A comparison of hydrological models with different level of complexity in Alpine regions in the context of climate change

Francesca Carletti[1], Adrien Michel[1,2], Francesca Casale[3], Daniele Bocchiola[3], Michael Lehning[1,2], and Mathias Bavay[1]

[1]WSL Institute for Snow and Avalanche Research (SLF), Davos, Switzerland
[2]School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[3]Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy

Correspondence: F. Carletti (fcarletti94@gmail.com)

Abstract. This study compares the ability of two degree-day models (Poli-Hydro and a degree-day implementation of Alpine3D) and one full energy-balance melt model (Alpine3D) to predict the discharge on two partly glacierized Alpine catchments of different size and intensity of exploitation, under present conditions and climate change as projected at the end of the century. For present climate, the magnitude of snow melt predicted by Poli-Hydro is sensibly lower than the one predicted by the other

5 melt schemes, and the melting season is delayed by one month. This difference can be explained by the combined effect of the reduced complexity of the melting scheme and the reduced computational temporal resolution. The degree-day implementation of Alpine3D reproduces a melt season closer to the one obtained with its full solver; in fact, the onset of the degree-day mode still depends upon the full energy-balance solver, thus not bringing any particular benefit in terms of inputs and computational load, unlike with Poli-Hydro. Under climate change conditions, Alpine3D is more sensitive than Poli-Hydro, reproducing dis-

10 charge curves and volumes shifted by one month earlier as a consequence of the earlier onset of snow melt. Despite their benefits, the coarser temporal computational resolution and the fixed monthly degree-days of simpler melt models like Poli-Hydro make them controversial to use for climate change applications with respect to energy-balance ones. Nevertheless, under strong river regulation, the influence of calibration might even overshadow the benefits of a full energy-balance scheme.

 Keywords: Climate change, alpine regions, alpine hydrology, hydrological models, degree-day models, energy-balance
15 models

1 Introduction

The hydrology of high Alpine catchments is dominated by the melt of seasonal snow cover and glaciers, and thus particularly sensitive to climate change (Barnett et al., 2005). 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 (Yvon-Durocher

20 et al., 2010; Schaefli et al., 2007). Change in summer discharge in Alpine areas will also increase the sensitivity to air temperature, enhancing the warming of Alpine rivers with climate change (Michel et al., 2021a). 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 (Huss et al., 2008; Bavay et al., 2009; Magnusson et al., 2011; Zhang et al., 2012; Farinotti et al., 2012; Gallice et al., 2016). 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 (Zappa et al., 2003; Magnusson et al., 2011; Kobierska et al., 2013; Bavera et al., 2014). Degree-day models might be preferred because they reduce the computational load and require simpler, commonly-available input data (Zappa et al., 2003). However, when considering climate change, the use of such models may be disputable since the value of the calibrated parameters required by this type of models may change under different climatic conditions (Hock, 2005; Magnusson et al., 2010). 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.

Model comparisons were performed before on partly glacierized catchments. The work of Magnusson et al. (2011) showed accurate runoff simulations provided by the energy-balance model Alpine3D (Lehning et al., 2006) during the snow melt season and reduced performances during the glacier ice ablation phase. On the other hand, the degree-day model (based on a distributed degree-day approach proposed by Hock (1999)) showed poor performance in reproducing snow melt, and the simulated total runoff was considerably overestimated during the snow melt phase. However, runoff was accurately reproduced in the ice melt season.

Kobierska et al. (2013) compared Alpine3D runoff predictions with those obtained with the degree-day model PREVAH (Viviroli et al., 2009). Their results showed a lower sensibility of PREVAH to climate change, which was accentuated in summer, when glacierized parts of the basin show 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 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. (2018) used Alpine3D as a single distributed model 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 for 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 few variables (wind speed and long-wave radiation) which were not representative for the catchment's weather, and as a consequence discharge was significantly overestimated.

In this paper, we build upon the work of Shakoor et al. (2018) in the sense that Alpine3D will be used to simulate both energy-balance and degree-day melt schemes, coupled with the hydrological model StreamFlow (Gallice et al., 2016) for discharge computation. Additionally, a spatially semi-distributed degree-day model called Poli-Hydro (Bocchiola et al., 2018; Casale et al., 2020) will be used. With this, we aim at getting a better understanding of the conditions under which one kind

5 of melt scheme and/or hydrological model outperforms the other, to corroborate the aforementioned findings. An additional aim of this paper 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 of nowadays. However, it would be simplistic to focus on the melting schemes alone in order to assess models' suitability for climate change studies basing on their performances in the present climatic conditions.

10 Thus, another 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 changed climate.

The study is performed for two Alpine catchments which are substantially different in size and quality of data coverage. The first one is the small, almost-natural Dischma catchment, where many studies have previously been conducted, e.g. Gallice

15 et al. (2016); Wever et al. (2017); Brauchli et al. (2017). The second one is the bigger, trans-boundary 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. Conversely, the small, natural Dischma catchment is densely monitored by means of a large number of meteorological high-altitude stations in the surroundings.

20 This paper presents the results of hydrological discharge simulation and its major components, i.e. precipitation and snow and glacial melt, with a focus on the melt dynamics. It is organized as follows. In Section 2, we present the study areas and the available data for model calibration, validation and climate change. 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 NSE and KGE. Finally, in Section 4.3, we discuss

25 models' suitability for climate change impact studies, in terms of melt scheme and relative weight of calibration in the current climatic conditions.

2 Catchments and data

2.1 Mera

The Mera catchment (Fig. 1) nests the homonymous river flowing across Switzerland and Italy. The source is located near Piz

30 Mungiroi, canton Grisons, Switzerland, in the relatively dry Central Alps. From its source, it flows east towards Maloja-pass, then turns west through Val Bregaglia flowing into Lobbia dam reservoir and crosses the border to Italy in Castasegna. Mera ends in Lake Como, located into the pre-Alpine region of Italy. The basin spans across a surface of 560 km^2, 3% of which is glacierized. The length of the river is 50 km, 24 of which flow on Swiss territory. Elevation ranges from 3217 to 202 m a.s.l.
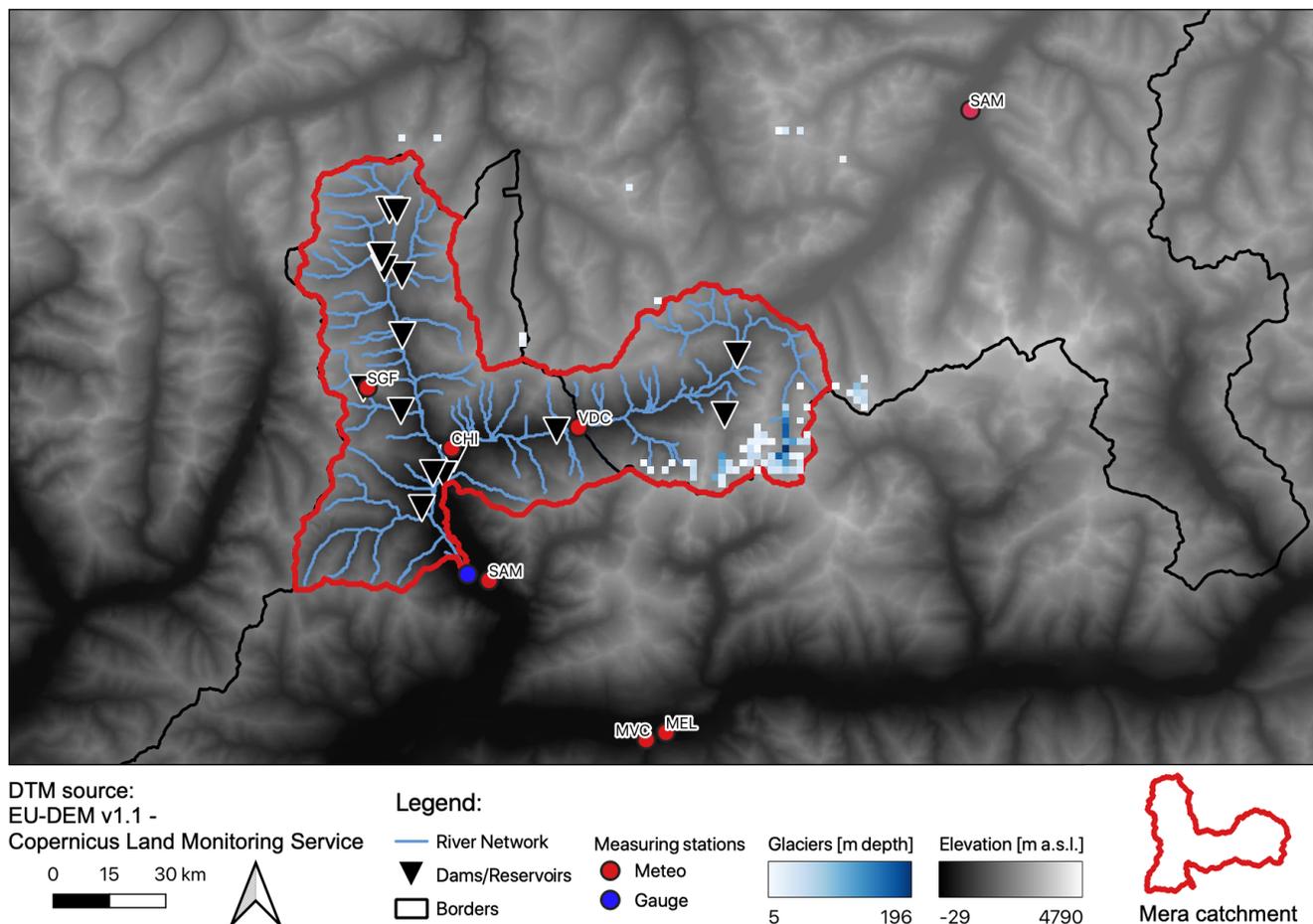
Figure 1. Mera catchment.

The area is particularly important for hydropower production, and its water resources are largely exploited. Hydropower plants and dams are sketched and described in Fig. 1 and Tab. 1.

2.2 Dischma

Dischmabach is a steep, high-Alpine catchment located in the eastern part of the Swiss Alps (Fig. 2). It has two equal source
5 streams whose source areas are Scalettapass and Fuorcla da Grialetsch. Both source streams unite at Dürrboden; the Dischmabach then flows 15 km north-west into the Landwasser in Davos Dorf. The climatological region of the catchment is in the transition zone between the wet Northern Alps and the dry Central Alps. The aspect is S/SE to N/NW. The catchment's surface is 43.3 km^2 including a small glacierized part of 2.1% of the basin's extension. Elevation ranges from 1668 to 3146 m a.s.l. with a mean altitude of 2378 m a.s.l.. In the lowest part, snow accumulates between November and February, whereas in
10 the highest part the accumulation season ends in April (Zappa et al., 2003).

Table 1. Description of hydropower plants and dams exploiting Mera water resources. Stream Liro is one of the main tributaries of river Mera and joins the latter in Prata Camportaccio.

Hydropower plants							
Name	**Lat** [°]	**Lon** [°]	**Elevation** [m.a.s.l.]	**Type**	**Plant capacity** [MW]	**Hydraulic drop** [m]	**Discharge concession** [m3/s]
Isolato Spluga	46.443	9.335	1255	Reservoir	42.8	641.6	0.97
Isolato Madesimo	46.444	9.336	1255	Basin	16.2	275.8	1.4
Prestone	46.393	9.353	1080	Basin	23.5	184.9	5.18
San Bernardo	46.345	9.35	1060	Reservoir	34.2	1014	0.64
Lobbia	46.375	9.659	1441	Basin	-	-	-
Chiavenna	46.314	9.397	280	Basin	66.9	329.8	6.83
Prata Camportaccio	46.304	9.388	280	Streamflow	3.3	18.76	6.83
Mese	46.305	9.378	286	Basin	172.6	769.45	8.4
Gordona	46.283	9.367	245	Streamflow	14.6	36.45	15.65
Dams							
Name	**Lat** [°]	**Lon** [°]	**Elevation** [m.a.s.l.]	**River**	**Hydropower plant of reference**		
Cardanello	46.473	9.344	1904	Liro	Isolato Spluga		
Stuetta	46.472	9.351	1897	Liro	Isolato Spluga		
Madesimo	46.431	9.354	1533	Liro	Isolato Madesimo		
Isolato	46.436	9.338	1243	Liro	Prestone		
Prestone	46.393	9.353	1080	Liro	Mese		
Truzzo	46.36	9.316	2078	Liro	San Bernardo, Mese		
Albigna	46.337	9.646	2155	Mera (CH)	Lobbia		
Villa di Chiavenna	46.33	9.492	635	Mera (IT)	Chiavenna, Prata Camportaccio		

2.3 Hydrological data

2.3.1 FOEN data

The discharge gauge of Kriegsmatte is the only measuring station within the Dischma catchment. The gauge is located at 1668 m a.s.l. on the right side of the stream. Data monitoring started in 1963. Quality controlled discharge at hourly resolution is provided by the Swiss Federal Office for the Environment (FOEN, 2018).
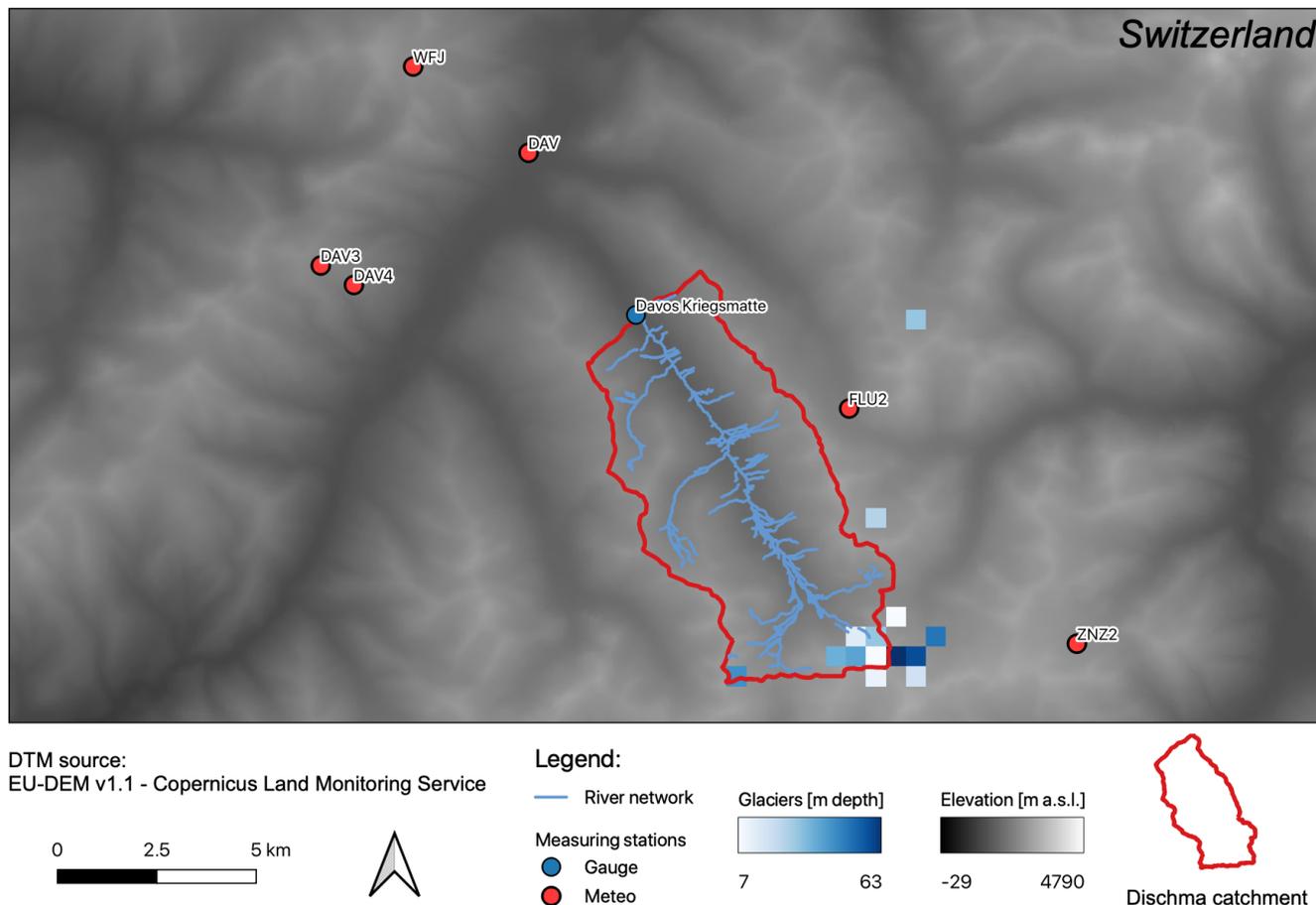
Figure 2. Dischma catchment.

2.3.2 Adda Consortium data

Despite its size, discharge data for the Mera catchment are only available at Samolaco, close to Mezzola Lake, in one of the final sections of the river. Discharge data are available from the Adda Consortium (https://www.addaconsorzio.it/) but with a fee. The Consortium provides the hydrometric level and the rating curve from which discharge is evaluated. Data are available at semi-hourly scale and eventually aggregated at daily scale. Discharge modeling here may be slightly disturbed by hydropower regulation: specifically, there is a shift in volumes at the hourly scale on working days (Monday to Friday). However, at the daily scale and at longer time scales, streamflows are not largely disturbed overall, and hydrological modeling exercise provides acceptable results (Fuso et al., 2021). Unfortunately, due to their difficult availability, there is no specific data on water use by plants of Valchiavenna.

Table 2. Details about Swiss meteo stations (IMIS and MCH). The first column of the table indicates the associated river (D and M for Dischma and Mera, respectively). Contrary to ARPA in Tab. 3, all variables started to be recorded automatically in the same year 1981.

Stream	Station	Abbreviation	Network	Lat [°]	Lon [°]	Elevation (m.a.s.l.)	Data availability	Variables
D	Hanengretji	DAV3	IMIS	46.788	9.774	2455	1998-2018	TA, PSUM, RH, VW
D	Frauentobel	DAV4	IMIS	46.784	9.785	2330	2002-2018	TA, PSUM, RH, VW
D	Flüelapass	FLU2	IMIS	46.753	9.946	2394	2003-2018	TA, PSUM, RH, VW
D	Pülschezza	ZNZ2	IMIS	10.019	46.697	2677	2002-2018	TA, PSUM, RH, VW
D	Davos	DAV	MCH	9.844	46.813	1594	1981-2008	TA, PSUM, RH, ISWR, VW
D	Weissfluhjoch	WFJ	MCH	9.806	46.833	2691	1981-2018	TA, PSUM, RH, ISWR, VW
M	Samedan	SAM	MCH	46.533	9.867	1708	1981-2018	TA, PSUM, RH, ISWR, VW

2.4 Meteorological data

2.4.1 MeteoSwiss stations

Meteorological data over Switzerland are partly acquired from the MeteoSwiss (MCH) automatic monitoring network (IDAWEB, 2019). Since no MCH station is installed within the Dischma catchment itself, the nearby stations of Davos (DAV) and Weiss-

5   fluhjoch (WFJ) provide daily values for air temperature (TA), precipitation (PSUM), wind velocity (VW), relative humidity (RH) and incoming shortwave solar radiation (ISWR). The same variables, except for PSUM, are used for the Swiss side of the Mera catchment from the station Samedan (SAM) in the upper Engadine.

2.4.2 IMIS stations

The second part of the meteorological data used for Switzerland comes from the Intercantonal Measurement and Information

10   System (IMIS) station network (IMIS, 2019), operated by the WSL Institute for Snow and Avalanche research (SLF). The IMIS station network does not provide ISWR measures. Additionally, IMIS rain gauges are not heated, therefore winter precipitation is obtained indirectly based on snow depth variations and snow settling computed by the SNOWPACK model (**?**). The information used in this study comes from the IMIS stations located in Davos (DAV), Flüelapass (FLU) and Zernez (ZNZ) (see Fig. 2). All data coming from Swiss stations (IMIS, MCH) are presented in Tab. 2.

15   ### 2.4.3 ARPA Lombardia stations

The first automatic meteorological stations was installed in the Italian region of Lombardy only in the late 80s. Thus, Lombardy can only count on observational time series which, at most, date back to those years. Nowadays, ARPA counts on a network of

Table 3. Details about Italian meteo stations (ARPA). The * symbol next to some of the available variables indicates that such variables started to be recorded later than the first recorded year for TA and PSUM.

Station	Abbreviation	Lat [°]	Lon [°]	Elevation [m a.s.l.]	Data availability	Variables
Chiavenna Via Cerletti	CHI	46.321	9.395	333	1994-2009	TA, PSUM, RH*
Morbegno Eliporto	MEL	46.136	9.584	230	2016-2018	TA, PSUM, RH, HS, ISWR, VW
Morbegno Via Cortivacci	MVC	46.132	9.566	262	1999-2018	TA, PSUM, RH*, ISWR*, VW*
Samolaco	SML	46.236	9.427	206	1995-2018	TA, PSUM, RH, ISWR*, VW*
San Giacomo Filippo	SGF	46.361	9.319	2064	2005-2018	TA, PSUM, RH*, VW*
Villa di Chiavenna	VDC	46.332	9.512	665	2003-2018	TA, PSUM

318 meteorological stations, but the vast majority of them was only installed in the last two decades. As a consequence, there are only few and sparse robust observational time series of the past 30 years in Valchiavenna. Moreover, historical time series often contain a massive amount of missing data, due to interruption. The stations used in this work are presented in Tab. 3 and locations are shown in Fig. 1.

2.5 Climate change data

2.5.1 CH2018 data

The climate change section of this paper depends upon the CH2018 climate change scenarios (MeteoSwiss, 2018). These scenarios were derived from the new EURO-CORDEX ensemble of climate simulations with Regional Climate Models (RCMs) (Jacob et al., 2014; Kotlarski et al., 2014). The RCM simulations of EURO-CORDEX were performed using a common model domain centered on western Europe. RCMs dynamically downscale coarser global projections (from Global Climate Models, GCMs) into a resolution which is more suitable to represent the topography of Switzerland. However, RCM simulations from EURO-CORDEX do provide information at a spatial resolution of 0.11 or 0.44° km, which is still too coarse for local impact assessments and might lead to biases, especially for complex orographies. CH2018 applied quantile mapping for spatial downscaling, to both observations at station's scale and gridded observations at 2 km to derive site-specific projections. Thus, CH2018 can provide projections at daily resolution for the MCH automatic weather station network. Available variables are the mean, minimum and maximum near-surface air temperature, relative humidity, wind speed, cumulative precipitation and incoming shortwave radiation. 68 model chain outputs are provided under three Representative Concentration Pathways:

Table 4. Details about the considered climate models. GCM = Global Climate Model; RCM = Regional Climate Model; RCP = Representative Concentration Pathway.

Institute	GCM	RCM	Resolution	RCP
CLMCOM	HADGEM	CCLM5	EUR44	8.5
DMI	ECEARTH	HIRHAM	EUR11	2.6, 4.5, 8.5
KNMI	ECEARTH	RACMO	EUR44	4.5, 8.5
KNMI	HADGEM	RACMO	EUR44	2.6, 4.5, 8.5
SMHI	CCCMA	RCA	EUR44	4.5, 8.5
SMHI	MIROC	RCA	EUR44	2.6, 4.5, 8.5
SMHI	MPIESM	RCA	EUR44	2.6, 4.5, 8.5

RCP8.5, RCP4.5 and RCP2.6. In this paper, we considered a selected subset of 17 out of the original ensemble, resumed in Tab. 4. The subset was selected to represent the spread between models but to limit computational demand at the same time.

2.5.2 Quantile mapping on MCH, IMIS and ARPA data

The availability of CH2018 downscaled scenarios is restricted to the MCH station network. We adopt the methodology presented by Rajczak et al. (2016) and apply it on IMIS stations (Michel et al., 2021b). This Quantile Mapping (QM) methodology enables the spatial transfer of the future climate projections from the set of CH2018 MCH stations to a different network. We extend the data set from Michel et al. (2021b) to ARPA network to have consistent climate change data for all our stations.

The procedure is summarized in the following: as a first step, the so-called Most Representative Station (MRS) is selected out of the MCH CH2018 network for each ARPA station. The MRS is selected by maximising a combined correlation score between observed mean daily temperature and cumulative daily precipitation time series. The combined correlation score is computed over a time span which depends on the first available observations within the ARPA network (see Tab. 3). Climate scenarios are bias-corrected to match the long-term observational MCH CH2018 time series and, as a second step, they are spatially transferred from the selected MRS to the sparsely observed site within the ARPA network. For further details on this method, we refer readers to Rajczak et al. (2016). On the one hand, Rajczak et al. (2016) performed a robust validation on this method using long-term, multivariate measurements, revealing satisfactory performance even in topographically complex terrains. On the other hand, limitations for the application of the method remain and concern different aspects besides the well-known uncertainties related to climate projections themselves (Hawkins and Sutton, 2009, 2011). The main untailed uncertainties can be related to (1) possible non-stationarities of transfer functions under climate change (Christensen et al., 2008; Boberg and Christensen, 2012; Maraun, 2012; Bellprat et al., 2013), (2) the fact that climate change signals are imposed by the location of the MRS, (3) the fact that the spatial transfer also bears uncertainties, especially for spatially and temporally heterogeneous variables such as precipitation. Nonetheless, this method was specifically tailored to improve climate projections for sites affected by data scarcity, also addressing the fact that the closest station isn't necessarily the most representative for

the observed one. The QM spatial transfer is thus applied to the ensemble of ARPA stations, resulting in climate projections spanning to the end of the century.

Climate change simulations presented in Section 4.3 are run for the hydrological years 1991-2000 and 2081-2090. We will refer to the first decade as of "reference period", whereas to the last one as of "climate change period". For simplicity, we
5   use full decade names further in the paper (e.g. 1990-2000 for the hydrological decade 1991-2000, meaning October 1990 to September 2000).

2.5.3   Temporal downscaling

CH2018 scenarios for Switzerland only provide data at daily resolution. However, such temporal resolution can be too coarse for this data to be used as input to physically-based models such as Alpine3D and StreamFlow. This paper follows the method
10   recently proposed by Michel et al. (2021b), which provides hourly downscaled climate change time series for MCH and IMIS stations. The same approach was used to obtain hourly climate projections for ARPA stations by first disaggregating the MRS time series (according to the approach by Rajczak et al. (2016) described in the previous Section 2.5.2) and then applying the trend to ARPA stations. This approach is based on the delta-change method, which was also used in the previous CH2011 scenario (MeteoSwiss, 2011). In Michel et al. (2021b), this procedure is further developed, especially in the sense of the quality
15   assessment of the generated time series and the validation of the parameters. This enhanced method highlights that the original approach used in CH2011 did not represent correctly the seasonal cycle of the climate change scenarios. Furthermore, the method originally developed was validated only for temperature and precipitation, whereas the new method is validated on the 5 variables that are used in this paper. Time periods of 10 years are used (using the period 2005-2015 for historical baseline). The work of Michel et al. (2021a) shows that using 10 years time periods for applications such as in this study leads to similar
20   results than using 30 year periods. This finding is highly relevant for this paper for a twofold reason. On the one hand, this allows to include data from IMIS stations. Although IMIS data are not available over the last 30 years to well represent climatic cycles, the 10 years available have proven to improve discharge simulations over Alpine areas (Schlögl et al., 2016). On the other hand, without this finding, no climate change study would be possible within poorly gauged catchments whose surveying does not date 30 years back, such as Mera.

25   ## 2.6   Geographical data

2.6.1   DEM and land use

On the Swiss side, the Digital Elevation Model (DEM) used is the DTM25 dataset at 25 m resolution provided by Swisstopo, which is then resampled to the desired resolution of 500 m. Land cover data are derived from the 2006 version of the Copernicus Corine Land Cover (Agency, 2013) dataset at 100 m resolution, then upscaled to 500 m. Corine Land Cover classes are
30   translated into the land cover classes available in A3D, listed in Michel et al. (2021b). On the Italian side, the DEM is provided by NASA (https://earthdata.nasa.gov/) and land cover data by CLC. The initial resolution of the DEM is 30 m, then upscaled to 500 m.

For the flow routing model StreamFlow (see Section 3.1.2), watershed, sub-watershed and stream network are derived using the TauDEM software (Tarboton, 1997) coupled with a wrapper (Michel et al., 2021a) to force it to reproduce exactly the stream network provided by the FOEN (of the Environment, 2013).

Conversely, Poli-Hydro model simulates the hydrological balance and the flow routing for the catchment area (through flow

5 directions and flow accumulations) which is identified from the DEM using ArcGIS software (ESRI, 2012). The DEM used is previously upscaled to match the scale of interest (500 m).

It is important to underline here that the two approaches prepare the DEM in different ways, thus justifying slight differences in basin shape hereafter.

2.6.2 Glacier data

10 For both catchments, we use glacier height maps from Zekollari et al. (2019). By means of an innovative model called Glo-GEMflow, the authors have developed the first regional Alpine glacier modelling making use of high-resolution climate model outputs from EURO-CORDEX (Jacob et al., 2014; Kotlarski et al., 2014). Glacier height needs to be understood as the ice depth above the surface. Such maps are used as initial condition of each past and future (climate change) simulation. For historical periods, all calibrations are run with the 2005 glacier map as initial condition. Glacier evolution is then treated and

15 simulated according to the two models and melt schemes. A due consideration is that glacier maps do overwrite the CLC land cover classes, and in the case that a pixel is classified as glacier by CLC but not by Zekollari et al. (2019), that pixel is simply converted into bare rock.

3 Methods

3.1 Models description

20 In this study, two different models are compared: the degree-day model Poli-Hydro (PH hereafter) and the process-based model chain Alpine3D+StreamFlow (A3D+SF hereafter). Interestingly, both models have been used recently to perform climate change studies. In Michel et al. (2021a), A3D+SF chain is used to investigate future water temperature of Swiss rivers under climate change. Fuso et al. (2021) used PH to evaluate future hydrological scenarios for the Lake Como catchment in Italy according to three GCMs of the Sixth Assessment Report (AR6) of the IPCC driven by four shared-socio-economic-pathways-

25 based scenarios. PH solves within the same model the surface mass and energy balance in order to obtain soil runoff and the hydrological routing. For the process based side, A3D is first run in order to obtain the soil runoff, and in a second step SF is run for the hydrological routing using A3D output as input.

The forcing meteorological data at station location are extrapolated to the grids in A3D using distance weighted method with vertical lapse rate corrections (see details in Michel et al. (2021a)). The grids of forcing data are written out and also used in

30 PH, in order to use exactly the same forcing data. Longwave radiation is also computed in A3D using TA, RH and cloud cover

derived from ISWR using the approach described in Omstedt (1990). The grids are produced at 500 m resolution. The same forcing grids are then used in PH.

3.1.1 Alpine3D

Alpine3D model (Lehning et al., 2006) is a deterministic and spatially distributed model designed for high-resolution simula-
5 tion of snow processes in topographically complex areas. In A3D, the SNOWPACK model (Lehning et al., 2002) is applied to each cell of the catchment (here a 500x500 m grid). SNOWPACK is a physical 1-D multi-layer snow and soil model for-merly developed for avalanche warning. It comes with a detailed description of the snow micro-structure and it resolves phase changes in snow on a physical basis. SNOWPACK also contains a two-layer canopy module simulating the micro-meteorology in the forest, the evapotranspiration, and the interaction between trees and snow, including snow interception (Gouttevin et al.,
10 2015).

A3D can also be run with a simpler melt-factor energy balance mode (Shakoor et al., 2018), this setup is called A3D$_{DD}$ hereafter. In A3D$_{DD}$, the melt rate is linearly linked to air temperature by the Temperature Melt Factor (TMF). In A3D, A3D$_{DD}$ can be used within SNOWPACK as an energy boundary condition at the snow-atmosphere interface. When a melt phase occurs, i.e. when water and ice are coexisting in the snow element and air temperature is larger than snow surface temperature, the
15 energy entering the snowpack is computed with A3D$_{DD}$. Radiative fluxes are set to zero in the uppermost snow element and the turbulent fluxes are not computed anymore. When the snowpack is not in a melting phase, the net energy flux at the snow surface is computed by solving the standard energy-balance boundary condition (Schlögl et al., 2016). Ice and snow are not distinguished by A3D$_{DD}$ melt model: a single TMF value is used for both of them, although such approach has been shown to be oversimplified for glacierized catchments (Hock, 1999). The energy balance (EB) in case of melting is computed by A3D$_{DD}$
20 as of Equation 1.

$$EB = TMF(T_{air} - T_{snow}) \tag{1}$$

In Equation 1, TMF is the Temperature Melt Factor; T_{air} and T_{snow} the air and snow temperature, respectively.

Within A3D, the SNOWPACK model is run for every cell of the catchment at a 15 minutes time step, while the forcing boundary conditions are updated on an hourly basis. A3D also has a radiative sub-module enabling the consideration of topo-
25 graphic shading, terrain reflection and atmospheric cloudiness on shortwave radiation distribution. Topographic effects deeply influence radiation balance in mountain regions. In fact, the intensity of incoming and outgoing radiative fluxes depends on many factors such as surface inclination angle, shading and surface properties (von Rütte et al., 2021). In mountain terrain, in-coming longwave radiation decreases with elevation because higher areas have an enhanced angular exposition to the radiating sky, which is colder than the terrain (Lehning et al., 2006). A3D is run with the same setup described in detail in Michel et al.
30 (2021a), to which we refer for further details.

3.1.2 StreamFlow

StreamFlow is a semi-distributed hydrological model, described in (Gallice et al., 2016). The runoff output produced in A3D is collected into each sub-watershed in SF and the residence time in the soil is determined with an approach using two linear reservoirs. The water is then routed to the river. While SF offers multiple routing scheme, the simple instant routing scheme

5   shows a similar performance as more complicated ones for Alpine catchments studies (Michel et al., 2021a), thus we use this scheme in the present work. SF is run at hourly timestep and at a resolution of 100 m. Further details about the setup used for SF can be found in (Michel et al., 2021a).

3.1.3 Poli-Hydro

Poli-Hydro is a spatially semi-distributed, cell based (here 500x500 m) hydrological model, suitable for the simulation of high-

10   altitude, poorly gauged catchments (Bocchiola et al., 2018; Casale et al., 2020). It is able to reproduce deposition of snow and ablation of snow and ice through the accumulation of thermal time, i.e. the daily sum of degree-days. The model works at a daily time scale and it is based on the mass conservation equation. The mass balance involves liquid and solid precipitation, snow melt, glacial ablation, effective evapotranspiration and groundwater discharge. The model takes as input data a DEM, a land cover map, and daily values for air temperature and total precipitation. As mentioned before, input data are the meteorological

15   forcing grids extrapolated by A3D. This model conceives the formation of flow by means of two mechanisms: superficial and groundwater discharge.

    Snow melt and glacial ablation are estimated by means of a melt factor (Equation 2). Here, the model takes into consideration contributions from shortwave radiation and temperature (Aili et al., 2019; Bombelli et al., 2019; Stucchi et al., 2019).

$$M_s = DD_s(T - T_t) + RMF_s(1 - \alpha_s)q_{sw}$$
$$M_i = DD_i(T - T_t) + RMF_i(1 - \alpha_i)q_{sw}$$

(2)

20   In Equation 2, M_s and M_i are the melt factors for snow and ice respectively; T the mean daily temperature; DD_s and DD_i the temperature melt factors for snow and ice (Tab. 7); T_t the threshold temperature equal to 0°C, RMF_s and RMF_i the radiation melt factors; $\alpha_{s,i}$ the snow and ice albedo (Soncini et al. (2017)); q_{sw} the shortwave radiation flux.

    Ice and snow melt occur only when the average daily temperature is higher than the threshold temperature. Ice melt occurs on ice covered domain cells, once snow melt is completed.

25   Overland flow Q_s is only generated in the condition of saturated soil (Equation 3).

$$Q_s = S_{t+\Delta t} - S_{max} \quad if \ \ S_{t+\Delta t} > S_{max}$$
$$Q_s = 0 \quad\quad\quad\quad\ if \ \ S_{t+\Delta t} \leq S_{max}$$

(3)

In Equation 3, S is the soil water content evaluated from the mass balance equation for each time step, and S_{max} is the greatest potential soil storage calculated by the Curve Number method (Usda., Scs., 1986), from a land cover map for each cell of catchment. The sub-superficial discharge Q_g is computed as of Equation 4

$$Q_g = K \left(\frac{S}{S_{max}} \right)^k \tag{4}$$

5   In Equation 4, K is the saturated permeability and k is a power exponent.

Then, for the flow routing, two parallel systems are considered, one for superficial flow and one for sub-surface flow. Two instantaneous unit hydrograms $u(t)$ (IUH) are evaluated for each cell using the Nash approach (Rosso, 1984) for superficial and sub-surface flow respectively.

3.2 Statistical scores

10   Statistical scores are used to assess the models' predictive performance. In this work, three metrics are used: the RMSE for the snow cover analysis in Section 4.2.1, the Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe (1970)) and the Kling-Gupta Efficiency (KGE, Gupta et al. (2009)) for the comparison of predicted discharge in section 4.2.

With respect to NSE, the value $NSE = 0$ is commonly regarded as an inherent benchmark for models' performance (Schaefli and Gupta, 2007; Knoben et al., 2019), as it describes the situation where the average of the observations has the same

15 explanatory power as the model's predictions.

The KGE score further extends NSE into three different components: the correlation r, the flow variability error α and the mean flow bias β. It is important to remark that, as explained in Knoben et al. (2019), it is impossible to define within KGE a benchmark value that acts as a threshold between a "good" and a "bad" model. To keep consistency with NSE, here we will consider as a benchmark the case where $Q_{sim} = \overline{Q_{obs}}$, which yields $KGE = -0.41$ (Knoben et al., 2019) and corresponds to

20 $NSE = 0$.

The need for considering the KGE in addition to the widely-used NSE mostly arises from two important drawbacks of the NSE. In the first place, The use of the observed mean annual as a benchmark can be a mediocre predictor (Schaefli and Gupta, 2007), especially for strongly seasonal discharge time series, as those dominated by snow melt, leading to the overestimation of the model efficiency. Secondly, the NSE computes squared differences between the observed and predicted values. As a

25 result, the metric becomes excessively sensitive to extreme values by enhancing higher magnitude streamflows and neglecting lower ones (Legates and McCabe, 1999; Criss and Winston, 2008). It is important to note that PH is calibrated using the NSE metric as the score to maximise, while SF uses the KGE metric, so each model might be biased toward better performance in its respective calibration metric.

Table 5. StreamFlow calibration parameters.

		KGE	Maximum infiltration rate (mm d $_{-1}$)	Upper reservoir τ (d)	Lower reservoir τ (d)
A3D	**Mera**	0.74	3.829e+01	8.441e+00	4.398e+02
	Dischma	0.91	9.598e+00	2.408e+01	2.926e+02
A3D$_{dd}$	**Mera**	0.77	6.319e+01	7.459e+00	1.231e+02
	Dischma	0.88	3.764e+00	2.181e+01	3.922e+02

4 Results and discussion

4.1 Calibration

4.1.1 StreamFlow

To calibrate the model SF, 10'000 calibrations runs are performed over the calibration period October 2010 to September
2014 (the validation period being from October 2014 to September 2018). The following parameters are calibrated: maximum
infiltration rate, upper reservoir residence time (τ), lower reservoir τ, and fraction of water lost to deep soil infiltration. The
calibration is achieved with a Monte Carlo approach using 10'000 random parameter sets and the corresponding model runs
over the historical period. The random sets are drawn from uniform distributions, with bounds corresponding to standard values
available in the literature (see Gallice et al. (2016)). The performance is assessed using the Kling-Gupta Efficiency (KGE). SF
is calibrated separately for the run with A3D in full energy balance mode and for the run with A3D$_{DD}$.

For the Mera catchment, a KGE value of 0.74 is obtained for the calibration period (0.77 with A3D$_{DD}$). For the Dischma
catchment, a value of KGE=0.91 results (0.88 with A3D$_{DD}$). The values of the parameters obtained are given in Tab. 5.

4.1.2 Poli-Hydro model

Before the overall model calibration, degree-days for snow melting are calibrated using a two-step method. First, the simu-
lated Snow Cover Area (SCA) is compared against satellite MODIS images (Snow Cover Extent, Copernicus Global Land
Service, available online: https://land.copernicus.eu/global/products/sce). Such validation is performed once every 15 days,
i.e. the overpassing frequency of MODIS satellite on the area of Valchiavenna. This method gives information about the SCA
extension, although it is not possible to verify snow depth. To cover this, snow depth data measured at AWS are collected. For
Mera catchment, snow depth data for the entire simulation period is only available at the station San Giacomo Filippo (SGF,
Lat 46.361°, Lon 9.319°, 2064 m a.s.l.). Monthly values of snow degree-days are reported in Tab. 6.

The simulated snow water equivalent is compared with measured snow depth at the AWS. Within Dischma catchment itself,
MCH/IMIS snow depth data are not available, therefore only SCA is verified. Detailed approach and results are reported for
Mera catchment in Fuso et al. (2021). Modelled snow validation is performed comparing observed and modelled time series
of snow depth. To this aim, the daily modelled snow water equivalent is converted to snow depth using the approach proposed

Table 6. Snow degree-days used by Poli-Hydro model for Mera and Dischma catchments.

DD_s mm $C^{-1}d^{-1}$	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mera	1.14	1.07	2.00	2.00	1.80	1.60	0.40	0.40	0.50	0.50	1.80	1.14
Dischma	1.14	1.07	2.00	3.50	9.00	9.00	0.5	0.5	1.00	0.80	1.00	1.14

by Martinec and Rango (1989) and more recently implemented in Aili et al. (2019). It is important to remark that the approach by Martinec and Rango (1989) is only used to verify the modelled snow water equivalent against the measured snow depth. PH model does not model physical processes and energy fluxes within the snowpack layers. For additional details on this procedure, we refer to the aforementioned papers.

5   Poli-Hydro model is then fed with a set of fixed calibration parameters from other previous studies on the same areas of central Alps and Valtellina (Soncini et al., 2017; Aili et al., 2019; Fuso et al., 2021). The following parameters are calibrated: soil permeability, thermal factors, saturated permeability K, power exponent k and lag times (superficial t_s and sub-surface t_g). The initial values of K and k are taken based on literature research on the study area (Aili et al., 2019), and then (manually) iteratively tuned by fitting modeled against observed discharge. Lag times are assessed according to power law functions of the

10   basins' contributing area, both for superficial and sub-superficial flow (Casale et al., 2020). Model parameters are listed with reference in Tab. 7 for Mera and Dischma catchment. Calibration is performed from October 2010 to September 2014 whereas validation from October 2015 to September 2018.

For model calibration and validation, simulated discharge at each catchment's closure section is compared with the observed value. The Percentage Bias (PBIAS) and the Nash-Sutcliffe Efficiency (NSE) are computed during the calibration: the absolute

15   value of the PB is minimized and the NSE is maximized. Results are reported in Tab. 8.

4.2   Model comparison

As mentioned in the Introduction, the use of degree-day models might be preferred for the lighter computational load and the lowest amount of input data required, but on the other hand, their use for climate change impact studies is disputable.

With the aim of performing climate change impact studies, the objective is to assess the models' efficiency in reproducing

20   discharge and total volumes at the river section of interest. However, high-performance discharge simulations strongly depend on the correct representation of runoff formation. Specifically, runoff in Alpine catchments can be separated in two main components: the fast runoff, which is mainly due to precipitation and surface flow, and the slow runoff, generated by snow and ice melt and sub-surface flow.

Thus, in this section we compare the models presented in Sections 3.1.1, 3.1.2 and 3.1.3 in three different ways: (1) by

25   means of a snow cover analysis, (2) before flow routing (referring to the water volume released at each pixel of the catchment domain) considering the different contributions to the total runoff (precipitation, snow melt and glacier melt) and (3) after flow routing (referring to discharge and volumes). When discussing runoff contributions, before the flow routing, we refer to models

Table 7. Calibration parameters for Poli-Hydro model.

Description	Parameter	Unit	Mera	Dischma	Reference
Threshold melt temperature	T_t	°C	0	0	Soncini et al. (2017), Aili et al. (2019), Bombelli et al. (2019), Fuso et al. (2021)
Wilting point	Θ_W	-	0.15	0.15	Soncini et al. (2017), Aili et al. (2019), Bombelli et al. (2019), Fuso et al. (2021)
Field capacity	Θ_L	-	0.35	0.35	Soncini et al. (2017), Aili et al. (2019), Bombelli et al. (2019), Fuso et al. (2021)
Saturation	Θ_S	-	0.45	0.45	Stucchi et al. (2019),
Glacier degree-day	DD_i	$mm\,°C^{-1}d^{-1}$	5	5	Bombelli et al. (2019), Fuso et al. (2021)
Radiation melt factor (snow)	RMF_s	$mm\,d^{-1}W^{-1}m^2$	$2.10e^{-3}$	$2.10e^{-3}$	Soncini et al. (2017)
Radiation melt factor (glacier)	RMF_i	$mm\,d^{-1}W^{-1}m^2$	$1.5e^{-3}$	$1.5e^{-3}$	Soncini et al. (2017), Aili et al. (2019)
Power Q_g	k	-	1.2	1.1	Fuso et al. (2021), Calibration
Soil permeability	K	$mm\,d^{-1}$	4	3.5	Aili et al. (2019), Calibration
Sub-superficial lag time	t_g	h	270	85	Calibration
Superficial lag time	t_s	h	230	60	Calibration
Free debris cover glacier albedo	α_i	-	0.3	0.3	Soncini et al. (2017), Aili et al. (2019), Bombelli et al. (2019)
Snow albedo	α_s	-	0.7	0.7	Soncini et al. (2017), Aili et al. (2019), Bombelli et al. (2019)

Table 8. Poli-Hydro calibration scores.

	Mera	**Dischma**
NSE	0.5	0.36
PBIAS	-0.5%	-0.82%
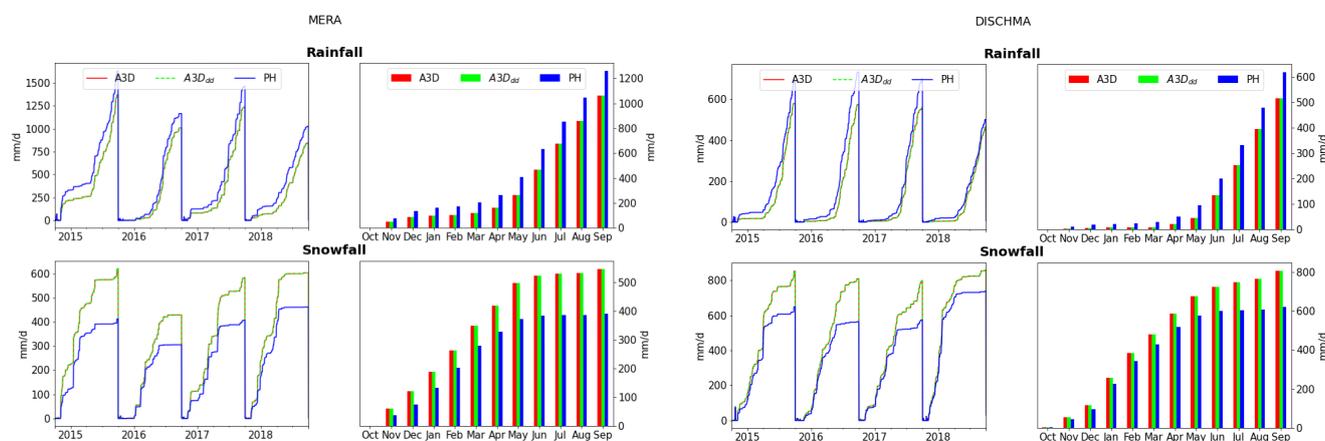
Figure 3. Comparison of solid and liquid precipitation in A3D/A3D$_{DD}$ and PH for Mera (left) and Dischma catchment (right).

as A3D, A3D$_{DD}$ and PH respectively for Alpine3D (in its full solver and degree-day version) and Poli-Hydro (Sections 3.1.1, and 3.1.3). Conversely, for discharge and volumes, after the flow routing, we refer to the same models as SF, SF$_{DD}$ and PH$_R$ respectively for Streamflow (forced with A3D and A3D$_{DD)}$ and the flow routing scheme of PH (Sections 3.1.2 and flow routing scheme in Section 3.1.3).

5   **4.2.1   Snow cover analysis**

In the first place, the two models are implemented with different temperature thresholds for rain-snow separation, 0°C for PH (Fuso et al., 2021) and 1.5°C for A3D and A3D$_{DD}$ (Michel et al., 2021a), 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).

    Figure 4 shows observed and modelled snow height and snow water equivalent within the Mera domain in the cell where the
10 high-elevation station of San Giacomo Filippo is located. Snow water equivalent (i.e. snow density) measures are not available for this station. A similar analysis was not possible for Dischma, because the MCH/IMIS dataset used for this paper did not contain snow height or snow water equivalent measurements within the basin itself. Models' performance in reproducing snow height is rated by means of RMSE. Snow height is best reproduced by A3D, which gives the lowest RMSE (top right of Fig. 4 panel (a)). A3D$_{DD}$ tends to underestimate snow height, however, RMSE is only slightly higher than A3D. PH shows the
15 highest RMSE. Common errors for A3D seem to be related to compaction, for A3D$_{DD}$ to an excessively fast melting, whereas PH shows errors in accumulation and melting. The poorer performance of PH may be explained by the relative simplicity of the
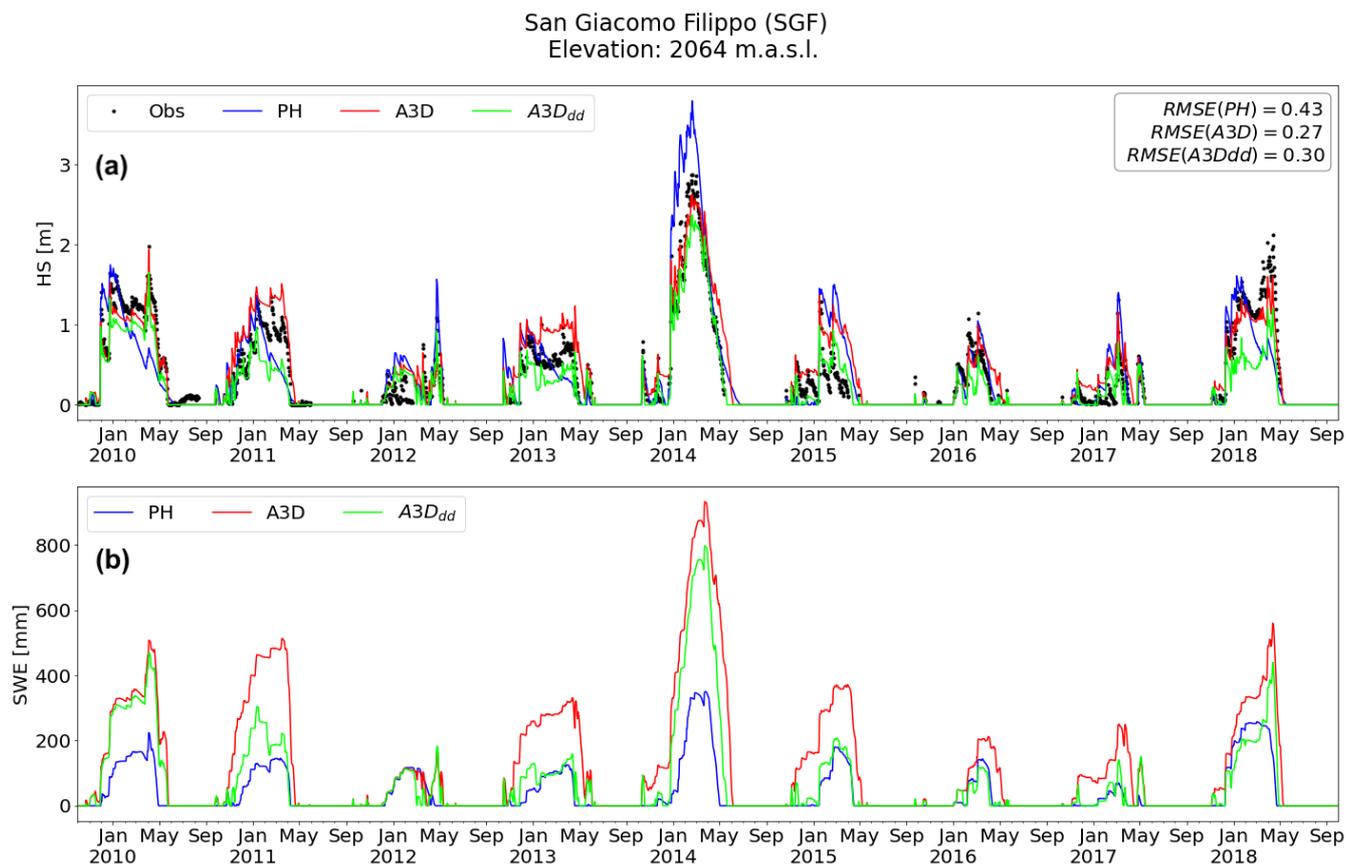
Figure 4. Snow cover analysis for high-altitude station of San Giacomo Filippo within Mera catchment. Panel (a) shows observed and modelled HS, panel (b) shows modelled SWE.

assessment of snow depth from snow water equivalent by means of the approach from Martinec (1991) mentioned in Section 4.1.2. PH systematically simulates lower snow water equivalent than A3D. Exceptions are the years 2010 and 2014, for which snow water equivalent predicted by A3D$_{DD}$ is more in agreement with PH rather than A3D. Despite measured snow water equivalent not being available here, literature generally agrees on A3D often outperforming simpler models in reproducing it,

5   for example in Terzago et al. (2020) among others. With reference to Section 4.1.2, it is important to remark that the station of San Giacomo Filippo is used to calibrate/validate snow water equivalent in model PH but not in A3D/A3D$_{DD}$. On the one hand the comparison is uneven, but on the other hand it conceives even more relevance to A3D and A3D$_{DD}$ outperforming PH there.

Modelled snow water equivalent at basin scale is shown in Fig. 5 and Fig. 6. As already observed in Fig. 4, A3D reproduces systematically higher snow water equivalent compared to A3D$_{DD}$ and PH. Snow water equivalent peaks are equally reproduced

10   by all models in terms of seasonality. However, regardless of the catchment and year, PH pictures a longer melting season. The reason might be a protracted snow melt process due to PH's degree-day scheme and the resolution of meteorological forcing: as mentioned in Section 3.1.3, the melting scheme only takes into account a mean daily temperature value, which needs to be

higher than the melting threshold for any snow melt to be represented. This condition is likely to be met only after melting has already set in reality, delaying the process of snow ablation and leading to more snow mass predicted for summer months. These findings are confirmed by the work of Terzago et al. (2020), where melt models of lower complexity did show higher biases against observations with respect to more complex ones like SNOWPACK. In Terzago et al. (2020), high biases for

5   simpler models are related both to their underestimation of snow water equivalent peak values and to the protraction of snow melt season that they reproduce. Maps of summer months in Fig. 5 and Fig. 6 corroborate that PH reproduces higher snow water equivalent at high elevation areas, especially in early summer.

4.2.2   Runoff formation

The different components of runoff are shown in Fig. 7 and Fig. 8. Throughout Section 4.2.1, we underlined major differences

10   among models in how they reproduce the snow cover. Such differences lead to discrepancies in modeled snow melt. As a consequence of the model's structure described in Section 3.1.3 and 4.2.1, PH reproduces a delayed and reduced melt season compared to both versions of A3D. Interestingly, $A3D_{DD}$ reproduces a melt season closer to the one obtained with the full solver and more important melt events during the winter.

An altered reproduction of snow melt seasonality affects models' performance in reproducing discharge. In Fig. 7 and Fig.

15   8, observed and modeled discharge after flow routing and monthly-averaged volumes are shown. For the Mera catchment, the general behaviour of the three models is quite similar, being characterized by an underestimation of the observed discharge over the accumulation season and overestimation over the melting season. Besides a similar general pattern, SF reproduces the observed flow with more accuracy during the melting season (regardless of being forced with A3D in its full-solver or degree-day version), whereas PH_R shows a constant bias towards higher discharges. In addition, PH_R does not distinguish snow

20   melt or glacier melt dominated seasons, as contributions are smoothed and do not show a predominant presence throughout the melting season. The effect of this is a poorer representation of the observed discharge, where peaks are smoothed and discharge is generally overestimated. For the nivo-glacial regime of the Dischma catchment, the difference between the models is even more evident. The different timings of snow melt have an important impact on predicted discharge: PH_R delays the spring snow-melt-induced discharge by one month compared to observations, and additionally it fails to represent the accentuated discharge

25   peak forced by snow melt characterizing nival rivers (Déry et al., 2009). The consequence is a poor reproduction of discharge regime and volumes throughout the entire melting season, with severe underestimations in early spring and, correspondingly, overestimation in summer and autumn. Conversely, as a consequence of the correct representation of the melt dynamics, both versions of A3D almost equally reproduce discharge and volumes accurately.

4.2.3   Statistical performance of runoff simulations

30   In this section, performance metrics described in 3.2 are used to evaluate discharge simulations. Despite showing that simpler melt schemes conceive an altered representation of runoff seasonality, the previous section did not quantify such alterations.

Table 9 shows models' performance scores over the validation period. In general, better performances are obtained over Dischma catchment than over Mera, in agreement with what can be seen in Fig. 7 and Fig. 8. The Mera river being regulated
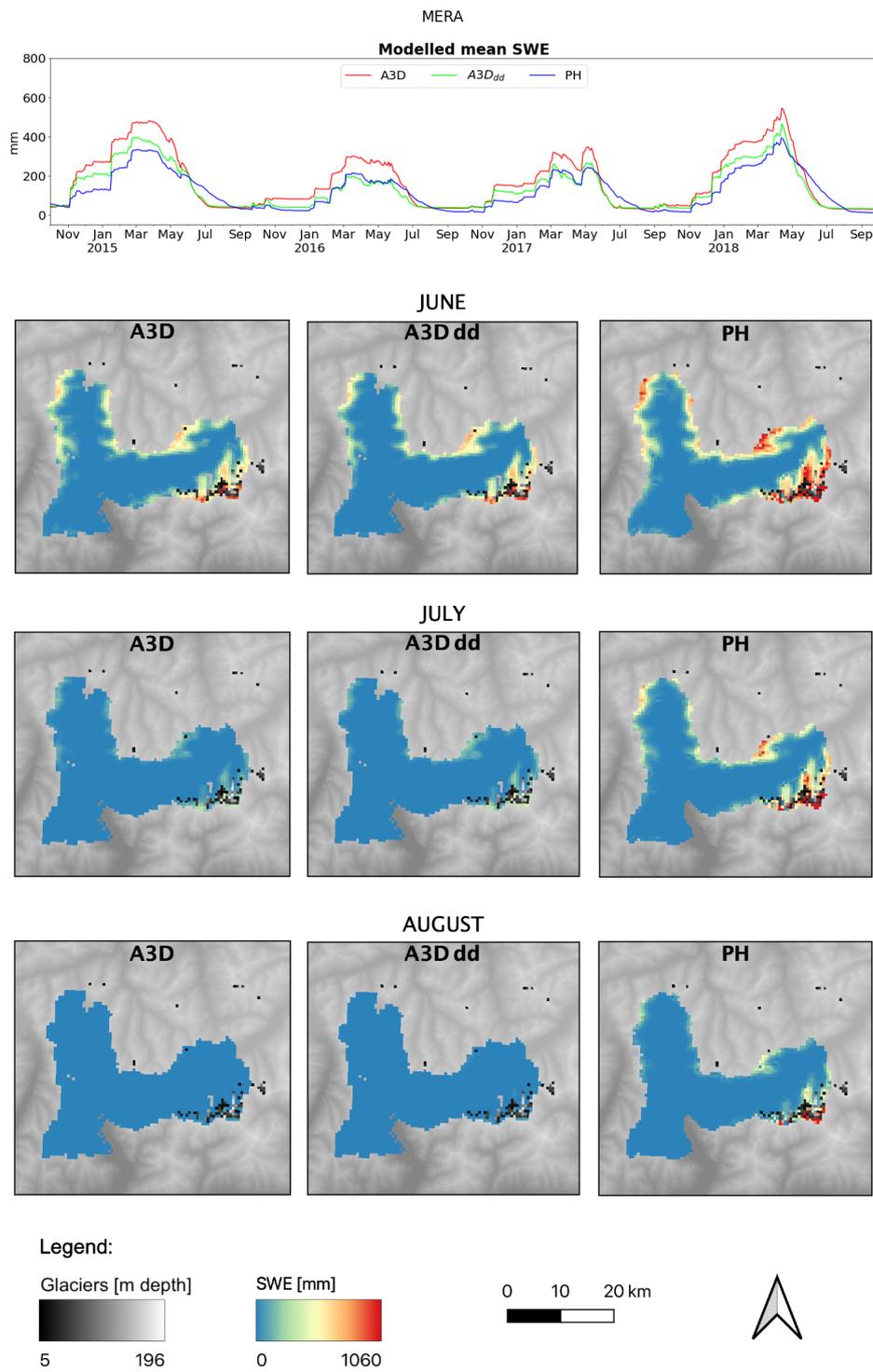
Figure 5. Top: modelled mean SWE over Mera basin for validation period. Bottom: SWE maps for summer months over Mera basin for calibration and validation periods.
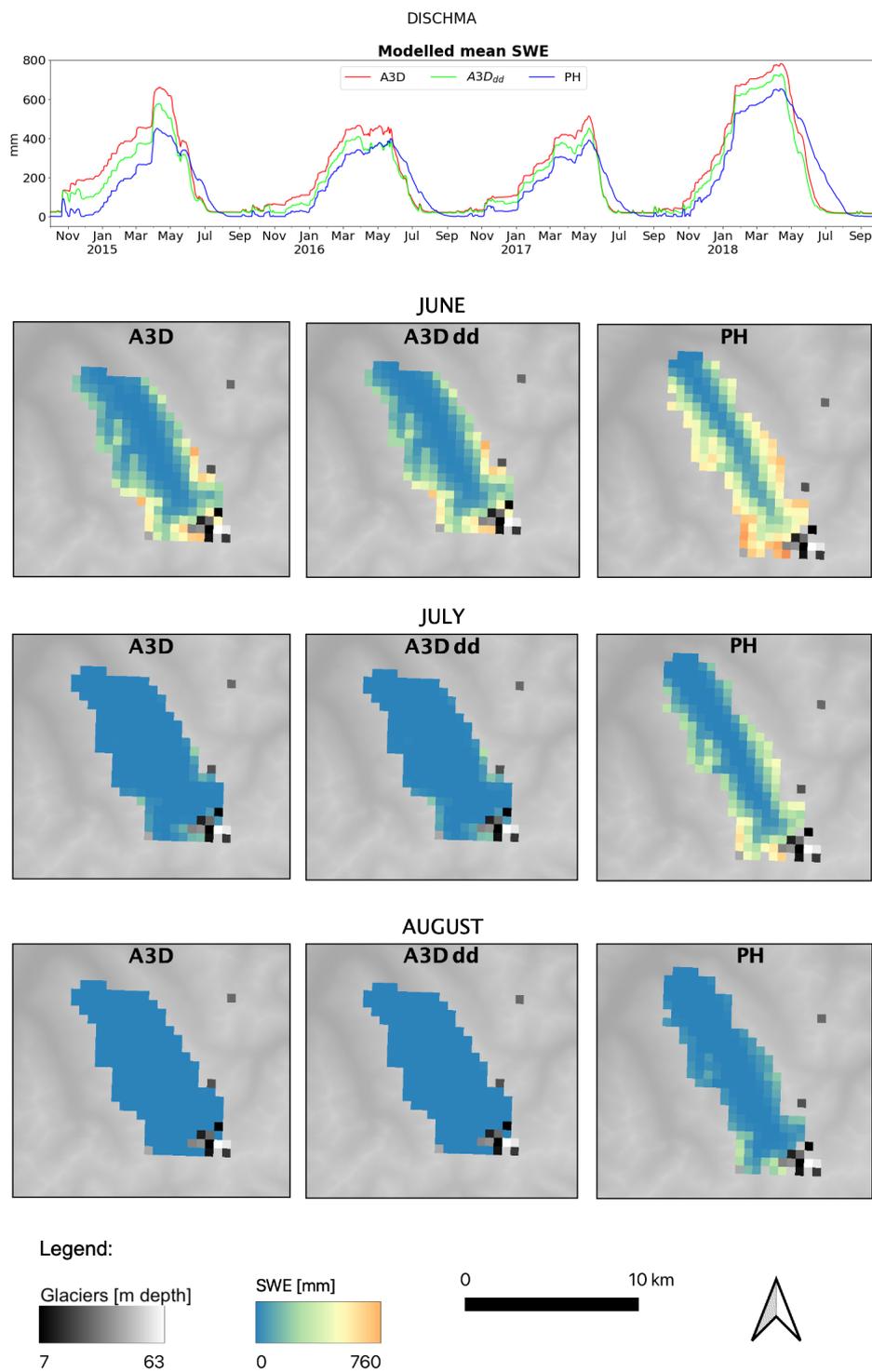
Figure 6. Top: modelled mean SWE over Dischma basin for validation period. Bottom: SWE maps for summer months over Mera basin for calibration and validation periods.
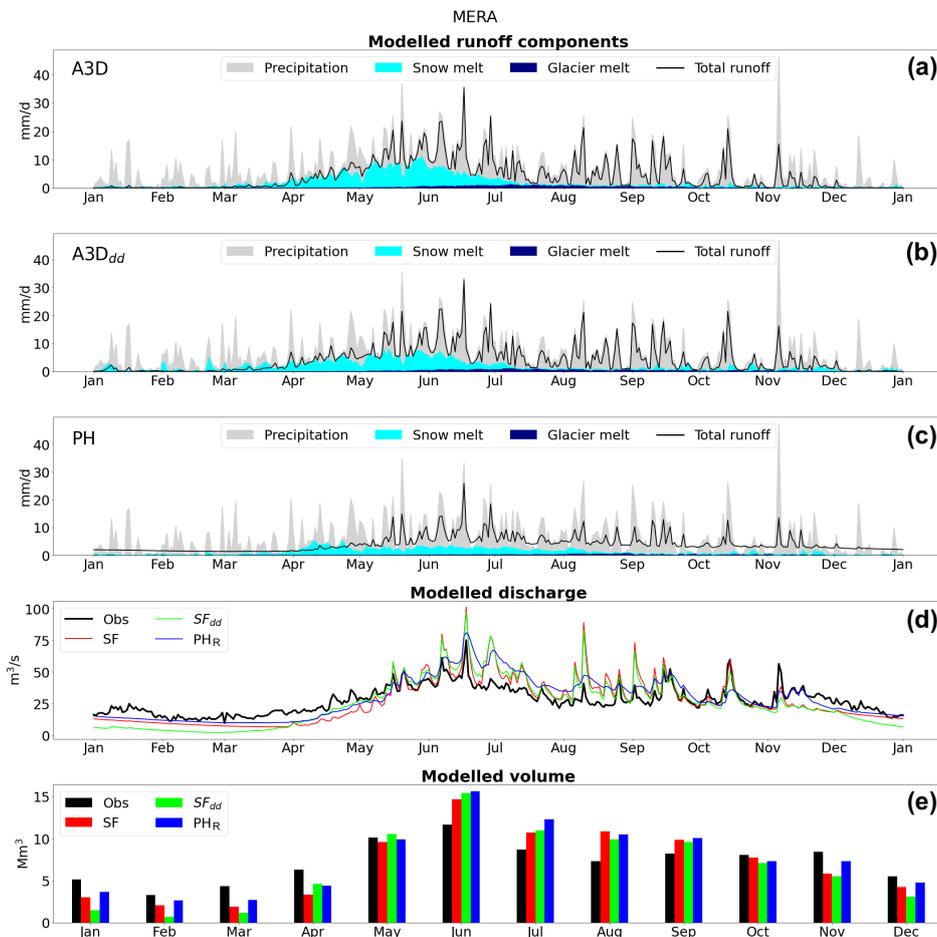
Figure 7. Panels (a), (b), (c): Modelled runoff components over the Mera catchment reproduced with the full solver of Alpine3D (A3D), the degree-day melting scheme of Alpine3D (A3D$_{DD}$) and Poli-Hydro (PH), respectively. Panels (d) and (e): Respectively discharge and monthly volumes modelled by StreamFlow using output from the full solver of Alpine3D (SF) or its degree-day scheme (SF$_{DD}$) and modelled by Poli-Hydro (PH$_R$) compared to measurements. Results for validation years 2015-2018, averaged by day of the year and by month.

Table 9. Performance scores over validation period for discharge in Mera and Dischma catchments, for StreamFlow forced by the full solver output of Alpine3D (SF), the degree-day melting scheme of Alpine3D (SF$_{dd}$) and Poli-Hydro (PH$_R$)

	Mera					**Dischma**				
	NSE	**KGE**	**r**	α	β	**NSE**	**KGE**	**r**	α	β
SF	0.14	0.60	0.67	1.22	0.96	0.89	0.89	0.94	0.91	0.96
SF$_{dd}$	0.20	0.62	0.69	1.21	0.92	0.84	0.85	0.92	0.90	0.92
PH$_R$	0.37	0.69	0.75	1.17	1.05	0.61	0.78	0.79	0.93	1.10
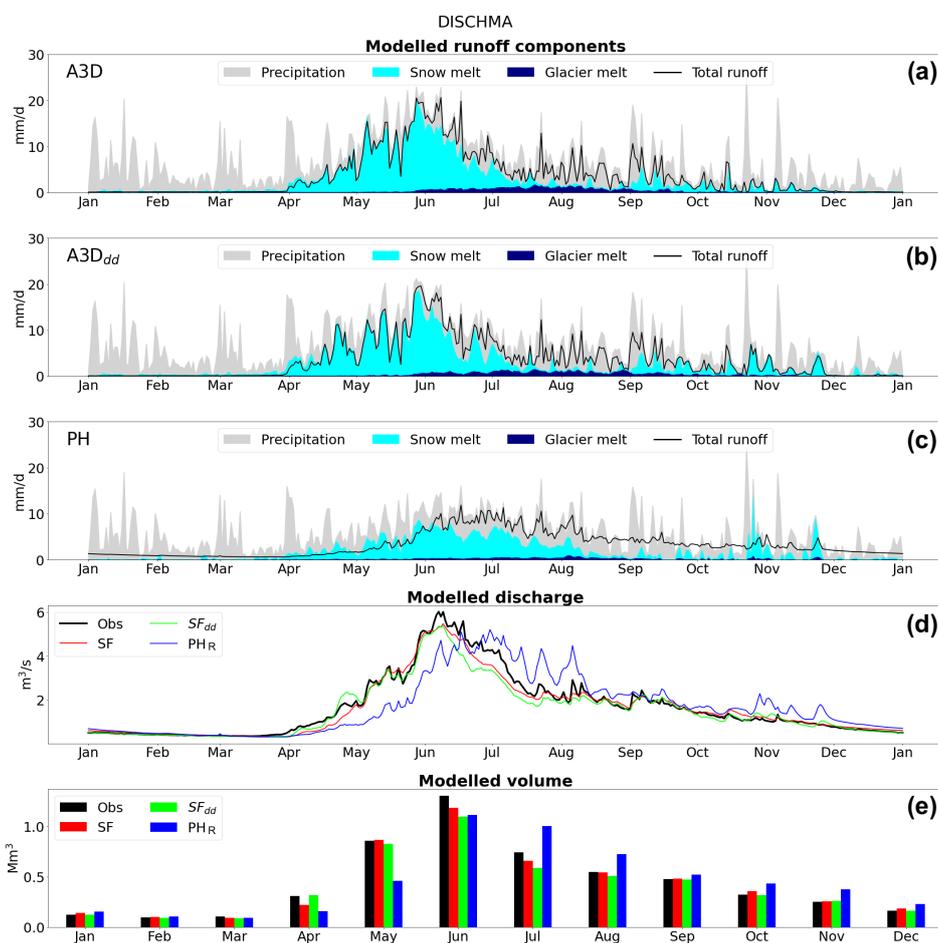
Figure 8. Panels (a), (b), (c): Modelled runoff components over the Dischma catchment reproduced with the full solver of Alpine3D (A3D), the degree-day melting scheme of Alpine3D (A3D$_{DD}$) and Poli-Hydro (PH), respectively. Panels (d) and (e): Respectively discharge and monthly volumes modelled by StreamFlow using output from the full solver of Alpine3D (SF) or its degree-day scheme (SF$_{DD}$) and modelled by Poli-Hydro (PH$_R$) compared to measurements. Results for validation years 2015-2018, averaged by day of the year and by month.

by multiple dams (see Fig. 1 and Tab. 1), this result is not surprising and illustrates the need of explicitly considering such infrastructures in models (which is not the case for the two models used here). With PH_R, better scores are obtained both in terms of annual NSE and KGE over the Mera, although we previously showed that the predicted melt dynamics is substantially wrong and Fig. 7 shows a clear overestimation of runoff during the summer season. Table 9 also shows the r, α and β

5   components of KGE scores. Over the Mera, flow variability predicted by the models is higher than the observed one (positive α), suggesting again the influence of the regulation.

Over the Dischma catchment, both indicators show a lower performance of PH_R compared to SF, as we can expect since PH_R tends to delay the melt season by one month. As mentioned earlier, for catchments with strong seasonal signals using quality metrics on a yearly basis is not optimal. Many authors have addressed this issue by using those metrics on a seasonal basis

10   (Garrick et al., 1978; Martinec and Rango, 1989; Legates and McCabe, 1999; Schaefli and Gupta, 2007). We thus evaluate the models' performances by dividing the validation interval according to seasons. With this approach, the observed discharge is no longer averaged over the ensemble of validation years, but rather over the ensemble of validation seasons. As a consequence, the mean flow as a benchmark case gains specific significance as a function of the considered season. The seasonal performance analysis (Fig. 9) shows that again lesser scores are obtained for Mera catchment. Overall, both versions of the SF+A3D chain

15   outperform $PH+PH_R$ in almost all seasons and in both catchments. PH_R obtains remarkably low scores in winter. The delayed melt season of PH for the Dischma catchment is also well visible in the scores obtained in spring season by PH_R. Interestingly, both versions of SF-A3D obtain similar scores.

4.2.4 Discussion

The annual scores shown in Tab. 9 suggest that the simpler PH_R model outperforms the A3D+SF model chain in the Mera

20   catchment, while obtaining decent results over the Dischma catchment. However, from Fig. 7 and Fig. 8, the superiority of PH_R model is not evident in the Mera, whereas for the Dischma the error on the melt season timing cannot be considered as a satisfactory result, contrarily to what the annual NSE (0.61) and KGE (0.78) values may suggest. The seasonal analysis of the metrics shows a completely different picture with a clear superiority of the A3D+SF model chain, especially over the Dischma catchment. This highlights the limitation of using such metrics on an annual basis and the benefits of narrowing them on a

25   seasonal perspective.

In general, both versions of A3D used reproduce the melt dynamics more correctly than PH, the main issues with PH being the shifted seasonality of the melting and the absence of a peak behaviour. Indeed, PH tends to reproduce a rather constant melt. This result is similar to the findings by Magnusson et al. (2010). The degree-day version of A3D obtains results surprisingly close to the full energy balance solver. In the Mera catchment, snow melt events are more important in winter with $A3D_{DD}$,

30   leading to a reduced snow height and thus snow melt later in the year. In the Dischma catchment, the melt is slightly enhanced in spring and thus reduced in summer in $A3D_{DD}$. Besides, the in-depth snow cover analysis in Section 4.2.1 shows the superiority of both versions of A3D against PH, which however does not model snow depth directly. However, the difference in discharge between the two versions of A3D are rather low and the seasonal scores obtained are similar. This shows that part of the
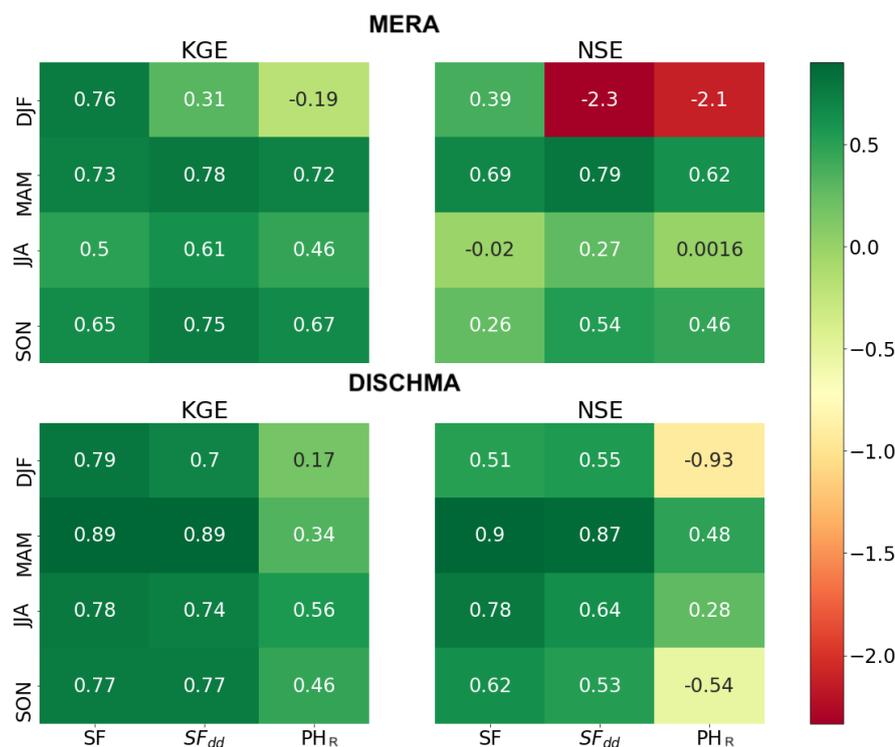
Figure 9. Seasonal KGE (left) and NSE (right) scores obtained over the Mera (top) and Dischma catchment (bottom). Scores are shown for the two versions of A3D and for PH.

difference between outputs of A3D can be absorbed in the calibration of the soil reservoir residence time in SF, leading in the end to similar discharge simulations.

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
5   means that if the mean is lower than the melting threshold, the model does not simulate any snow melt, 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. Previous studies showed that adopting shorter time steps in hydrological modelling can be beneficial for runoff simulations, not only for short-duration events, but also for the analysis of outputs at a larger time scale (Ficchi et al., 2016). This however requires to have forcing time series at hourly resolution, which is not always the case,
10   especially for climate change scenarios. In addition to the shorter time step, the degree-day scheme of A3D benefits from the enhanced complexity of the model. In particular, this scheme is enabled only when the melting starts (i.e. liquid and solid water coexist in the snowpack). The onset of melting is thus still determined by the full energy balance solver. Moreover, A3D$_{DD}$ remains a multi-layer snow model, which is generally desired for representing snow processes in a physically based model, given the significant vertical variations of snow properties Arduini et al. (2019). This version of A3D is thus not at all similar

to a simple degree-day model such as PH. It is important to note that, unlike PH, the degree-day scheme in A3D does not bring any particular advantage in terms of required input variables and computational speed, but it is only used here for sensitivity analysis purposes.

All models used here obtain lower performance in the Mera catchment than in the Dischma. The explanation is twofold. First, the Mera catchment is highly regulated by dams, which is not accounted for in the models. The errors of the models (overestimation of runoff in summer and underestimation in winter) match the pattern of dams usage: water retention in summer to produce electricity in winter during the demand peak. Besides regulation, another key point affecting the representation of discharge in the Mera catchment could be the scarcity of input data. Lack of data might be an impairment for models with higher degree of complexity like A3D, as they are more sensitive to poor interpolation of meteorological data (Bougamont et al., 2007; Magnusson et al., 2011; Schlögl et al., 2016). The scarcity of meteorological input forcing in the Mera catchment, compared to the Dischma, could explain why A3D does not outperform PH model there. In addition to the different representation of snow, SF and PH$_R$ use completely different water routing schemes. While SF uses a semi-distributed approach, the approach of PH$_R$ is purely conceptual. The important difference we observe in the snow melt seasonality between the two models do not allow comparison on the performance of the water routing schemes. Such comparison should be done by forcing both water routing model with the same input, which was not possible here since the water routing module of PH is deeply linked to the snow melt module in the model. In summary, from models comparison over the validation period, we conclude that:

- Both versions of A3D outperform PH in the simulation of snow cover. The PH model is effectively underestimating the contribution of snow and ice melt during the melting season;

- Such underestimation has strong repercussions on the correct representation of the sharp and narrow discharge peak characterising nival rivers' hydrograms;

- The degree-day version of A3D obtains slightly lower performances in the representation of snow, but these differences are compensated during the calibration of SF and the simulated runoff is then very similar;

- Input data scarcity in the Mera catchment leads to lower performances with A3D, while all models fails at correctly representing the discharge in the Mera catchment (possibly also in response to water regulation from dams).

For the upcoming application to climate change, the differences observed between PH and A3D must be considered. Indeed, the model PH, despite an extensive calibration and the usage of ad-hoc values for many parameters corresponding to nowadays conditions (see Tab. 7), exhibits a poorer representation of the snow melt contribution to runoff. This is a strong indication that snow melt may be incorrectly captured in the future by such models when the climate regime will exhibit substantial change. Additionally, the temperature melt factors are assumed to be constant for the whole basins and throughout the years, even if this approach could be wrong, especially for climate change applications. However, large basins such as Mera often range from low to high and inaccessible elevations, where data accessibility is likely limited, thus justifying the need to use simpler schemes for such applications. A comparison of the full solver of A3D and PH when forced with climate change scenarios is presented in the next section.

4.3 Climate change

Climate change simulations are run over the period 2080-2090 and forced by the set of RCMs listed in Tab. 4. In the following, we present the predicted evolution of precipitation, snow melt (as of the most significant seasonal forcing for total runoff and the major difference between the two melt schemes), discharge and total volumes with the aim of discussing the sensitivity

5 of the models to climate change. Glacier melt will not be covered because its influence by 2050-2060 and later is predicted to be extremely small due to glacier retreat. In the following, A3D's degree-day configuration will not be used anymore since (1) as mentioned before, there are no real computational advantages in using such scheme with respect to the full solver and (2) differences with respect to A3D are very small, although they could as well increase in a future changed climate. For precipitation and snow melt, climate change simulations are compared against the mean over the reference period 1990-2000,

10 whereas for discharge and volumes they are compared to the observed time series. The reference period is used here with the only aim of showing seasonality changes induced by climate change, as the models' differences in reproducing melt were already discussed throughout Section 4.2.

Predicted evolution of precipitation, snow melt, discharge and volume patterns are shown in Fig. 10 and Fig. 11 for Mera and Dischma respectively. Boxplots in Fig. 12 and Fig. 13 show deviations of the RCP8.5 mean from the mean over the reference

15 period.

Both models show identical precipitation predictions, thus excluding possible downscaling errors (Fig. 10 and Fig. 11, panels (a)).

Changes in snow melt predicted by A3D and PH are consistent at a seasonal level: with respect to the reference mean over 1990-2000, a net increase is predicted in spring (most evidently in the nivo-glacial Dischma catchment), a net decrease in

20 late summer (August to September) and late autumn (October to November), an increase during winter months (December to March). Such findings are in line with previous studies, e.g. Bavay et al. (2013) and Kobierska et al. (2013). Although tendencies are similar amongst the two melt models, the magnitude of changes is rather different. The energy-balance model A3D appears to be more sensitive to climate change than the degree-day model PH: not only the predicted snow melt is significantly higher (up to 50%), but also changes with respect to the reference period are more marked, regardless of the catchment

25 (see Fig. 12 and Fig. 13, panels (b)). Differences in melting magnitude and change are mostly ascribed to spring and summer months. To explain this, the basis of the degree-day melt scheme needs to be discussed again. The main physical processes involved in the energy balance are incoming and outgoing longwave radiation, absorbed shortwave radiation, turbulent heat fluxes and melt (Ohmura, 2001; Zappa et al., 2003). The degree-day estimation approach of snow/glacial melt of PH relies on air temperature, which is considered to be a good predictor for the majority of energy fluxes taking part in the balance,

30 and on absorbed shortwave radiation (see Equation 2), which is incompletely described by air temperature and whose effect is included through calibrated parameters. Therefore, there are two substantial limitations in the use of a degree-day melting scheme for climate change applications. On the one hand, snow degree-days parameters of PH are calibrated and fixed for each month (see Tab. 6). As a consequence, the model gains less sensitivity towards air temperature changes and it cannot capture changes in melt seasonality under climate change conditions, leading to possibly altered results. On the other hand, the gener-
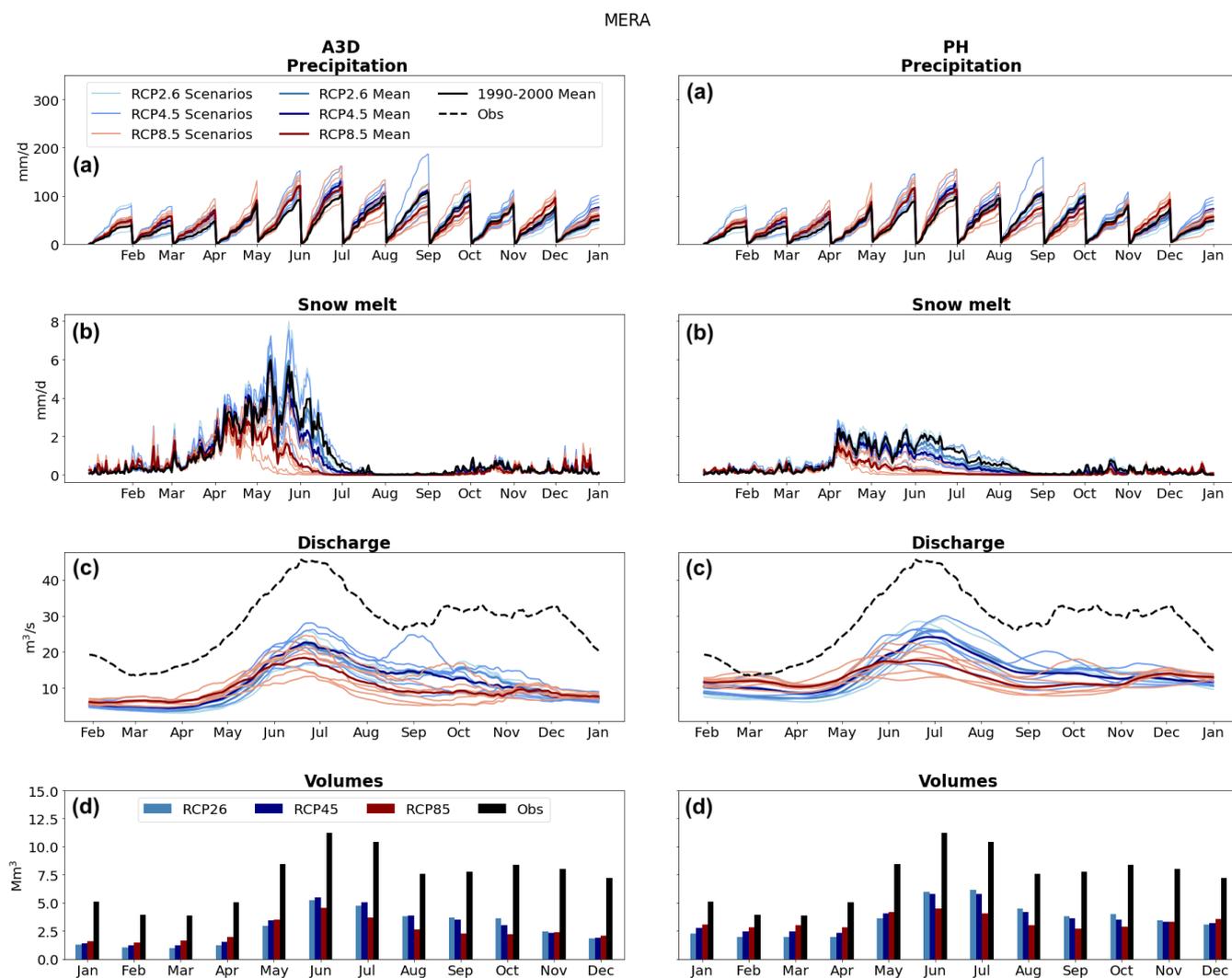
Figure 10. Climate change scenarios over Mera catchment for 2080-2090 against the reference mean over 1990-2000 for precipitation (panels (a)), snow melt (panels (b)), discharge (panels (c)) and volumes (panels (d)) predicted by A3D (left panels) and PH (right panels). Precipitation and volumes are presented as average monthly cumulates. Snow melt is presented as the average year. Discharge is presented as the average year, smoothed by a moving average with a window of 30 days.
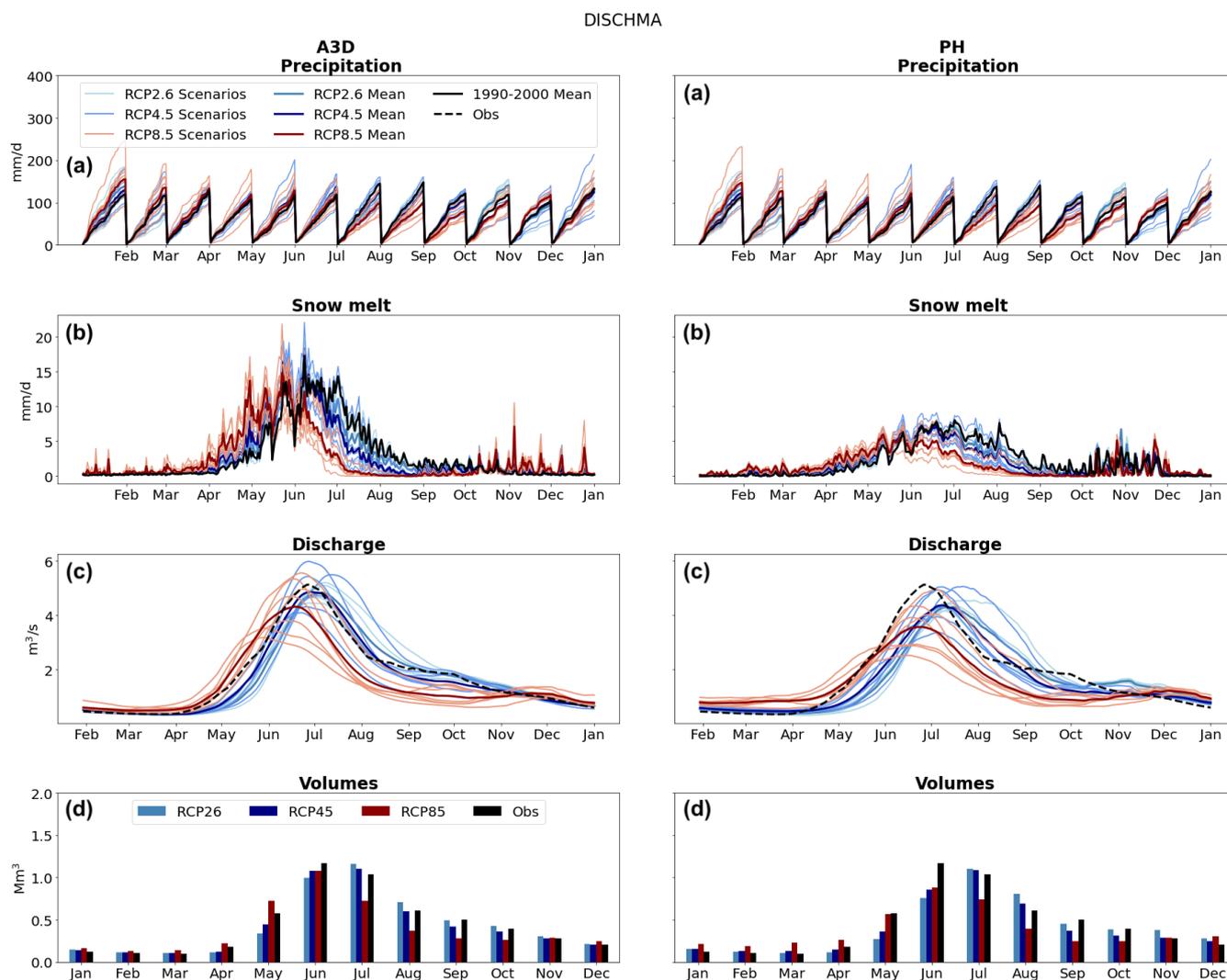
Figure 11. Climate change scenarios over Dischma catchment for 2080-2090 against the reference mean over 1990-2000 for precipitation (panels (a)), snow melt (panels (b)), discharge (panels (c)) and volumes (panels (d)) predicted by A3D (left panels) and PH (right panels). Precipitation and volumes are presented as average monthly cumulates. Snow melt is presented as the average year. Discharge is presented as the average year, smoothed by a moving average with a window of 30 days.
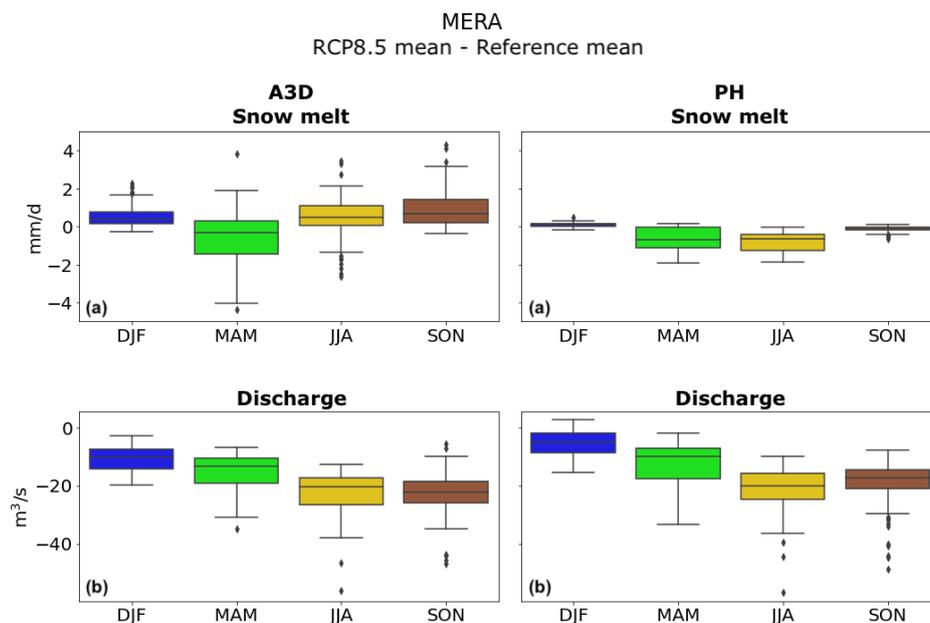
Figure 12. Boxplot illustrating the difference between the 2080-2090 RCP8.5 scenario mean and the reference mean (1990-2000) for snow melt (panels (a)) and discharge (panels (b)) predicted by A3D (left panels) and PH (right panels) for Mera catchment.
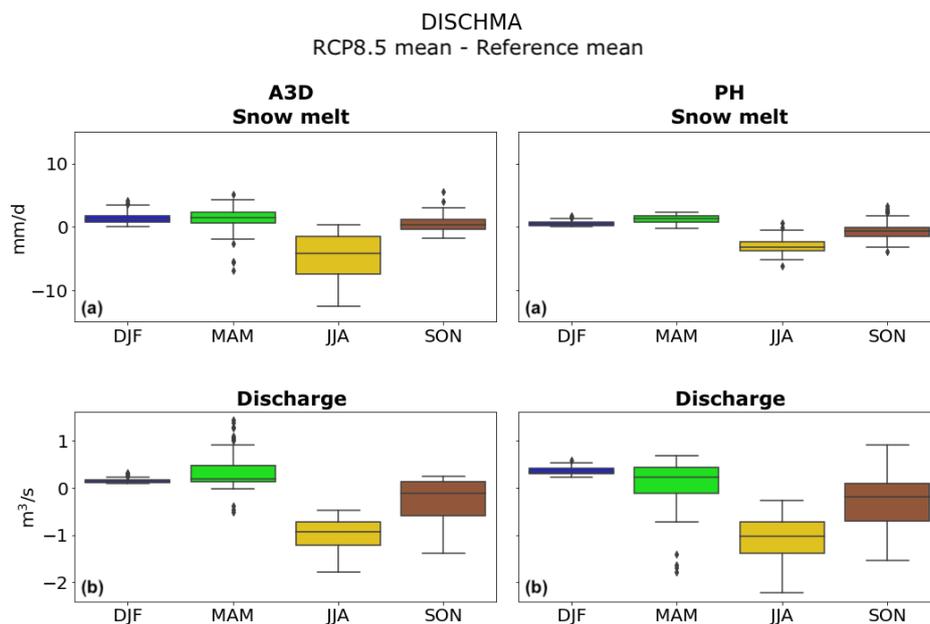
Figure 13. Boxplot illustrating the difference between the 2080-2090 RCP8.5 scenario mean and the reference mean (1990-2000) for snow melt (panels (a)) and discharge (panels (b)) predicted by A3D (left panels) and PH (right panels) for Dischma catchment.

ally higher temperatures predicted for the end of the century may lead to (1) less snowfall and a faster aging of the snow cover and (2) an earlier melt of the seasonal snow cover. Both effects induce a lower mean albedo over the melting period, as old snow has lower albedo than new snow. Within A3D, snow albedo is computed at each time step with the specific sub-module of SNOWPACK. Conversely, albedo is set as a fixed parameter within the model PH (see Tab. 7), hence any albedo decrease

5   could not be captured. As a consequence, the contribution of net shortwave radiation is likely underestimated, leading to lower melt rates in spring and summer.

The predicted shift in melting season affects the seasonality of the discharge regime for both catchments. Figure 10 and 11, panels (c), show an increase in late autumn and winter discharge for both catchments, which is likely linked to the enhanced snow melt in late autumn months. Such findings are in line with many previous studies about climate change in mountain

10   catchments, Michel et al. (2021a) among those. However, changes in discharge seasonality are more pronounced when looking at A3D over Dischma catchment. Even if the discharge peak timing is not expected to shift significantly, the RCP8.5 mean reproduces a discharge curve which is shifted by one month with respect to the average observations (see Figure 11, left panel (c)). The same pattern is reproduced by the predicted monthly volumes (see Figure 11, left panel (d)). A greater magnitude of change in spring and summer months predicted by A3D is also shown in Fig. 12 and Fig. 13, especially for the nivo-glacial

15   Dischma catchment.

Another interesting aspect that this analysis brings to light concerns predicted discharge magnitude for the Mera catchment. As shown by panels (c) and (d) in Fig. 10 illustrating discharge and volumes, both models reproduce a significant reduction in river's streamflow magnitude. This behaviour is only ascribed to Mera river as no mass seems to be lost for Dischma (Fig. 11, panels (c) and (d)), where the curve is shifted by new climatic conditions but the volumes are overall preserved.

20   Discharge magnitude reductions go as far as -85% for RCP2.6, -82% for RCP4.5, -86% for RCP8.5 predicted by A3D, and -77%, -78% and -82% predicted by PH$_R$. We ascribe this peculiar inability of both models to correctly reproduce base flow in Mera catchment under climate change to the river's regulation and to human activity in general. If both models show coherent volumes for a nearly natural catchment such as Dischma, it is likely that for a strongly regulated catchment such as Mera, the calibration process is the one that most influences the models' ability to reproduce volumes correctly. Under changed climatic

25   condition, the parameters that have been calibrated or fixed under current climate may fail to reproduce the base flow correctly.

To test this hypothesis, one may refer to the predicted evapotranspiration over the catchments. At the time when these simulations were run, this output was not yet implemented within A3D, so that it is impossible to make a comparison with the output from PH. In the meantime, this information was outputed (Michel et al., 2021a), but nevertheless it is shown there that evapotranspiration term will not have a great influence in cryospheric areas.

30   To conclude, our findings are twofold. On the one hand, the analysis on the almost natural Dischma catchment confirmed what previous studies discussed before (Hock, 2005; Magnusson et al., 2010), i.e. that the use of degree-day models for future hydro-climatic studies is questionable since they rely on parameters which are calibrated in current climatic conditions and on a partial representation of the energy balance, whose inner physical processes won't change under changing climate. On the other hand, the case of Mera suggested that the use of either complex or simple melt models coupled with routing modules

35   alone might be disputable when applied to strongly regulated, extensive catchments, because the calibration process might

influence models' predictive ability more profoundly than the energy balance alone. Henceforth, our guess is that hydropower reservoir operation cannot be neglected in similar contexts and need to be accounted for in modelling.

5   Conclusions

This paper compares the discharge response of two Alpine catchments to present conditions and climate change, predicted by
5 one energy balance and two degree-day melt models: A3D, A3D$_{DD}$ and PH respectively. The two catchments, Mera and Dischma, are different in size and extent of water resources exploitation. Under present climatic condition, 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 snow melt 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,
10 hydropower operations (i.e. water release in winter to produce electricity when demand is peaking and subsequent withholding in summer for accumulation purposes) can reduce model's 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 significant difference between models for Mera, with scores being satisfactory but not outstanding. The explanation is twofold. On the one
15 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 is a bigger problem 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
20 outperform PH$_R$ in about all seasons and all catchments. Interestingly, in terms of snow melt 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, so it is always determined by a full energy balance model in the first place and it 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 advantages
25 of being a simplification of a multi-layer snow model in the first place.

Under climate change, end-of-century changes in snow melt 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
30 monthly degree-days compromise the model's ability to perceive seasonal changes in snow melt and (2) albedo changes cannot be captured, thus the contribute of net shortwave radiation might be underestimated. Such findings are consistent with previous literature findings.

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 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 influence. Our interpretation is that the calibration process for strongly regulated catch-

5 ments as Mera overshadows the benefits of a full energy balance scheme showing good performances in reproducing snow melt. Thus, we conclude that anthropization and hydropower exploitation of Alpine catchments might significantly hinder models' performances under climate change regardless the sophistication level of the melting scheme, as calibration could embody the limiting factor.

Author contributions.

10 FCarletti participated in the design of the study, performed the data analysis and wrote the paper. AM performed the Alpine3D and StreamFlow simulations, significantly helped the design of the study and participated to the analysis and the redaction. FCasale performed Poli-Hydro simulations and participated to the analysis and the redaction. ML and DB helped design the study, supported in data analysis and re-reading. This paper builds upon previous work done by MB and AB. All the authors discussed the results and contributed to the writing of the final paper.

15 *Competing interests.*

The authors declare that they have no conflict of interest.

Acknowledgements. The work presented in the paper was in fulfilment of the Interreg Project GERIKO-MERA, kindly acknowledged.

The Federal Office of Meteorology and Climatology (MeteoSwiss), the WSL Institute for Snow and Avalanche Research (SLF), the Swiss Federal Office of Topography (Swisstopo), the Federal Office for the Environment (FOEN), ARPA Lombardia are acknowledged for the free

20 access to their data.

The vast majority of this work was performed with open and free languages and softwares (mainly bash, Python, MeteoIO, SNOWPACK, Alpine3D, StreamFlow, TauDEM, GDAL, and QGIS, along with countless libraries), and all the authors acknowledge the open-source community for its invaluable contribution to science.

Alpine3D simulations were performed on the Piz Daint supercomputer of the Swiss National Supercomputing Center (CSCS). The CSCS

25 technical team is acknowledged for his help and support during this project.

The Climate Lab of Politecnico di Milano is acknowledged for support with the use of Poli-Hydro model.

Harry Zekollari is acknowledged for performing the GloGEMflow simulations.

Code and data availability.

All codes and results developed and produced throughout this paper are available upon request to the main author.

The source code for MeteoIO, SNOWPACK, Alpine3D, and StreamFlow are available at https://models.slf.ch. The following versions have been used in this work: MeteoIO 3.0.0 (rev 2723), SNOWPACK 3.6.1 (rev 1878), Alpine3D 3.2.1 (rev 570), and

5   StreamFlow 1.2.2 (rev 368).

The source code of the model GloGEMflow can be obtained upon request to the corresponding author.

The downscaled climate change scenarios are available on Envidat (https://www.envidat.ch/#/metadata/climate-change-scenarios-at-hourly-resolution).

References

Agency, E. E.: CORINE Land Cover (CLC) 2006, https://land.copernicus.eu/ pan-european/corine-land-cover/clc-2006, 2013.

Aili, T., Soncini, A., Bianchi, A., Diolaiuti, G., D'Agata, C., and Bocchiola, D.: Assessing water resources under climate change in high-altitude catchments: a methodology and an application in the Italian Alps, Theoretical and Applied Climatology, 135, 135–156, https://doi.org/10.1007/s00704-017-2366-4, https://doi.org/10.1007/s00704-017-2366-4, 2019.

Arduini, G., Balsamo, G., Dutra, E., Day, J., Sandu, I., Boussetta, S., and Haiden, T.: Impact of a Multi-Layer Snow Scheme on Near-Surface Weather Forecasts, Journal of Advances in Modeling Earth Systems, 11, https://doi.org/10.1029/2019MS001725, 2019.

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.

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.

Bavay, M., Grünewald, T., and Lehning, M.: Response of snow cover and runoff to climate change in high Alpine catchments of Eastern Switzerland, Advances in Water Resources, 55, 4–16, https://doi.org/https://doi.org/10.1016/j.advwatres.2012.12.009, http://www.sciencedirect.com/science/article/pii/S0309170812003193, 2013.

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.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C.: Physical constraints for temperature biases in climate models, Geophysical Research Letters, 40, 4042–4047, https://doi.org/https://doi.org/10.1002/grl.50737, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50737, 2013.

Boberg, F. and Christensen, J.: Overestimation of Mediterranean summer temperature projections due to model deficiencies, Nature Climate Change, 2, 433–436, https://doi.org/10.1038/nclimate1454, 2012.

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.

Bombelli, G. M., Soncini, A., Bianchi, A., and Bocchiola, D.: Potentially modified hydropower production under climate change in the Italian Alps, Hydrological Processes, 33, 2355–2372, https://doi.org/https://doi.org/10.1002/hyp.13473, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.13473, 2019.

Bougamont, M., Bamber, J., Ridley, J., Gladstone, R., Grueell, W., Hanna, E., Payne, A., and Rutt, I.: Impact of model physics on estimating the surface mass balance of the Greenland Ice Sheet, Geophys. Res. Lett., 34, L17 501, https://doi.org/10.1029/2007GL030700, 2007.

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.

Casale, F., Bombelli, G. M., Monti, R., and Bocchiola, D.: Hydropower potential in the Kabul River under climate change scenarios in 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.

Christensen, J., Boberg, F., Christensen, O., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophysical Research Letters, https://doi.org/10.1029/2008GL035694, 2008.

Criss, R. and Winston, W.: Do Nash Values Have Value? Discussion and Alternate Proposals, Hydrological Processes, 22, 2723–2725, https://doi.org/10.1002/hyp.7072, 2008.

5 Déry, S. J., Stahl, K., Moore, R. D., Whitfield, P. H., Menounos, B., and Burford, J. E.: Detection of runoff timing changes in pluvial, nival, and glacial rivers of western Canada, Water Resources Research, 45, https://doi.org/10.1029/2008WR006975, 2009.

ESRI: Environmental Systems Research Institute (ESRI) / ArcGIS Release 10.8.1, 2012.

Farinotti, D., Usselmann, 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 Ficchì, A., Perrin, C., and Andréassian, V.: Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events, Journal of Hydrology, 538, https://doi.org/10.1016/j.jhydrol.2016.04.016, 2016.

FOEN: Hydro-CH2018, https://www.bafu.admin.ch/bafu/en/home/topics/water/info-specialists/state-of-waterbodies/state-of-watercourses/water-flow-and-flow-regime-in-watercourses/climate-change-and-hydrology/hydro-ch2018.html, 2018.

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

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.

Garrick, M., Cunnane, C., and Nash, J. E.: A criterion of efficiency for rainfall-runoff models, Journal of Hydrology, 36, 375–
20 381, https://doi.org/https://doi.org/10.1016/0022-1694(78)90155-5, http://www.sciencedirect.com/science/article/pii/0022169478901555, 1978.

Gouttevin, I., Lehning, M., Jonas, T., Gustafsson, D., and Mölder, M.: A two-layer canopy model with thermal inertia for an improved snowpack energy balance below needleleaf forest (model SNOWPACK, version 3.2.1, revision 741), Geoscientific Model Development, 8, 2379–2398, https://doi.org/10.5194/gmd-8-2379-2015, https://gmd.copernicus.org/articles/8/2379/2015/, 2015.

25 Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, http://dx.doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095–1107, https://doi.org/10.1175/2009BAMS2607.1, 2009.

30 Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, Climate Dynamics, 37, 407–418, https://doi.org/10.1007/s00382-010-0810-6, 2011.

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.

Hock, R.: Glacier melt: a review of processes and their modelling, Progress in Physical Geography: Earth and Environment, 29, 362–391,
35 https://doi.org/10.1191/0309133305pp453ra, https://doi.org/10.1191/0309133305pp453ra, 2005.

https://www.addaconsorzio.it/: Adda Consortium, https://www.addaconsorzio.it/.

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.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., et al.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional environmental change, 14, 563–578, 2014.

Knoben, W. J., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta
5   efficiency scores, Hydrology and Earth System Sciences, 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

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.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard,
10   E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014, https://gmd.copernicus.org/articles/7/1297/2014/, 2014.

Legates, D. R. and McCabe, G. J.: Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation, Water Resources Research, 35, 233–241, https://doi.org/10.1029/1998WR900018, 1999.

15   Lehning, M., Bartelt, P., Brown, B., and Fierz, C.: A physical SNOWPACK model for the Swiss avalanche warning Part III: Meteorological forcing, thin layer formation and evaluation, Cold Regions Science and Technology, 35, 169–184, https://doi.org/10.1016/S0165-232X(02)00072-1, 2002.

Lehning, M., Völksch, I., Gustafsson, D., Nguyen, T. A., Stähli, M., and Zappa, M.: ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology, Hydrological Processes, 20, 2111–2128, https://doi.org/10.1002/hyp.6204, http:
20   //infoscience.epfl.ch/record/170171, 2006.

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

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.

25   Maraun, D.: Nonstationarities of Regional Climate Model Biases in European Seasonal Mean Temperature and Precipitation Sums, Geophysical Research Letters, 39, 6706–, https://doi.org/10.1029/2012GL051210, 2012.

Martinec, J.: Indirect evaluation of snow reserves in mountain basins, Snow, 205, 111, document Type: Conference Paper Or Compendium Article, 1991.

Martinec, J. and Rango, A.: Merits of statistical criteria for the performance of hydrological models, JAWRA Journal of the American Water
30   Resources Association, 25, 421–432, https://doi.org/10.1111/j.1752-1688.1989.tb03079.x, 1989.

MeteoSwiss: CH2011, Swiss Climate Change Scenarios, Technical Report, National Centre for Climate Services, Zurich, 2011.

MeteoSwiss: CH2018, Climate Scenarios for Switzerland, Technical Report, National Centre for Climate Services, Zurich, 2018.

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.

Michel, A., Sharma, V., Lehning, M., and Huwald, H.: Climate change scenarios at hourly time-step over Switzerland from an enhanced temporal downscaling approach, International Journal of Climatology, 41, 3503–3522, https://doi.org/https://doi.org/10.1002/joc.7032, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.7032, 2021b.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/https://doi.org/10.1016/0022-1694(70)90255-6, http://www.sciencedirect.com/science/article/pii/0022169470902556, 1970.

of the Environment, S. F. O.: L'ordre des cours d'eau selon Strahler pour le réseau hydrographique numérique au 1:25'000 de la Su-
5 isse, https://www.bafu.admin.ch/bafu/fr/home/themes/eaux/etat/cartes/reseau-hydrographique-suisse/ reseau-hydrographique–ordre-des-cours-deau-pour-le-reseau-hydro.html, 2013.

Ohmura, A.: Physical Basis for the Temperature-Based Melt-Index Method, Journal of Applied Meteorology, 40, 753 – 761, https://doi.org/10.1175/1520-0450(2001)040<0753:PBFTTB>2.0.CO;2, https://journals.ametsoc.org/view/journals/apme/40/4/1520-0450_2001_040_0753_pbfttb_2.0.co_2.xml, 2001.

10 Omstedt, A.: A coupled one-dimensional sea ice–ocean model applied to a semi-enclosed basin, Tellus A: Dynamic Meteorology and Oceanography, 42, 568–582, https://doi.org/10.3402/tellusa.v42i5.11899, https://doi.org/10.3402/tellusa.v42i5.11899, 1990.

Rajczak, J., Kotlarski, S., Salzmann, N., and Schär, C.: Robust climate scenarios for sites with sparse observations: A two-step bias correction approach, International Journal of Climatology, 36, 1226–1243, https://doi.org/10.1002/joc.4417, 2016.

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

Schaefli, B. and Gupta, H.: Do Nash values have value?, Hydrological Processes, 21, https://doi.org/10.1002/hyp.6825, 2007.

Schaefli, 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.

20 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.

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.

Soncini, A., Bocchiola, D., Azzoni, R., and Diolaiuti, G.: A methodology for monitoring and modeling of high altitude Alpine catchments, Progress in Physical Geography: Earth and Environment, 41, 393–420, https://doi.org/10.1177/0309133317710832, https://doi.org/10.1177/0309133317710832, 2017.

30 Stucchi, L., Bombelli, G. M., Bianchi, A., and Bocchiola, D.: Hydropower from the Alpine Cryosphere in the Era of Climate Change: The Case of the Sabbione Storage Plant in Italy, Water, 11, https://doi.org/10.3390/w11081599, https://www.mdpi.com/2073-4441/11/8/1599, 2019.

Tarboton, D.: TauDEM, Utah State University http://hydrology.usu.edu/taudem/taudem5/, 1997.

Terzago, S., Andreoli, V., Arduini, G., Balsamo, G., Campo, L., Cassardo, C., Cremonese, E., Dolia, D., Gabellani, S., Hardenberg, J.,
35 Morra di Cella, U., Palazzi, E., Piazzi, G., Pogliotti, P., and Provenzale, A.: Sensitivity of snow models to the accuracy of meteorological forcings in mountain environments, Hydrology and Earth System Sciences, 24, 4061–4090, https://doi.org/10.5194/hess-24-4061-2020, 2020.

IDAWEB: MeteoSwiss, Federal Office of Meteorology and Climatolgy, 2019.

IMIS: WSL Institute for Snow and Avalanche Research, SLF, IMIS measuring network, 2019.

Usda., Scs., .: Urban Hydrology for Small, Soil Conservation, p. 164, http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle: Urban+Hydrology+for+Small+watersheds{#}1, 1986.

Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R.: An introduction to the hydrological modelling system PREVAH and its pre- and post-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.

von Rütte, F., Kahl, A., Rohrer, J., and Lehning, M.: How Forward-Scattering Snow and Terrain Change the Alpine Radiation Balance With Application to Solar Panels, Journal of Geophysical Research: Atmospheres, 126, e2020JD034 333, https://doi.org/https://doi.org/10.1029/2020JD034333, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD034333, e2020JD034333 2020JD034333, 2021.

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.

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.

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.

Zekollari, H., Huss, M., and Farinotti, D.: Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble, The Cryosphere, 13, 1125–1146, https://doi.org/10.5194/tc-13-1125-2019, https://tc.copernicus.org/articles/13/1125/2019/, 2019.

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.