretical and Applied Climatology*, 139, 1415–1434, <https://doi.org/10.1007/s00704-019-03052-y>, <https://doi.org/10.1007/s00704-019-03052-y>, 2020.

studies have previously been conducted due to its dense monitoring by means of high-altitude stations in the surroundings^{51,52,53}. The second one is the bigger, transboundary Mera catchment, which originates and partly flows across Switzerland and then stretches to Valchiavenna in Italy. The Mera catchment is approximately 10 times larger than the Dischma catchment and its resources are highly exploited through hydropower operations. Here, meteorological observations and gauging are rather sparse.

The paper is organized as follows. In Section 2, we present the study areas and the available data for model calibration, validation and impact study. Then, in Section 3, models are described in terms of their different melt schemes. In Section 4, calibration results are presented and a model comparison is performed. In the same section, we rate the models' performance in reproducing runoff by means of performance metrics. Finally, in Section 4.3, we discuss models' suitability for climate change impact studies, in terms of melt scheme and relative weight of calibration in the current climatic conditions.

On Section 2:

“68 model chain outputs are provided under three Representative Concentration Pathways: RCP8.5, RCP4.5 and RCP2.6. In this paper, we considered a selected subset of 17 out of the original ensemble”. Do the models you selected in this study outperformed others? Is there any assessment of the historical performance of the GCMs and RCMs before they are selected for the study area? Please explicit the reason why you choose the subset.

Thank you for suggesting this. The following paragraph will be added in Section 2.5.1, explaining why we selected the model chains listed in Tab. 4.

These model chains have been chosen because they capture both high and low climate change signals for air temperature and precipitation. Besides, the same subset of representative model chains has been used by previous studies^{54,55}.

In the report⁵⁵ (Table 10.1), it is indicated for each model chain if it lies in the upper or lower hand or in the middle of all the chains used.

⁵¹ Gallice, A., Bavay, M., Brauchli, T., Comola, F., Lehning, M., and Huwald, H.: StreamFlow 1.0: an extension to the spatially distributed snow model Alpine3D for hydrological modelling and deterministic stream temperature prediction, *Geoscientific Model Development*, 9, 4491–4519, <https://doi.org/10.5194/gmd-9-4491-2016>, <https://www.geosci-model-dev.net/9/4491/2016/>, 2016.

⁵² Wever, N., Comola, F., Bavay, M., and Lehning, M.: Simulating the influence of snow surface processes on soil moisture dynamics and streamflow generation in an alpine catchment, *Hydrology and Earth System Sciences*, 21, 4053–4071, <https://doi.org/10.5194/hess-21-4053-2017>, 2017.

⁵³ Brauchli, T., Trujillo, E., Huwald, H., and Lehning, M.: Influence of Slope-Scale Snowmelt on Catchment Response Simulated With the Alpine3D Model, *Water Resources Research*, 53, 10 723–10 739, <https://doi.org/10.1002/2017WR021278>, 2017

⁵⁴ Epting, J., Michel, A., Affolter, A., and Huggenberger, P. Climate change effects on groundwater recharge and temperatures in Swiss alluvial aquifers. *Journal of Hydrology X*, 11:100071, DOI: 10.1016/j.hydroa.2020.100071, 2021

⁵⁵ CH2018 (2018), CH2018 – Climate Scenarios for Switzerland, Technical Report, National Centre for Climate Services, Zurich, 271 pp. ISBN: 978-3-9525031-4-0

On Section 3:

The the model description part, two models are introduced separately. Since the title is compare the models with different levels of complexity. I think more focus could be paid on the summarizing the overall differences in terms of, for instance, the models structure and modules, hypothesis, parameters and etc. And how the complexity differences are embodied in the models. I think it would be easier for readers to obtain the most important information about the differences of the models.

Thank you for this input. Our opinion is that separated presentation is functional to introduce and describe the characteristics of each model, although we agree with the Reviewer about the fact that a summary of the main differences in terms of structure, hypothesis, parameters should be added to help the reader. We will add the following Table 2 to be recalled in Section 3.1, before each model is presented individually.

Table 2: Summary of the structure of the models and their main characteristics.

	A3D	A3D_{DD}	PH
Spatial resolution	Snow model: 500 m Hydrological model: 100 m		500 m
Temporal resolution	Snow model: 15 minutes Hydrological model: 1 hour		1 day
Snow model	SNOWPACK	SNOWPACK in Degree-Day mode	Melt factors from Degree-Days calibrated monthly
Type of snow model	Full Energy-Balance	Hybrid: Full Energy-Balance with Degree-Day mode	Degree-Day
Snow calibration	No	No	Mera: Yes Dischma: No
Hydrological model	StreamFlow with simple instant routing		Nash approach for IUH propagation ⁵⁶
Type of hydrological model	Semi-distributed		
Calibrated parameters (snow model)	None		Snow Degree-Days
Calibrated parameters (hydrological model)	Maximum infiltration rate Upper reservoir residence time Lower reservoir residence time Fraction of lost water due to deep soil		Exponent for sub-superficial flow Soil permeability Superficial lag time

⁵⁶ Rosso, R.: Nash Model Relation to Horton Order Ratios, Water Resources Research - WATER RESOUR RES, 20, 914–920, <https://doi.org/10.1029/WR020i007p00914>, 1984.

	<i>infiltration</i>	<i>Sub-superficial lag time</i>
--	---------------------	---------------------------------

Please give the equation for the calculation of the statistical scores RMSE, NSE and KGE in this section.

We will add equations for statistical scores in Section 3.2.

On section 4:

Did you do calibration for the A3D? If not, please clarify the reasons. If so, please list the parameters and their ranges for the calibration of the A3D model, and the calibration results for the A3D model.

From Michel et al. (2021)⁵⁷:

"[...] Each Alpine3D simulation is started in July, and the first 3 months serve as spin-up. Before formal parameter calibration in StreamFlow, multiple model runs of Alpine3D are performed with different values of the precipitation vertical lapse rate to adjust the yearly total mass balance in Alpine catchments. In addition, modelled snow heights are compared to measurements to assess the capacity of Alpine3D in reproducing observed snow season dynamics in terms of season duration. Alpine3D has therefore undergone some parameter adjustment but is not calibrated in a strict sense. [...]"

As discussed with Reviewer #1, we agree that this is never clearly stated throughout the paper. Our proposed amendment is quoted here from the discussion with Reviewer #1.

With respect to the modelling chain Alpine3D-Streamflow, snow cover is not calibrated: only the hydrological part is. We certainly agree that this is not clearly specified in Section 4.1.1, undoubtedly leading to confusion or lack of clarity in this respect. We will do our best to cover this lack during the revision process by better-detailing Section 4.1.1 into a Snow Module (new Section 4.1.1.1) and a Flow Propagation Module (new Section 4.1.1.2).

The calibration scores for the PH model listed in table 8 is not ideal, especially in Dischma catchment with only 0.36 measured in NSE. I just wondered how much credit could we give to the models? Though there are analysis for the performance of the model simulation. Could you add the comments on the major contribution for such errors? I strongly recommend adding some references to support the results and it is necessary to make an explanation for the errors. It would be helpful for the readers to interpret the results if the explanation is given.

Thank you for bringing up this point. While dealing with this point, we realised that we reported wrong values of NSE scores for the model PH. We list the correct ones in the

⁵⁷ Michel A, Schaefli B, Wever N, Zekollari H, Lehning M, Huwald H. Future water temperature of rivers in Switzerland under climate change investigated with physics-based models. Hydrol Earth Syst Sci. 2022;26(4):1063-1087. doi:10.5194/hess-26-1063-2022

following Table 3.

Table 3: Updated NSE values for model PH.

Calibration		Validation	
Mera	Dischma	Mera	Dischma
0.67	0.48	0.43	0.71

It is certainly true that the performance of Poli-Hydro in the calibration phase is not ideal for Dischma. A partial explanation is given at the end of Section 4.2.3 in Figure 9 when seasonal statistical performances are analysed. However, we realised that the level of detail is not sufficient and that explanations should have been made even earlier in the face of fairly low statistical performance in the calibration phase. Our proposition is to bring this point up earlier in the text to justify such a low score, addressing which processes are not well captured by the model, and to treat Figure 9 later as validation and not as a new finding.

Our interpretation is that such a low score is largely attributable to the spatial computational resolution of 500 m. Our proposed amendment is reported hereafter from the discussion with Reviewer #1, to be added to Section 2.6.

For the model Alpine3D over the Dischma catchment, the computational resolution was chosen referring to the work of Schlögl et al., 2016⁵⁸. In this paper, the authors test the effects of Alpine3D input variation on Snow Water Equivalent (SWE) quantification, and a big effort is spent on testing different horizontal DEM resolutions. The authors selected four different resolutions (25, 200, 500, 1000 m) for the DEM grid and land cover data. Results show that downscaling from a horizontal resolution of 500 m to one of 25 m, the relative difference in SWE decreases by only 3% approximately. Considering such findings, we decided that this simplification would have been acceptable for the scopes of our paper. Besides, the focus of our paper is finally the estimation of the discharge, and the benefits of a slightly more accurate SWE quantification risk to be lost in the flow routing process – especially during the calibration. In addition, Alpine3D being a complex model, a 100 m resolution over the ~5 times larger Mera catchment, i.e., a computational cost multiplied by 25, is technically not doable. For consistency, a resolution of 500 m was kept in Poli-Hydro as well. Given that we had no previous similar sensitivity studies over the Mera, we tried an alternative calibration there with a resolution of 100 m using the model Poli-Hydro. This has not led to significant improvements in NSE and PBIAS, in the face of considerably higher computational times (See Table 4).

Table 4: Sensitivity analysis on calibration scores over Mera catchment with Poli-Hydro.

	Resolution = 500 m	Resolution = 100 m

⁵⁸ Schlögl, S., Marty, C., Bavay, M., and Lehning, M.: Sensitivity of Alpine3D modeled snow cover to modifications in DEM resolution, station coverage and meteorological input quantities, Environmental Modelling & Software, 83, 387–396, <https://doi.org/https://doi.org/10.1016/j.envsoft.2016.02.017>, <https://www.sciencedirect.com/science/article/pii/S1364815216300378>, 2016.

	<i>NSE</i>	<i>PBIAS</i>	<i>Time</i>	<i>NSE</i>	<i>PBIAS</i>	<i>Time</i>
<i>Mera</i>	0.67	+5.90%	2 minutes	0.65	-4.29%	5 hours

Later on, in Section 4.1.2, the following paragraph will be added.

We are aware that the calibration score obtained with the model PH over the Dischma catchment is not ideal, even though it is still considerably higher than zero thus holding more explanatory power than the time series mean. The explanation we give for this unsatisfactory score is twofold. On the one hand, there is the spatial resolution of the computations. In Section 2.6, several reasons have been listed why a resolution of 500 m might be the best compromise for this case study. Such resolution, which may be functional for full Energy-Balance schemes and/or large catchments, may prove to be not optimal for Degree-Day schemes applied to catchments like Dischma. Indeed, as explained in Section 2.2, Dischma is a small and steep catchment with a very significant altitude range. It could be the case that, contrary to what might happen in a larger and less topographically complex basin like Mera, on the same domain cell the elevation difference could vary a lot. To flatten this difference implies flattening temperature variation, thus snowmelt dynamics, within a scheme (i.e. the Degree-Day) that already flattens temperature variations within the same day. On the other hand, unlike in the case of Mera, due to the lack of MCH/IMIS snow height measurements within the basin, snow is not calibrated over Dischma by the model PH. All hydrological parameters are therefore calibrated with a single objective function (i.e. the measured discharge), so achieving convergence becomes more complex resulting in poorer performance scores. However, we believe that the resolution of 500 m is still adequate for our case study, and despite these drawbacks on the specific case of the Dischma, we believe that it does not invalidate the general findings of our work.

“PHR delays the spring snowmelt-induced discharge by one month compared to observations” Why does the PH reproduce a delayed melt season? It’s noticed that the PH also has a lower snow melt volume. How do could it be explained in terms of model structures, mechanisms and hypothesis differences?

The explanation is twofold. On the one hand, there is the different rain-snow threshold temperature with which the two schemes are normally implemented. This point is explained at the very beginning of Section 4.2.1, with reference to Figure 3:

“[...] In the first place, the two models are implemented with different temperature thresholds for rain-snow separation, 0°C for PH (Fuso et al., 2021⁵⁹) and 1.5°C for A3D and A3D_{DD} (Michel et al., 2021a⁶⁰), as this is the way they are typically used. As a result, PH may simulate more liquid precipitation than A3D in winter which does not accumulate as snow (see Fig. 3). [...]”.

⁵⁹ Fuso, F., Casale, F., Giudici, F., and Bocchiola, D.: Future Hydrology of the Cryospheric Driven Lake Como Catchment in Italy under Climate Change Scenarios, *Climate*, 9, <https://doi.org/10.3390/cli9010008>, <https://www.mdpi.com/2225-1154/9/1/8>, 2021.

⁶⁰ Michel, A., Schaefli, B., Wever, N., Zekollari, H., Lehning, M., and Huwald, H.: Future water temperature of rivers in Switzerland under climate change investigated with physics-based models, *Hydrology and Earth System Sciences Discussions*, 2021, 1–45, 35 <https://doi.org/10.5194/hess-2021-194>, <https://hess.copernicus.org/preprints/hess-2021-194/>, 2021a.

On the other hand, PH is run at a daily resolution, whereas both versions of A3D are run sub-daily. This means that PH relies on a melting scheme that only considers the average daily temperature, which has repercussions on the melt dynamics, as explained later in Section 4.2.4.

“[...] A key point regarding the higher performances of both A3D versions on Dischma compared to PH is the higher temporal resolution at which the energy balance is solved. The melt scheme used by PH is based on the mean daily temperature, which means that if the mean is lower than the melting threshold, the model does not simulate any snowmelt, whereas temperatures might well reach higher values during the daytime and melting could happen instead. As a consequence, the melt during the spring season is delayed in PH. [...]”.

In the figure 7 and 8, It seems that the performance of the A3D and A3Ddd is very close to each other, although a simpler melt-factor energy balance mode is applied in the A3Ddd. Could it be interpreted as the differences of energy balance modules for the A3D model does not have a significant effect in simulating the runoff?

To cover this point, we'll add the following Figures (S1 and S2) to the Supplementary Material, not to overwhelm the already long paper.

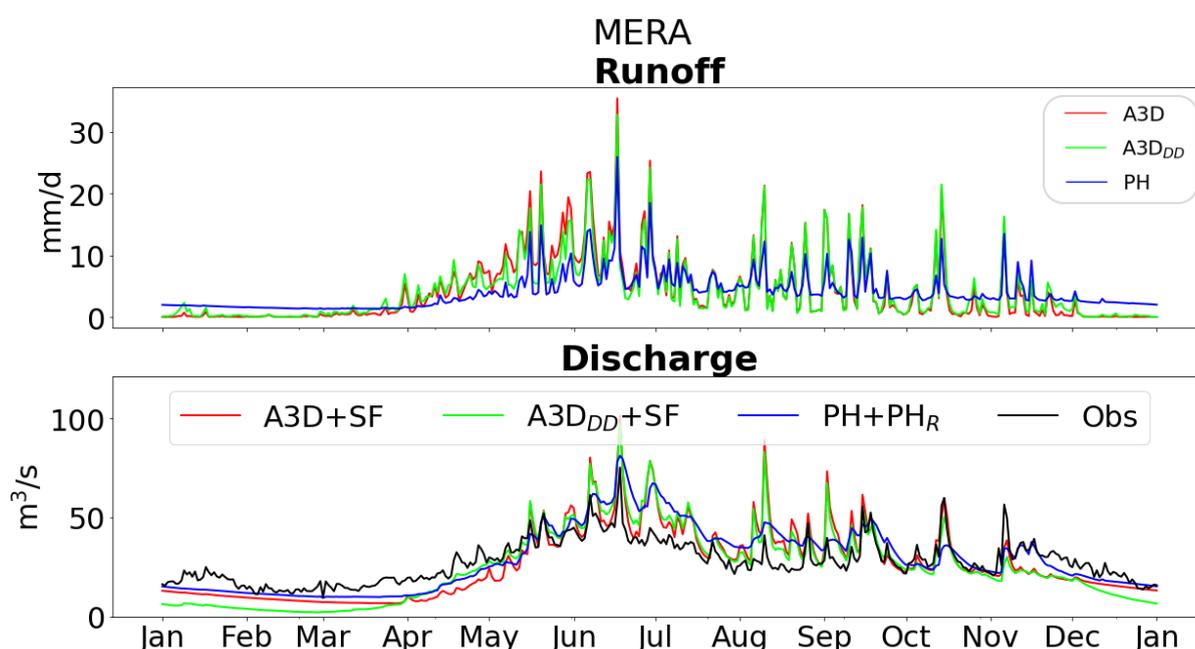


Figure S1: Runoff, as of before the hydrological model routing (top) and Discharge, as of after the hydrological model routing (bottom) predicted by models A3D/A3D+SF (red lines), A3D_{DD}/A3D_{DD}+SF (green lines) and PH/PH+PH_R (blue lines) versus observations (black line) on the validation period (2014-2018) over Mera catchment.

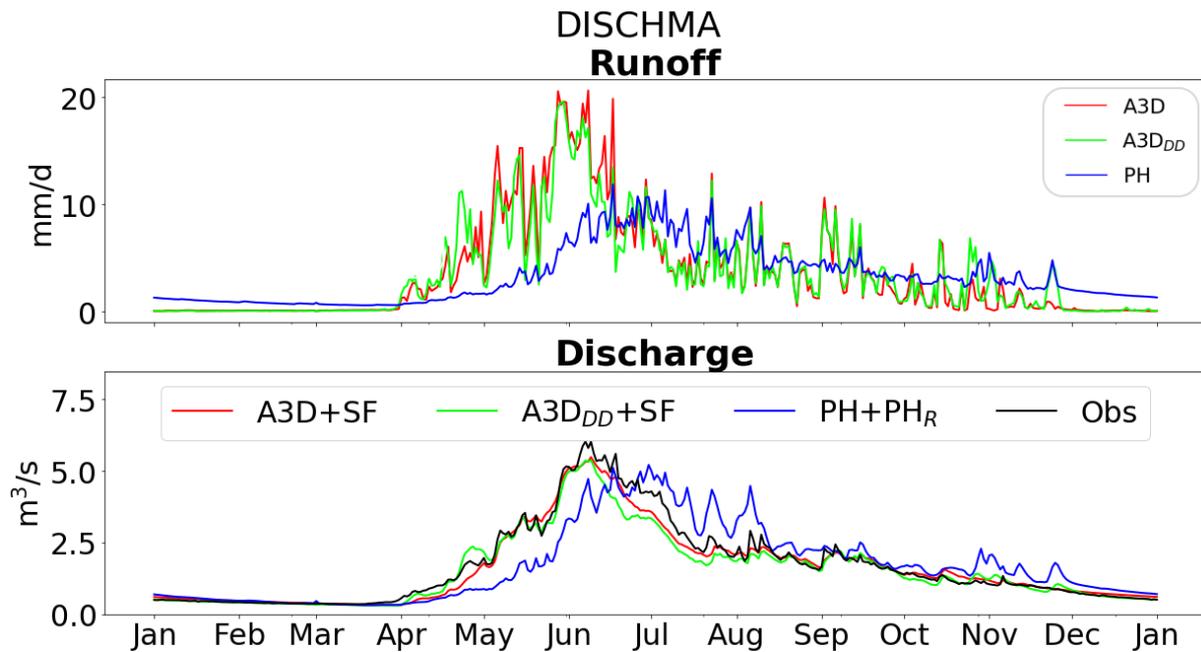


Figure S2: Runoff, as of before the hydrological model routing (top) and Discharge, as of after the hydrological model routing (bottom) predicted by models A3D/A3D+SF (red lines), A3D_{DD}/A3D_{DD}+SF (green lines) and PH/PH+PH_R (blue lines) versus observations (black line) on the validation period (2014-2018) over Dischma catchment.

Figure S1 and S2 show the runoff (i.e. the output from the models before the hydrological routing) and the discharge (i.e. after the hydrological routing) for Mera and Dischma respectively. In both cases, runoff and discharge predicted by A3D and its hybrid Degree-Day version A3D_{DD} are very similar to each other. If, on the one hand, it is true that A3D_{DD} computes the energy entering the snowpack with a simplified Degree-Day approach, on the other hand, the start-up of this hybrid mode still depends upon a full, multi-layer Energy-Balance scheme (i.e. when water and ice are coexisting in the snow element and air temperature is greater than snow surface temperature). Thus, A3D_{DD} is just performing a simplified computation, but only after benefiting from the full complexity of A3D.

Please add some interpretation of the α , β and components of the KGE scores in table 9.

The following paragraph will be added in Section 4.2.4.

KGE components r , α and β express errors in correlation, variability and mean, respectively.

Over the Mera, PH exhibits a better linear correlation with observed values, a slightly lower variability error and a slightly higher mean error with respect to both versions of A3D. The lower variability error can be explained by the less-accentuated snowmelt-generated discharge; whereas the higher mean error is likely due to the slightly higher baseflow simulated by PH, which may well be induced by calibration (S1). On the other hand, over the Dischma, the linear correlation modelled by PH is poorer with respect to both versions of A3D, and the mean error is higher, with variability error not changing significantly among models. Errors for correlation and mean are explained by the shifted discharge curve (S2) induced by the poorer accuracy of snowmelt simulation. Both versions of A3D have very similar values of correlation, variability error and bias.

The discussion part is suggested to be in a new section after all the results are listed.

We will list the Climate Change results in Section 4 and dedicate a new section (Section 5) entirely to the discussion, divided between present and Climate Change conditions.

On section 4:

The errors here are attributed to the dams, “ The explanation is twofold. First, the Mera catchment is highly regulated by dams, which is not accounted for in the models.” However, in section 2.3.2 you mentioned “Discharge modeling here may be slightly disturbed by hydropower regulation... However, at the daily scale and at longer time scales, streamflows are not largely disturbed overall, and hydrological modeling exercise provides acceptable results.”I think the arguments are controversial. Besides in the conclusion part, you also emphasized the effect of reservoir regulation on the discharge simulation. As far as I see, the impact of hydropower regulation could not be easily neglected for this study.

We agree that this sentence in Section 2.3.2:

“[...] However, at the daily scale and at longer time scales, streamflows are not largely disturbed overall, and hydrological modelling exercise provides acceptable results (Fuso et al., 2021). [...]”

is inaccurate and slightly contradictory to our conclusions. We will delete line 5 to line 9 of Section 2.3.2 and we will dedicate a new Section (2.3.3) to this matter, with the due literature review to justify our choices.

2.3.3 River regulation over Mera catchment

As mentioned in Section 2.1, this area is exploited quite intensely by a complex system of reservoirs and hydroelectric power stations (Tab. 1). As a consequence, discharge modelling here may be disturbed due to regulation: specifically, there is a shift in volumes at the hourly scale on working days (Monday to Friday). The problem of regulation data availability over Valchiavenna has already been addressed directly or indirectly in the literature^{61,62,63}. In the work of Giudici et al., dams and plants of Valchiavenna are simply not considered. In the work of Fuso et al., the authors underline overall agreement between observed and modelled discharge at the monthly scale, which deteriorates at the daily scale, as a result of regulation. The Climate Lab of Politecnico di Milano has made several attempts at trying to model reservoir operations in Valchiavenna, but given the impossibility of verifying the assumptions made about reservoir management, the results have never been published. Given these points, and considering the main obstacle of the lack of data in this respect, we decided to neglect the influence of reservoir operation on discharge modelling of the Mera

⁶¹ Fuso, F., Casale, F., Giudici, F., and Bocchiola, D.: Future Hydrology of the Cryospheric Driven Lake Como Catchment in Italy under Climate Change Scenarios, *Climate*, 9, <https://doi.org/10.3390/cli9010008>, <https://www.mdpi.com/2225-1154/9/1/8>, 2021.

⁶² Giudici F, Anghileri D, Castelletti A, Burlando P. Descriptive or normative: How does reservoir operations modeling influence hydrological simulations under climate change? *J Hydrol.* 2021;595:125996. doi:<https://doi.org/10.1016/j.jhydrol.2021.125996>

⁶³ Maruffi L, Stucchi L, Casale F, Bocchiola D. Soil erosion and sediment transport under climate change for Mera River, in Italian Alps of Valchiavenna. *Sci Total Environ.* 2022;806:150651. doi:<https://doi.org/10.1016/j.scitotenv.2021.150651>

river.

It's interesting to notice that on average the peak of snow melt and discharges in RCP2.6 is higher than those in RCP8.5. With higher temperature increase in RCP 8.5, what makes the peaks of the discharge and snow melt being less?

Thank you for pointing this out, we should have specified it. Our explanation is that, with increasing temperatures, it is likely that more precipitation will fall as rain instead of snow. As a consequence, snow will only accumulate at high to very high elevations (where low temperatures may even reduce or slow down the melting), with mid-to-high elevations experiencing considerably less snowfall and snow accumulation, thus less snowmelt and snowmelt-induced discharge.

On section 5:

“Our interpretation is that the calibration process for strongly regulated catchments as Mera overshadows the benefits of a full energy balance scheme showing good performances in reproducing snow melt.” Maybe it's true in this case that the calibration offset the errors from regulation to some extent. But I think as the conclusion it is more important to know implication from the study. In which case the calibration could overshadow the benefit of the physical scheme? Could benefit from the calibration also be applicable under climate change scenarios, and what is the limitation of the models through the comparison?

We propose a new Conclusion that takes into account these comments on Section 5.

This paper compares the discharge response of two Alpine catchments to present conditions and climate change, predicted by one Energy-Balance, a hybrid Degree-Day version of an Energy-Balance, and one Degree-Day melt model: A3D, A3D_{DD} and PH respectively. The two catchments of this case study, Mera and Dischma, are different in size, data availability and extent of water resources exploitation by human activities.

Under current climatic conditions, both the full Energy-Balance and the Degree-Day versions of A3D outperform PH in reproducing the melt dynamics, especially over the almost-natural, nivo-glacial Dischma catchment, where snowmelt is severely underestimated by PH. Over the Mera catchment, monthly volumes are underestimated in winter and overestimated in summer by all models, suggesting that regardless of the melt scheme, hydropower operations can reduce models' discharge predicting capacity. The superiority of both versions of A3D compared to PH is particularly evident when analyzing snow depth and spatial distribution. In terms of predicted discharge, seasonal performance scores over the entire validation period don't show a significant difference between models for Mera, with scores being satisfactory but not outstanding. The explanation is twofold. On the one hand, flow regulation might alter monthly volumes relatively, but the impact on daily flow regimes is certainly heavy, thus hindering each model and melt scheme in reaching high scores at all. On the other hand, data scarcity over Mera might be more problematic for the more complex Energy-Balance approach, which may explain why A3D does not outperform the simpler melt scheme there. Conversely, performance metrics over the well-gauged, almost-natural Dischma catchment show better performance for both versions of A3D+SF over PH_R. Seasonal scores, however, show that both versions of A3D+SF chain outperform PH_R in about all seasons and all catchments. Interestingly, in terms of snowmelt

magnitude/seasonality and discharge, results from the Degree-Day version of A3D+SF are very similar to those obtained from its full energy balance one. However, the scheme A3D_{DD} is only enabled at the melting onset, which is still computed by a full Energy-Balance model. This implementation cannot be compared to a simpler Degree-Day model as PH, which lacks A3D's predicting capacity but brings desirable advantages such as reduced input detail and computational load. However, A3D_{DD} also carries the advantage of being a simplification of a multi-layer snow model in the first place.

Under climate change, end-of-century changes in snowmelt seasonality predicted by A3D and PH are qualitatively the same: a net increase in spring and winter, a net decrease in summer and autumn. However, A3D's melt scheme appears to be more sensitive to climate change than PH's, as the discharge curve predicted by A3D+SF is shifted by one month under RCP8.5 scenario. Likely, the use of a degree-day melt scheme like PH for climate change studies is not suitable, since (1) fixed monthly degree-days compromise the model's ability to perceive seasonal changes in snowmelt and (2) albedo changes cannot be captured, thus the contribution of net shortwave radiation might be underestimated. Such results are consistent with previous climate change studies.

The newest finding of this paper is brought to light when analyzing the predicted discharge for Mera catchment under climate change, as both models and melt schemes substantially fail in reproducing the base flow there. The same behaviour is not observed for the almost-natural Dischma catchment, and the analysis of precipitation input and considerations about evapotranspiration allowed us to exclude other possible influences. Our interpretation is that over Mera, the calibration process didn't only parametrise fixed physical properties of the basin (which are not supposed to change significantly over time), but also anthropogenic disturbances. Such disturbances are likely absorbed in the calibration generating overfitting, so that as soon as the conditions are altered (i.e. under Climate Change), the modelled parametrization fails. Thus, we conclude that the calibration process for strongly regulated catchments as Mera might even overshadow the benefits of a full Energy-Balance scheme, but with the result that the obtained parametrization could be of no use under changed conditions like Climate Change. Moreover, we believe that there would be no interest in applying any benefit deriving from the calibration to Climate Change impact studies because there is no certainty whether they would still be valid in the future, whereas the physics is certainly not expected to change.

The greatest limitation of this model comparison case study is certainly data scarcity over the Mera catchment. A denser monitoring network for input meteorological data would have likely contributed to more accurate results. Dams operation data, if available, could have been used to validate assumptions of water retention and release, to couple PH with energy production plans as has been done in previous studies. On the other hand, for SF such implementation is not yet available. Despite such limitations, the model comparison was still possible and brought to light many interesting aspects for future developments on the modelling and monitoring of Alpine catchments highly exploited for hydropower production in the context of a changing climate.