TopoCLIM: Rapid topography-based downscaling of regional climate model output in complex terrain v1.0

Joel Fiddes[1], Kristoffer Aalstad[2], and Michael Lehning[1,3]

[1]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[2]Department of Geosciences, University of Oslo, P.O. Box 1047, Blindern, 0316 Oslo, Norway
[3]CRYOS, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Correspondence: Joel Fiddes (joel.fiddes@slf.ch)

Abstract. This study describes and evaluates a new downscaling scheme that specifically addresses the need for hillslope-scale atmospheric forcing timeseries for modeling the local impact of regional climate change projections on the land surface in complex terrain. The method has a global scope in that it does not rely directly on surface observations and is able to generate the full suite of model forcing variables required for hydrological and land surface modeling at hourly timesteps. It achieves this by utilising the previously published TopoSCALE scheme to generate synthetic observations of current climate at the hillslope scale, while accounting for a broad range of surface-atmosphere interactions. These synthetic observations are then used to debias (downscale) CORDEX climate variables using the quantile mapping method. A further temporal disaggregation step produces sub-daily fields. This approach has the advantages of other empirical-statistical methods, namely speed of use, while avoiding the need for ground data, which is often limited. It is therefore a suitable method for a wide range of remote regions where ground data is absent, incomplete, or not of sufficient length. The approach is evaluated using a network of high elevation stations across the Swiss Alps and a test application of modelling climate change impacts on Alpine snow cover is given.

1 Introduction

Climate change has caused, and will continue to cause significant changes in the global cryosphere with increasing impacts likely in a wide range of domains (Hock et al., 2019). Where observational records of sufficient length exist, we are able to quantify these changes. Such records are often curated as part of national or international networks e.g. the World Meteorological Organisation's Global Cryosphere Watch, World Glacier Monitoring Service or the Global Climate Observing System. However, such locations are globally sparse, with particular observational gaps in remote regions due to technical difficulties and resources required to maintain monitoring infrastructure.

In order to obtain possible scenarios of future conditions we are reliant on climate models. However, for meaningful impact studies climate time-series are often required at higher spatial and temporal resolutions than currently available from Global or even Regional Climate Models (GCMs/ RCMs). This is especially the case in heterogeneous terrain such as mountain regions where topographic and therefore climate variability is high over short horizontal distances. High surface variability requires

modelling at the hillslope scale (c. 100 m) in order to adequately capture fluxes and stores of energy, water and carbon (Fan et al., 2019). Various methods of downscaling can be utilised to achieve this goal. Dynamical downscaling typically applies an RCM or numerical weather prediction model (e.g. **(alias?)**) at high resolution over a limited area in order to obtain more detailed process representation. This requires no additional data beyond a boundary forcing, yet is computationally costly (normally a supercomputer is required) and is therefore generally applied to limited domains and/or time periods. In addition, an extensive set of boundary fields are required in order to set up a run that are not normally available via standard distribution portals such as Earth System Grid Federation (ESGF), further complicating possible studies. Empirical-statistical downscaling (ESD) approaches are typically computationally cheap to run yet require extensive and robust ground observations that are often either not available, of uncertain quality, or distributed unequally according to important gradients such as elevation. In addition, timeseries from many stations are not available at climate timescales - typically 30 years (Arguez and Vose, 2011), rendering the application of ESD methods problematic. Furthermore, many modern physically based impact models require a full suite of forcing variables to drive them, usually at sub-daily time-steps. Such requirements are rarely met by initiatives that have provided input for impact studies (Michel et al., 2021). It is increasingly recognized that analysis of extremes, and not just mean values, is required to fully quantify the impact of climate change (Katz and Brown, 1992). By definition this requires highly temporally resolved forcing, as climatic extremes often occur over short timescales, e.g. daily maximum temperature or storm peak that require sub-daily simulations.

Climate model timeseries, even from the latest generation of RCMs, typically exhibit bias (systematic deviations) when evaluated against observations (Ivanov et al., 2018; Kotlarski et al., 2014). These biases need to be corrected before climate timeseries can be used to force a locally applied impact model (Wood et al., 2004). However, in impact studies we are typically interested in a climate change signal (CCS), which is a quantitative measure of the difference between a future climate state and historical reference period (Themeßl et al., 2012). Bias correction (BC) can modify the CCS (Ivanov et al., 2018; Themeßl et al., 2012) which has been a subject of discussion and often seen as a deficiency in BC methods (Hempel et al., 2013). However, it is recognised that model biases typically do not cancel out in the calculation of a CCS and therefore its modification under BC has been interpreted recently as an enhancement rather than a deficit, particularly in intensity dependent biases which characterise variables such as precipitation (Gobiet et al., 2015).

There are a wide range of BC methods (Gutmann et al., 2014) with perhaps the most established and widely used being quantile mapping (QM) which has been shown to perform favourably in comparison studies (Teutschbein and Seibert, 2013; Themeßl et al., 2011) and found to cope well with with non-stationary conditions - removing the restrictive stationarity assumption in climate BC. It is also one of the few methods able to correct wet day frequency and intensity (Déqué, 2007). QM is a distribution based BC method that removes quantile dependent biases with respect to a reference period (Ivanov and Kotlarski, 2017), it therefore corrects the variance and not just the mean. It should be noted that if the climate variable and reference are at the same spatial scale a pure bias correction is applied, whereas if the reference is at a finer scale (e.g. a meteorological station) then an implicit downscaling is achieved, which also makes it a class of empirical-statistical downscaling (ESD) methods.

While concerns over the deterministic nature of QM (Maraun, 2013) and effect on multi-day statistics (Addor and Seibert, 2014) are acknowledged, it is widely accepted to be a pragmatic approach to satisfy the requirements of impact models (Rajczak

et al., 2016) with deficits shared by all statistically-based methods. A useful overview of the issues involved with respect to impact modelling can be obtained from Stocker et al. (2015).

All ESD/BC methods require climate scale timeseries of observations (typically 30 years). In remote regions lacking sufficient historical observations this requirement can be difficult to satisfy. Atmospheric reanalysis data-sets have been proposed as a means to compensate for missing or incomplete observations (Cao et al., 2019; Fiddes and Gruber, 2014) in order to provide a "best-guess" of the current state. Moreover, global reanalysis datasets can form the basis for impact studies with a global consistency.

In this study we address the problem of lack of impact-model ready (i.e. hillslope scale) climate timeseries with a new modelling framework called "TopoCLIM". We use the latest ECMWF global reanalysis dataset ERA5 (Hersbach et al., 2020) together with the downscaling method TopoSCALE (Fiddes and Gruber, 2014) to provide a robust assessment of local-scale meteorological forcing for the reference period. Importantly, using these pseudo-observations we are able to debias climate timeseries in regions lacking ground observations. Furthermore, this method provides a full suite of forcing data required to run a numerical model at sub-daily timesteps.

By coupling this to the subgrid clustering scheme TopoSUB (Fiddes and Gruber, 2012) and the snow model FSM (Essery, 2015), we demonstrate the ability of this scheme to efficiently generate hillslope scale climate change maps of snow cover over the entire Swiss Alps. We test this scheme both with a detailed evaluation at the Weissfluhjoch meteorological station, and across the Swiss network of high elevation stations (IMIS) which has a large spatial coverage.

2 Methods

2.1 Overview

The scheme is implemented in Python with several specific sub-routines implemented in R (e.g. the quantile mapping package). An overview of the processing pipeline is given in Figure 1 and can be summarised as a three-step process: (1) A quasi-physical topography-based downscaling method TopoSCALE (Fiddes and Gruber, 2014) generates hillslope scale (defined by the DEM resolution) forcing time-series for the reference period from the ERA5 reanalysis, (2) the BC method quantile mapping (Gudmundsson et al., 2012) is used to statistically downscale (debias) a climate time series at the given point for which we now have a reference from downscaled ERA5 forcing, (3) a disaggregation scheme (Förster et al., 2016) generates hourly climate timeseries based on observed sub-daily distributions of meteorological variables. We demonstrate this approach by downscaling CORDEX RCM data at hillslope scale and additionally generalising this to a map product using the subgrid scheme, TopoSUB (Fiddes and Gruber, 2012), which efficiently spatialises 1D model results (multiple subgrid samples per CORDEX gridbox) to a map domain according to important dimensions of land surface heterogeneity. In this way, hillslope resolution (100 m) map results are generated with a laptop-feasible number of subgrid simulations (in this case 100) per large-scale CORDEX gridbox. The overall philosophy of this approach is to develop methods that are global in application and therefore can be used in data-poor regions, are efficient to run and repeat (i.e. the ability to rapidly repeat numerical experiments that have a relatively low computational cost), and yet address the key drivers of climate-surface variability in complex terrain.

2.2 Preprocessing

The CORDEX data download is achieved using a custom tool built around the ESGF Python client. This is not a trivial task due to the large data volumes involved and variable uptime of data nodes. All preprocessing of raw CORDEX data (Figure 1): concatenating NetCDF time series, extracting region of interest and regridding from rotated pole projections, is accomplished using standard tools from the Climate Data Operators (CDO) suite. The CDO tools are called from the preprocessing module of TopoCLIM and not used as standalone command line tools - enhancing the ease of use and reproducibility of the processing pipeline. As downloads are done by variable, it is possible that the full set of required forcing variables are not available for a given CORDEX model chain. If this is the case the model chain is excluded from any further processing.

Climate model calendars are often simplified for numerical reasons and are inherited from the parent GCM, these need to be correctly handled to produce comparable timeseries. Three calendars exist in the CORDEX data used here: "360-day" (every month is 30 days long), "365-day" (no leap year), and "standard" (complete Gregorian calendar). We convert all calendars to "standard" by linearly scaling dates to the standard Gregorian calendar and then gap filling missing data by linear interpolation.

For example, during the conversion from a "360-day" to a "standard" calendar, the output from the linear scaling will result in a 365 day timeseries (in the case of non-leap year) and be missing the following dates: January 31st, March 31st, June 1st, July 31st, September 30th and November 30th. In a second step these dates are gap-filled using linear interpolation.

2.3 Spatial downscaling of reanalysis data

ERA5 reanalysis “observations” are downscaled using the TopoSCALE scheme (Fiddes and Gruber, 2014) to be used as the reference data in the bias correction. We acknowledge that reanalyses are not true observations (Parker, 2016), yet by assimilating an extensive set of observations into a Numerical Weather Prediction (NWP) model, reanalyses are often considered to give the best possible view of the global climate (Dee et al., 2014). TopoSCALE performs a 3D interpolation of atmospheric fields available on pressure levels, to account for time varying lapse rates, and a topographic correction of radiative fluxes. The latter includes a cosine correction of incident direct shortwave radiation on a slope, adjustment of diffuse shortwave and longwave radiation by the sky view factor, and elevation correction of both longwave and direct shortwave. It has been extensively tested in various geographical regions and applications, e.g. permafrost in the European Alps (Fiddes et al., 2015), permafrost in the North Atlantic region (Westermann et al., 2015), Northern hemisphere permafrost (Obu et al., 2019), Antarctic permafrost (Obu et al., 2020), Arctic snow cover (Aalstad et al., 2018), Arctic climate change (Schuler and Østby, 2020), and Alpine snow cover (Fiddes et al., 2019). This approach enables us to provide a climate length pseudo-observation timeseries globally, while accounting for the main topographic effects on atmospheric forcing. We call this product T-MET throughout the text. It should be noted that this serves as our reference throughout the paper. A detailed validation and quantification of uncertainty of the TopoSCALE method is given in Fiddes and Gruber (2014) and therefore is not repeated here. We do however compare our results to the station data variables air temperature and snow depth across the IMIS network.

2.4 Quantile mapping

Bias correction through quantile mapping $Q : x \rightarrow x^{\star}$ is achieved as follows (Panofsky and Brier, 1968):

$$x^{\star} = Q(x) = F_o^{-1}\left[F_m\left(x\right)\right], \tag{1}$$

where the debiased output variable $x^{\star}$ is obtained by applying the quantile mapping function Q to the biased input variable x. This function is generally formed through the composition of the modeled cumulative distribution function (F_m) and the inverse of the observed cumulative distribution function (F_o). In our case, F_o is obtained from the pseudo-observations in the form of downscaled ERA5 data, while F_m is generated from the CORDEX output which we want to bias correct.

We use the R package "QMAP" for this purpose (Gudmundsson et al., 2012). Gudmundsson et al. (2012) compared different implementations of QM for daily precipitation data and found that a non-parametric empirical approach (as implemented in the cited package) outperforms implementations relying on theoretical distributional assumptions. While quantile mapping ensures that quantile biases are corrected in the CDF, it does not account for seasonally varying bias. It is therefore well suited to air temperature where we can be reasonably sure that winters are cold and summers hot, at least in mid to high latitudes. However, with precipitation the intra-annual distribution can be biased while the CDF may look reasonable (e.g. wet and dry season timing could be shifted). We address this with a two step approach called QM_MONTH. We split the data according to 12 temporal subsets corresponding to the months of the calendar year and run the QM algorithm on each subset, computing QM parameters separately for each month. These are applied throughout the historical and climate time series at the appropriate months. Similar approaches have been used successfully in other studies, e.g. Hanzer et al. (2018). Quantile mapping is performed on a subset of data during 1980-1995 to allow an evaluation to be performed over the time period 1996-2006. It should be noted that these periods are constraints imposed by the datasets used in this study and can be changed in other applications of the method.

While the variables are bias corrected independently, they are corrected towards a physically consistent dataset in the form of downscaled ERA5 data. We argue that the method does not produce physically inconsistent results, despite being univariate. The validation during the current climate also supports this claim (Table 3-5).

2.5 Temporal downscaling of climate time series

The final step in preparing the climate forcing is a temporal disaggregation to generate required sub-daily fields. Original hourly resolution T-MET are used to temporally disaggregate the downscaled (quantile mapped) daily climate timeseries. The "Melodist" package is used for this purpose (Förster et al., 2016). This disaggregates daily data based on observed sub-daily distributions. It should be noted that this assumes sub-daily distributions are stationary in a future climate, which quite possibly may not be the case. However, a greater source of uncertainty likely exists in the ability of ERA5 to reproduce short timescale local weather patterns which require convection resolving model resolutions (Liu et al., 2017) and appropriate physics.

Variables are disaggregated with the Melodist methods listed in Table 1 (Förster et al., 2016). Melodist does not however, provide methods for air pressure or longwave radiation, these are handled with the following procedure: taking advantage of the relationship between incoming longwave radiation (ILWR) and air temperature (TA),

ILWR = $\epsilon \times \sigma \times TA^4$

where σ is the Stefan-Boltzmann constant ($5.67 \times 10^{-8} Wm^{-2} K^4$) we diagnosed the daily all sky emissivity (ϵ). We then used ϵ as a daily scaling factor to convert disaggregated TA into ILWR. This procedure therefore assumes a constant ϵ at sub daily timestep (which of course will not normally be true) yet ensures that ILWR scales correctly with TA. Therefore higher TA lead to higher ILWR values and vice-versa. Air pressure is simply linearly interpolated to the sub-daily timestep.

3 Study region and data

3.1 Model domain

We consider two scales in this study (a) point-scale (meteorological stations: Weissfluhjoch and the IMIS network) and (b) regional (Swiss Alps), in order to illustrate typical applications of the scheme. A map of the study region and the location of the stations we used is given in Figure 2.

3.2 Climate data

The basic forcing comes from the regional climate model project CORDEX EUR-44 product (Jacob et al., 2014) at a nominal resolution of 44 km. The 44 km product was chosen over the 22 km product as this had many more model chains available. The EURO domain also has an 11 km product but this is not available globally, and therefore is not fit for the purpose of this study. Data was retrieved from the ESGF using an API and automated Python based tool developed in this study. This is an important step as the number of file downloads is large with dimensions being models $\times$ variables $\times$ scenarios $\times$ time periods. The fact that the ESGF consists of a distributed set of data nodes with variable uptime, further complicates the download process. We use daily data to force the scheme and retrieve historical data plus projections from two climate change scenarios, RCP2.6 which is a very stringent pathway, RCP2.6 requires that CO_2 emissions start declining by 2020 and go to zero by 2100 and RCP8.5 where emissions continue to rise throughout the 21st century. A full description of CORDEX variables is given in Table 1 and model chains used in Table 2. Daily data was chosen for the method (therefore requiring a temporal disaggregation step) as limited number of model chains are available at sub-daily resolutions. This is particularly the case outside of the EURO domain where use-cases for this method are envisaged. A higher number of model chains increases confidence in our results, by improved quantification of inter-model variability.

3.3 Reanalysis data

We use ECWMF's latest reanalysis product ERA5 (Hersbach et al., 2020), which uses version 41r2 of IFS which is the ECMWF NWP model. ERA5 represents an evolution over its predecessor, ERA-Interim, by increasing the model spatial resolution to 0.25 degrees, temporal resolution to hourly and the vertical model levels to 137 (37 pressure levels are stored). These reanalysis data are downscaled for the purpose of bias correction using the TopoSCALE scheme (Fiddes and Gruber, 2014) as described above (cf. Methods).

3.4 Topography

NASA's SRTM-3 90 m digital elevation model (DEM) is used as a topographical surface for TopoSCALE downscaling routines. Slope, aspect and sky-view factor (the portion of the sky hemisphere that is visible for a given DEM pixel) are derived (Dozier and Frew, 1990).

"A higher resolution DEM may also be used but likely does not add value as processes such as wind transport that operate on these scales are not included in the model (Mott et al., 2018). However, it should be noted that avalanching off steep slopes is accounted for by removing snow linearly above a slope threshold ((Fiddes et al., 2015)). Importantly in our scheme, higher resolution does not necessarily increase runtime either in point mode (trivially) or in spatial mode where the run-time in the main programme modules is related to number of TopoSUB clusters. Additionally, the scheme is designed to scale well by generating cluster forcings through array computations.

4 Simulation setup and evaluation

4.1 Simulation and maps

We use the Factorial Snow Model (FSM) (Essery, 2015) to simulate the snow cover and TopoSUB (Fiddes and Gruber, 2012) to spatialise results to a 2D map. FSM is a multi-physics ensemble model, however in this study we always use configuration 31 which is the most complex version of the model where all five parameterisations are switched on. TopoSUB is a topographic sampling scheme that reduces distributed modelling problems that explicitly model individual pixels to a lumped model several orders of magnitude smaller while considering the full range of topographic heterogeneity that exists. For example, a 100 m grid over a 25 km $\times$ 25 km domain would have 62500 pixels which would each represent a model simulation in a fully distributed setup of many surface models. TopoSUB would reduce this to 100-200 samples, each representing a single model run, representing a reduction in computation of around 3 orders of magnitude. It achieves this by using a k-means clustering algorithm (MacQueen, 1967) to perform a multidimensional classification on the model domain resulting in so-called TopoSUB clusters, consisting of pixels with similar terrain parameters. Predictors used in the clustering algorithm are elevation, slope, aspect and sky view factor. In this way it is possible to produce high resolution maps (e.g. DEM resolution) over large (regional) modelling domains, while explicitly including important drivers of surface-atmosphere processes. The scheme is implemented on an HPC cluster for efficiency in an "embarrassingly parallel" sense, i.e. no communication required between compute nodes. Typical setups use 100 nodes with run-times measured in minutes for full 1980-2100 runs, for example to produce the results given in Figure 7. The scheme also has a desktop mode which typically utilises 8 cores and typical runs will require a few hours. These indicative numbers are provided merely to give the reader an order-of-magnitude idea of how frugal this scheme is in terms of computation resources compared to dynamical downscaling.

4.2 Treatment of glaciers

Glacier zones are typically masked in studies of seasonal snow unless a glacier layer is explicitly accounted for in the model. In this study, however, we wanted to highlight and track changes in the glacier accumulation zones. These zones are areas where the annual surface mass balance is positive, leading to the formation or growth of glaciers. With our framework we are able to identify such zones, as we are able to model areas with perennial snow cover with FSM and these give a good indication of glacier accumulation zones under a given climate. However, we ignore glacier dynamics so we are not able to adequately map the spatio-temporal evolution of glaciers.

4.3 Evaluation

Predictive methods must by definition be evaluated on independent data from that which was used for calibration in order to correctly evaluate how applicable a model is beyond the data-space within which it was developed. Further, highly adaptable methods, such as the non-parametric techniques used in this study, are prone to over fitting. These issues are avoided in this study as we perform an independent evaluation using station data from the IMIS station network, whereas calibration or in this case downscaling, is performed using downscaled ERA5 fields. Note, none of the meteorological fields of the IMIS stations were assimilated during the production of ERA5. Snow depth results are evaluated by automatic snow depth measurements performed by sonic ranger (Campbell Scientific SR50), available from the Inter-cantonal Measurement and Information System (IMIS) station network at 30 minute intervals. This is a high elevation station network that forms the backbone of the national avalanche service in Switzerland.

The analysis in this study is organised as follows: The reference period is defined as 1981-2010 and future scenarios are analysed for climate periods 2031-2060 and 2070-2100. We assess the downscaling of all meteorological fields at the Weissfluhjoch IMIS station (Figure 3) which has the full suite of variables produced as compared to standard IMIS stations which lack a full radiation balance. Here, we additionally test the performance of the scheme at point-scale (Figure 5 and 6). We further assess air temperature and snow depth (as a proxy for precipitation) using the entire IMIS network (Figure 6). We do not assess precipitation directly as year-round datasets are not available from the IMIS network due to unheated gauges, and snow depth is often considered a more robust variable to measure at high elevation. After this we provide results using the scheme at various spatial scales (Figure 6-8).

5 Results and discussion

5.1 Evaluation at the Weissfluhjoch station

Figure 3 and Table 3 and 4 (statistics) show an evaluation of the scheme at the WFJ station both as a cumulative distribution function (left column) and a day of year (DOY) plot which averages all values in the timeseries for a given DOY (right column). Here we compare grid-box CORDEX ensemble mean (CLIM), a single parameter set quantile map run (QM) and monthly parameter set quantile map run (QM_MONTH) to T-MET and station measurements. This comparison is done over

the period 1996-2006 which corresponds to the overlapping time-frames of CORDEX-HIST, T-MET and station data, and also importantly does not include the period over which quantile mapping is run (1980-1995). We perform two sets of comparison, first with T-MET as a reference as this is the target of the quantile mapping (Table 3). This shows how the scheme behaves particularly in terms of the different QM and QM_MONTH implementations (Table 3). In the second, we compare results directly to the measurements at the station to get a global look at how the scheme performs (Table 4), of course this is then subject to residual error in the TopoSCALE downscaling scheme which produces T-MET (Table 5) and therefore must be interpreted carefully.

With respect to target T-MET (Figure 3, Table 3) percentage bias and RMSE are generally strongly decreased by the scheme in the standard QM mode and further by the QM_MONTH variant, especially where there is a time varying bias signal (e.g. shortwave radiation).

With respect to station measurements (Table 4) we see overall improvement in statistics by the scheme, however this is not always reflected in an improvement in statistics between QM and QM_MONTH as there is residual error between the station measurements and downscaled T-MET. The strongest improvements are variables which are downscaled according to model pressure levels in the TopoSCALE scheme (air temperature and relative humidity). Precipitation is the only parameter we do not see an improvement in the scheme with respect to the measurements, but this is expected due to high uncertainty in precipitation and the fact this is not addressed by the base TopoSCALE downscaling. This point is further discussed below in Section 5.4.

Figure 4 shows a typical point scale application generating a forcing timeseries, in this case for air temperature. The effect of the bias correction is shown with a clear improvement with respect to T-MET. Figure 5 gives a point scale example of snow depth evolution at WFJ. Available snow depth observations from WFJ show good agreement with both the historical period and first decade of the RCP runs in terms of lying within the model ensemble. It should be stressed that a quantitative comparison with individual years cannot be made as CORDEX variability (and climate models more generally) is not expected to be perfectly synchronised with observed variability at temporal resolutions of years to a decade. Note the stable/rising snow depth under RCP2.6 by end of century correlated to stabilised mid-century temperatures shown in Figure 4.

5.2 Evaluation for the IMIS network

Figure 6 gives the evolution of snow depth averaged across the IMIS network in terms of mean DOWY (day of water year) snow depth, starting 1 September. IMIS station measurements are given as reference, however it should be noted that the time period covered by each station is variable within the period 1996-2018. The shortest station record is 10 years. These are compared to T-MET snow depth by only using days present in the IMIS dataset. At this synoptic spatial and temporal scale there is good agreement with low bias and RMSE scores and high correlation (Table 4). The evolution of snow depth for future climate scenarios and two future periods is given. By mid-century under RCP2.6 a reduction in peak snow depth of around 20 cm is seen and further 25 cm under RCP8.5 with respect to the reference CORDEX historical period (CORDEX-HIST). Peak snow depth occurs around 30-40 days earlier. By late century RCP2.6 snow depth has not further deteriorated and in fact shown signs of possible recovery with peak snow depth moving towards that of the CORDEX-HIST period. RCP8.5, however, gives

a strong further reduction in peak snow depth of around 1 m with respect to CORDEX-HIST. The comparison between the CORDEX-HIST (1980-2006) and T-MET/ IMIS (1996-2018) is useful but should be treated with caution due to only partially overlapping periods (defined by availability of station measurements). The CORDEX-HIST period has a higher peak snow depth and later snow melt-out date, which could be due to less climate change as it represents an earlier period than IMIS and T-MET.

5.3 Climate change impacts on Alpine snow cover across Switzerland

As an example application of the full model pipeline, the results in Figure 7 were generated by feeding model results (TopoCLIM/FSM) to the TopoSUB spatial framework to generate transiently modelled snow depth maps at 100 m resolution. The ensemble mean is used in each scenario/ time period plot. We highlight again that by using this simple approach we implicitly model climate-viable glacier accumulation zones, where the snowpack does not melt-out by the end of summer. Mid-century results are comparable between RCP2.6 and RCP8.5 except for lower elevation areas such as the Jura mountains in the far west of Switzerland. By late-century the difference is large, as also seen in Figure 6 with a strong increase in the snowline elevation and reduction of snow depth even at high elevations. Glacier accumulation zones remain only in the high Valais (Mattertal) and Bernese Oberland around today's accumulation zone of the Aletsch Glacier. Figure 8 provides a more quantitative picture of snow cover-elevation dynamics. This hypsometry plot summarizes the pixels of Figure 7 by showing the mean and standard deviation of snow depth at each 50 m elevation band across the entire domain for three time periods and scenarios RCP2.6/8.5. As expected, a strong decrease in snow depth is seen at all elevations under RCP8.5. A slightly different story is seen under RCP2.6 with snow depth reducing up to mid-century followed by an increase at high elevations (above 3500 masl) by end of the century - this is a consistent message throughout this study and reflects the results of Figures 5, 6 and 7. It can be partially explained by stabilisation of air temperatures by mid-century (Figure 4) together with increased precipitation in the Alpine region (Jacob et al., 2014; Smiatek et al., 2016).

An interesting observation in this figure is that in all time periods and scenarios snow depth is limited both at low elevation by temperature and at high elevation by terrain, which tends to be steeper and therefore permits lower accumulations due to avalanching (permitted by the model, as discussed in Section 3.4). A final point of note in this figure is the truncation level indicating the accumulation zone elevation which is approximately 4000 m for HIST. Above this level seasonal snow is not possible, as it will not melt before the following winter season or the ground is too steep for snow accumulation. This upper limit of seasonal snow rises to around 4500masl over the 21st century, well above former glaciated surfaces. An implication of this for water resources is that while we lose a large quantity of glacier accumulation zones during the 21st century, irrespective of scenario, we will likely not lose seasonal snow water resources at those elevations.

Several previous studies have investigated the impacts of climate change upon Alpine snow cover (e.g. Verfaillie et al., 2018; Steger et al., 2012; Marty et al., 2017; Frei et al., 2018; Bender et al., 2020), however direct comparison is often problematic due to model resolution, analysis period, parent climate models and/or emissions scenarios used. This highlights the importance of model intercomparison studies whereby these important variables controlling model results can be standardised. Comparison to the previous works cited highlights the contribution of this study in that most were conducted either at an RCM resolution

of 25 - 12 km (e.g. Steger et al., 2012; Frei et al., 2018), local scale (e.g. Verfaillie et al., 2018; Bender et al., 2020) or reliant on in situ data (e.g. Marty et al., 2017; Bender et al., 2020). In this study, we demonstrate a method that generates results over large modelling domains at hillslope scale, a scale which is extremely important in regulating the stores and fluxes of water, energy, and carbon (e.g Fan et al., 2019), and therefore critical to modelling snow cover in mountainous terrain. Additionally, this approach does not rely on in situ data and therefore is appropriate for data-scarce regions.

5.4 Forcing uncertainty

The largest source of uncertainty in the scheme is the reanalysis forcing from ERA5. Both quantile mapping and disaggregation of fields to sub-daily timesteps are inherently constrained by the distribution of ERA5 fields. While ERA5 provides hourly data and therefore resolves the diurnal cycle, it remains a 0.25 degree model with a correspondingly smooth topographical surface representation and parameterisation of physical processes that occur on shorter length scales. Typical examples are convective precipitation and cold air pooling in valley bottoms (Cao et al., 2017; Liu et al., 2017), orographic enhancement of precipitation and wind fields (Gerber et al., 2018; Mott et al., 2018; Gutmann et al., 2016). As the density of observations that are assimilated in reanalyses varies globally we expect the performance of the TopoCLIM model pipeline to be a function of how well constrained ERA5 is in any given location. A full analysis and discussion of TopoSCALE uncertainties is given by Fiddes and Gruber (2014) and to some extent in Fiddes et al. (2019). The most uncertain variable is precipitation which is clearly a critical point for snow modelling studies. We do however show that there are no large scale biases in the precipitation field at least in our snow depth comparison across the IMIS network. We have shown in previous studies that variables driving the energy balance (air temperature, incoming shortwave radiation, incoming longwave radiation) are downscaled with good skill by TopoSCALE. One method we have explored to reduce (and quantify) meteorological forcing uncertainty and precipitation uncertainty in particular, is through Bayesian data assimilation of globally available satellite products using a particle batch smoother, which has shown promising results (Fiddes et al., 2019; Alonso-González et al., 2020). The coupling of this scheme with TopoCLIM will be the subject of subsequent work.

5.5 Evaluation uncertainty

Our station measurements are characterised by considerable uncertainties, a problem that is particularly acute for precipitation related fields such as snow depth, so we have to some degree a chicken and egg scenario when it comes to model validation, in that it's hard to untangle the origin of apparent errors. For example, stations tend to be situated in sheltered flat to concave topography. Here we expect there to be considerable preferential deposition (Lehning et al., 2008) from surrounding windblown slopes and ridge-lines (Grünewald and Lehning, 2015). In general we would expect stations therefore to be positively biased with respect to large scale precipitation fluxes. Additionally, certain stations will be exposed to very local climatic effects which are not represented in the large scale 0.25 degree resolution ERA5 forcing. Examples of such local effects include wind funnelling leading to scouring, enhanced Foehn effects and local orographic enhancement.

5.6 Snow model uncertainty

The snow model used in this study, FSM, is an intermediate complexity physically-based model and we do not expect it to perform as well with respect to snow densification as a more complex snow physics model such as SNOWPACK (Lehning et al., 2002; Wever et al., 2015). This can introduce uncertainty when using snow depth as a validation parameter. However, it should be noted that there is active discussion about snow model complexity and how this does not necessarily lead to improved performance (c.f. Magnusson et al., 2015). Snow water equivalent is a simpler modelling objective but a much harder measurement objective and therefore few sites are available - particularly with good coverage at regional scales, therefore limiting its applicability for large-scale evaluations.

6 Conclusions

In this study we have developed and tested a new scheme for downscaling regional climate projections, specifically designed to provide hillslope scale forcings for impact models. We take advantage of the now globally available CORDEX RCM data to develop a method with global scope. The scheme adheres to the approach of TopoSCALE upon which it builds, that is modelling tools that bridge the gap between relatively simple empirical approaches and full dynamical models that require extensive computing resources. It can be run both on desktop or cluster environments. The target application of this scheme is impact modelling in complex terrain where significant atmosphere-surface interactions need to be considered. A particular application is in remote areas where ground data may either not be present at all or not available for the duration of a climate normal period, meaning that traditional ESD methods are problematic to use. Another strength of this approach is that it produces continuous timeseries, such that it permits transient simulations in contrast to other parsimonious methods such as the delta-change approach, an important point for domains such as soil, ground-ice or glaciers where surface forcings drive processes over decadal timescales.

This framework is adaptable to any kind of meteorological input data (both the reference data and the future/past period). Here we have given an example with the reference based on downscaling ERA5, but it could equally be generated by downscaling other reanalyses such as MERRA or outputs from regional models such as WRF, COSMO, or ICAR. The method would also be applicable to other future projections such as those from CMIP6 which are coming online now. Exploring all these possibilities is beyond the scope of this paper, but it's useful to emphasize the potential applications of this framework. Another aspect to highlight is that we are downscaling an ensemble of CORDEX outputs which also gives us a better idea of the uncertainty in future climate projections and impacts on the snowpack (in this case). It should also be stated that as a framework other downscaling routines or bias correction approaches could be used. The modularity allows these options to be explored. The very efficient combination of steps as presented in this paper provides a powerful way of rapidly obtaining hillslope scale climate forcings anywhere on the globe.

As a final point, TopoCLIM can be used to bias correct any dataset that partly overlaps with the reference period. Thus, in addition to the future projections considered in this study, it would also be possible to use TopoCLIM to correct coarser-scale reanalysis data that stretch far back in time. A prime example would be ECMWF's 20th century reanalysis (ERA-20C) which

spans 1900-2010 and thus partly overlaps with ERA5 (1950-today) that is used to drive TopoSCALE to generate the reference data T-MET.

Code and data availability. Model code, documentation and example used in this study is archived at 10.16904/envidat.229 and Github repository is available at: https://github.com/joelfiddes/topoCLIM

Author contributions. JF devised the study conducted the analysis and wrote the manuscript, KA assisted with algorithm development and text, ML helped devise the study, assisted with analysis and text.

Competing interests. We declare no competing interests.

Acknowledgements. This study was conducted within the Swiss National Science Foundation project "Precipitation in extreme environments" (project number: 179130) and the WSL programme Climate Change impacts on Alpine Mass Movements (CCAMM). KA was funded by the European Space Agency Permafrost CCI project, and acknowledges support from the LATICE strategic research area at the University of Oslo.

We acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5. We also thank the climate modelling groups (listed in Table 2 of this paper) for producing and making available their model output. We also acknowledge the Earth System Grid Federation infrastructure an international effort led by the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison, the European Network for Earth System Modelling and other partners in the Global Organisation for Earth System Science Portals (GO-ESSP).

References

Aalstad, K., Westermann, S., Schuler, T. V., Boike, J., and Bertino, L.: Ensemble-based assimilation of fractional snow-covered area satellite retrievals to estimate the snow distribution at Arctic sites, The Cryosphere, 12, 247–270, 2018.

Addor, N. and Seibert, J.: Bias correction for hydrological impact studies – beyond the daily perspective, Hydrol. Process., 28, 4823–4828, https://doi.org/10.1002/hyp.10238, 2014.

Alonso-González, E., Gutmann, E., Aalstad, K., Fayad, A., and Gascoin, S.: Snowpack dynamics in the Lebanese mountains from quasi-dynamically downscaled ERA5 reanalysis updated by assimilating remotely-sensed fractional snow-covered area, Hydrology and Earth System Sciences Discussions, 2020, 1–31, https://doi.org/10.5194/hess-2020-335, https://hess.copernicus.org/preprints/hess-2020-335/, 2020.

Arguez, A. and Vose, R. S.: The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals, Bull. Am. Meteorol. Soc., 92, 699–704, https://doi.org/10.1175/2010BAMS2955.1, 2011.

Bender, E., Lehning, M., and Fiddes, J.: Changes in Climatology, Snow Cover, and Ground Temperatures at High Alpine Locations, Front Earth Sci., 8, 100, https://doi.org/10.3389/feart.2020.00100, 2020.

Cao, B., Gruber, S., and Zhang, T.: REDCAPP (v1.0): parameterizing valley inversions in air temperature data downscaled from reanalyses, Geoscientific Model Development, 10, 2905–2923, https://doi.org/10.5194/gmd-10-2905-2017, 2017.

Cao, B., Quan, X., Brown, N., Stewart-Jones, E., and Gruber, S.: GlobSim (v1. 0): deriving meteorological time series for point locations from multiple global reanalyses, Geoscientific Model Development, 12, 4661–4679, https://doi.org/10.5194/gmd-12-4661-2019, 2019.

Dee, D. P., Balmaseda, M., Balsamo, G., Engelen, R., Simmons, A. J., and Thépaut, J.-N.: Toward a Consistent Reanalysis of the Climate System, Bull. Am. Meteorol. Soc., 95, 1235–1248, https://doi.org/10.1175/BAMS-D-13-00043.1, 2014.

Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Glob. Planet. Change, 57, 16–26, https://doi.org/10.1016/j.gloplacha.2006.11.030, 2007.

Dozier, J. and Frew, J.: Rapid Calculation of Terrain Parameters For Radiation Modeling From Digital Elevation Data, Transactions on Geosciences and Remote Sensing, 28, 963–969, 1990.

Essery, R.: A factorial snowpack model (FSM 1.0), Geosci. Model Dev., 8, 3867–3876, https://doi.org/10.5194/gmd-8-3867-2015, 2015.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope hydrology in global change research and earth system modeling, Water Resour. Res., 55, 1737–1772, https://doi.org/10.1029/2018wr023903, 2019.

Fiddes, J. and Gruber, S.: TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales, Geoscientific Model Development, 5, 1245–1257, 2012.

Fiddes, J. and Gruber, S.: TopoSCALE v.1.0: downscaling gridded climate data in complex terrain, Geoscientific Model Development, 7, 387–405, 2014.

Fiddes, J., Endrizzi, S., and Gruber, S.: Large-area land surface simulations in heterogeneous terrain driven by global data sets: application to mountain permafrost, The Cryosphere, 9, 411–426, 2015.

Fiddes, J., Aalstad, K., and Westermann, S.: Hyper-resolution ensemble-based snow reanalysis in mountain regions using clustering, Hydrol. Earth Syst. Sci., 23, 4717–4736, 2019.

Förster, K., Hanzer, F., Winter, B., Marke, T., and Strasser, U.: An open-source MEteoroLOgical observation time series DISaggregation Tool (MELODIST v0.1.1), Geosci. Model Dev., 9, 2315–2333, https://doi.org/10.5194/gmd-9-2315-2016, 2016.

Frei, P., Kotlarski, S., Liniger, M. A., and Schär, C.: Future snowfall in the Alps: projections based on the EURO-CORDEX regional climate models, cryosphere, 12, 1–24, https://doi.org/10.5194/tc-12-1-2018, 2018.

Gerber, F., Besic, N., Sharma, V., Mott, R., Daniels, M., Gabella, M., Berne, A., Germann, U., and Lehning, M.: Spatial variability in snow precipitation and accumulation in COSMO–WRF simulations and radar estimations over complex terrain, cryosphere, 12, 3137–3160, https://doi.org/10.5194/tc-12-3137-2018, 2018.

Gobiet, A., Suklitsch, M., and Heinrich, G.: The effect of empirical-statistical correction of intensity-dependent model errors on the temperature climate change signal, Hydrol. Earth Syst. Sci., 19, 4055, https://doi.org/10.5194/hess-19-4055-2015, 2015.

Grünewald, T. and Lehning, M.: Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale: REPRESENTATIVENESS OF FLAT-FIELD SNOW DEPTH MEASUREMENTS, Hydrol. Process., 29, 1717–1728, https://doi.org/10.1002/hyp.10295, 2015.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012, 2012.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M.: An intercomparison of statistical downscaling methods used for water resource assessments in the United States, Water Resour. Res., 50, 7167–7186, https://doi.org/10.1002/2014wr015559, 2014.

Gutmann, E., Barstad, I., Clark, M., Arnold, J., and Rasmussen, R.: The Intermediate Complexity Atmospheric Research Model (ICAR), J. Hydrometeorol., 17, 957–973, https://doi.org/10.1175/JHM-D-15-0155.1, 2016.

Hanzer, F., Förster, K., Nemec, J., and Strasser, U.: Projected cryospheric and hydrological impacts of 21st century climate change in the Ötztal Alps (Austria) simulated using a physically based approach, Hydrology and Earth System Sciences 22 (2018), Nr. 2, 22, 1593–1614, https://doi.org/10.5194/hess-22-1593-2018, 2018.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, Earth Syst. Dyn., 4, 219–236, https://doi.org/10.5194/esd-4-219-2013, 2013.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quart. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hock, R., Rasul, G., Adler, C., Caceres, B., Gruber, S., Hirabayashi, Y., Jackson, M., Kaab, A., Kang, S., Kutuzov, S., Milner, A., Molau, U., Morin, S., Orlove, B., and Steltzer, H.: High Mountain Areas, pp. 131–202, 2019.

Ivanov, M. A. and Kotlarski, S.: Assessing distribution-based climate model bias correction methods over an alpine domain: added value and limitations, Int. J. Climatol., 37, 2633–2653, 2017.

Ivanov, M. A., Luterbacher, J., and Kotlarski, S.: Climate Model Biases and Modification of the Climate Change Signal by Intensity-Dependent Bias Correction, J. Clim., 31, 6591–6610, https://doi.org/10.1175/JCLI-D-17-0765.1, 2018.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N.,

Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional Environ. Change, 14, 563–578, https://doi.org/10.1007/s10113-013-0499-2, 2014.

Katz, R. W. and Brown, B. G.: Extreme events in a changing climate: Variability is more important than averages, Clim. Change, 21, 289–302, https://doi.org/10.1007/BF00139728, 1992.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Wulfmeyer, V., and others: Regional climate modeling on European scales : A joint standard evaluation of the EURO-CORDEX RCM ensemble, 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014, 2014.

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 Reg. Sci. Technol., 35, 169–184, https://doi.org/10.1016/S0165-232X(02)00072-1, 2002.

Lehning, M., Löwe, H., Ryser, M., and Raderschall, N.: Inhomogeneous precipitation distribution and snow transport in steep terrain: SNOW DRIFT AND INHOMOGENEOUS PRECIPITATION, Water Resour. Res., 44, 1–19, https://doi.org/10.1029/2007WR006545, 2008.

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., Chen, F., Chen, L., Clark, M., Dai, A., Dudhia, J., Eidhammer, T., Gochis, D., Gutmann, E., Kurkute, S., Li, Y., Thompson, G., and Yates, D.: Continental-scale convection-permitting modeling of the current and future climate of North America, Clim. Dyn., 49, 71–95, https://doi.org/10.1007/s00382-016-3327-9, 2017.

MacQueen, J.: Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, vol. 5.1, pp. 281–298, University of California Press, Berkeley, CA, 1967.

Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A., and Jonas, T.: Evaluating snow models with varying process representations for hydrological applications, Water Resour. Res., 51, 2707–2723, https://doi.org/10.1002/2014WR016498, 2015.

Maraun, D.: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, J. Clim., 26, 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1, 2013.

Marty, C., Schlögl, S., Bavay, M., and Lehning, M.: How much can we save? Impact of different emission scenarios on future snow cover in the Alps, cryosphere, 11, 517–529, https://doi.org/10.5194/tc-11-517-2017, 2017.

Michel, A., Sharma, V., Lehning, M., and Huwald, H.: Climate change scenarios at hourly time-step over Switzerland from an enhanced temporal downscaling approach, Int. J. Climatol., https://doi.org/10.1002/joc.7032, 2021.

Mott, R., Vionnet, V., and Grünewald, T.: The Seasonal Snow Cover Dynamics: Review on Wind-Driven Coupling Processes, Front Earth Sci. Chin., 6, 197, https://doi.org/10.3389/feart.2018.00197, 2018.

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M. O., Lewkowicz, A. G., Panda, S. K., Romanovsky, V., Way, R. G., Westergaard-Nielsen, A., Wu, T., Yamkhin, J., and Zou, D.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km^2 scale, Earth-Science Reviews, 193, 299–316, https://doi.org/https://doi.org/10.1016/j.earscirev.2019.04.023, 2019.

Obu, J., Westermann, S., Vieira, G., Abramov, A., Balks, M. R., Bartsch, A., Hrbáček, F., Kääb, A., and Ramos, M.: Pan-Antarctic map of near-surface permafrost temperatures at 1 km^2 scale, The Cryosphere, 14, 497–519, https://doi.org/10.5194/tc-14-497-2020, 2020.

Panofsky, H. A. and Brier, G. W.: Some Applications of Statistics to Meteorology, The Pennsylvania State University Press, 1968.

Parker, W. S.: Reanalyses and Observations: What's the Difference?, Bull. Am. Meteorol. Soc., 97, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1, 2016.

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

Schuler, T. V. and Østby, T. I.: Sval_Imp: a gridded forcing dataset for climate change impact research on Svalbard, Earth System Science Data, 12, 875–885, https://doi.org/10.5194/essd-12-875-2020, https://essd.copernicus.org/articles/12/875/2020/, 2020.

Smiatek, G., Kunstmann, H., and Senatore, A.: EURO-CORDEX regional climate model analysis for the Greater Alpine Region: Performance and expected future change: CLIMATE CHANGE IN THE GAR AREA, J. Geophys. Res. D: Atmos., 121, 7710–7728, https://doi.org/10.1002/2015JD024727, 2016.

Steger, C., Kotlarski, S., Jonas, T., and Schär, C.: Alpine snow cover in a changing climate: a regional climate model perspective, Clim. Dyn., 41, 735–754, https://doi.org/10.1007/s00382-012-1545-3, 2012.

Stocker, T., Qin, D., Plattner, G.-K., and Tignor, M.: Workshop Report of the Intergovernmental Panel on Climate Change Workshop on Regional Climate Projections and their Use in Impacts and Risk Analysis Studies, 2015.

Teutschbein, C. and Seibert, J.: Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?, Hydrol. Earth Syst. Sci., 17, 5061–5077, https://doi.org/10.5194/hess-17-5061-2013, 2013.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, Int. J. Climatol., 31, 1530–1544, https://doi.org/10.1002/joc.2168, 2011.

Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Clim. Change, 112, 449–468, https://doi.org/10.1007/s10584-011-0224-4, 2012.

Verfaillie, D., Lafaysse, M., Déqué, M., Eckert, N., Lejeune, Y., and Morin, S.: Multi-component ensembles of future meteorological and natural snow conditions for 1500 m altitude in the Chartreuse mountain range, Northern French Alps, cryosphere, 12, 1249–1271, https://doi.org/10.5194/tc-12-1249-2018, 2018.

Wang, X., Tolksdorf, V., Otto, M., and Scherer, D.: WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis, Int. J. Climatol., 41, 743–762, https://doi.org/10.1002/joc.6686, 2021.

Westermann, S., Østby, T. I., Gisnås, K., Schuler, T. V., and Etzelmüller, B.: A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data, The Cryosphere, 9, 1303–1319, https://doi.org/10.5194/tc-9-1303-2015, https://tc.copernicus.org/articles/9/1303/2015/, 2015.

Wever, N., Schmid, L., Heilig, A., Eisen, O., Fierz, C., and Lehning, M.: Verification of the multi-layer SNOWPACK model with different water transport schemes, The Cryosphere Discussions, 9, 2271–2293, https://doi.org/10.5194/tcd-9-2655-2015, 2015.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, Clim. Change, 62, 189–216, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e, 2004.

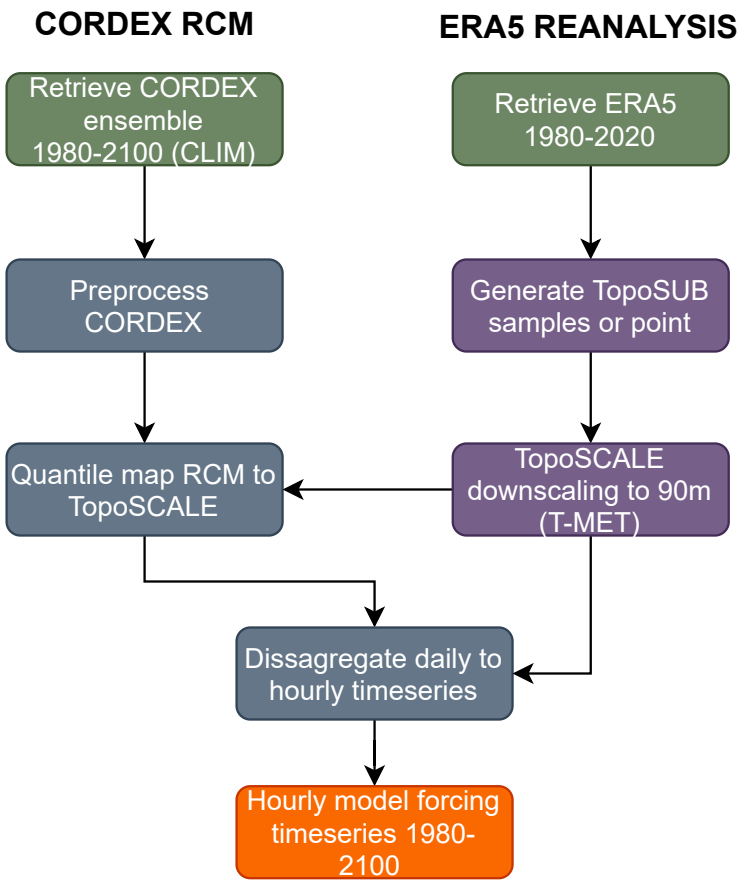

Figure 1. Schematic overview of the main TopoCLIM processes and experimental setup used in this study.

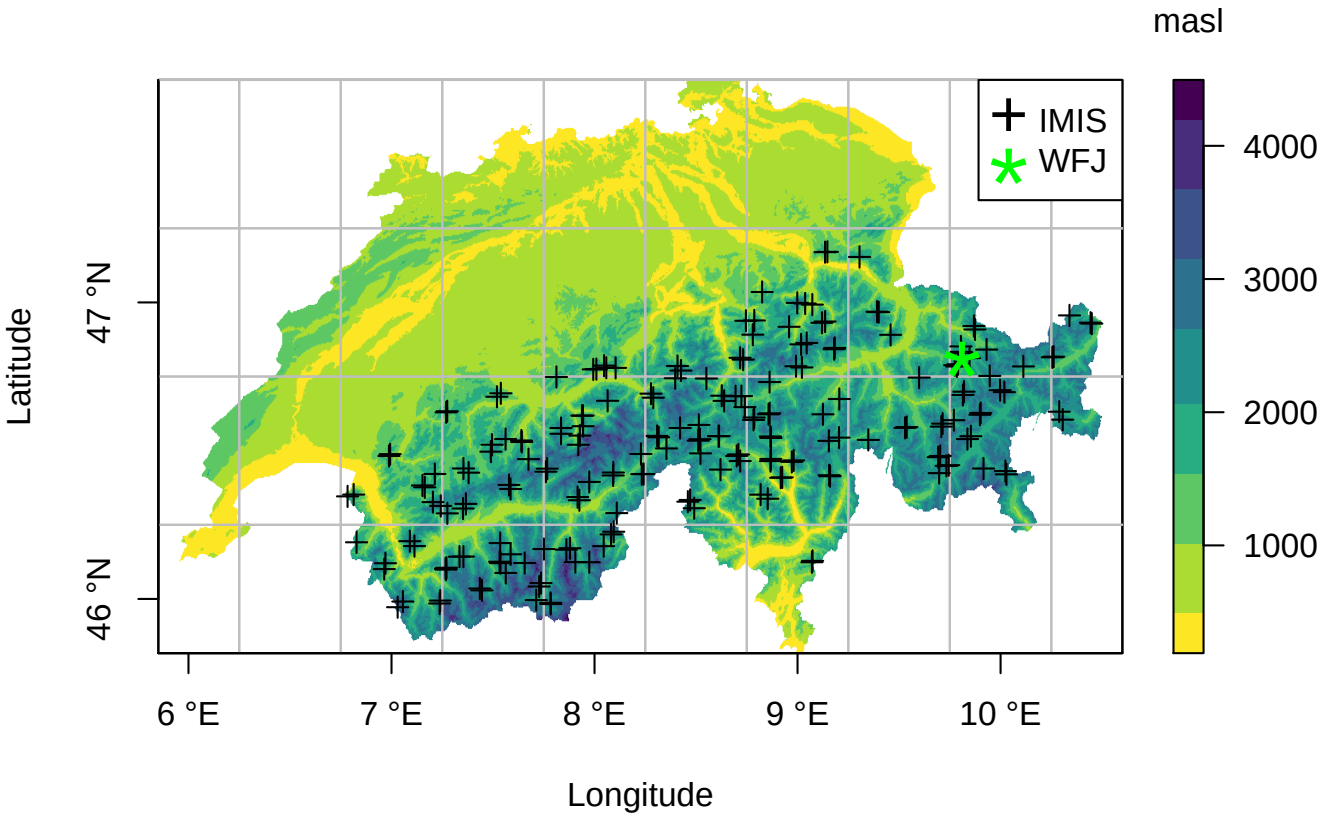

Figure 2. Study domain map with CORDEX EUR-44 grid (44 km) overlaid. IMIS stations and the Weissfluhjoch site (WFJ) are indicated.

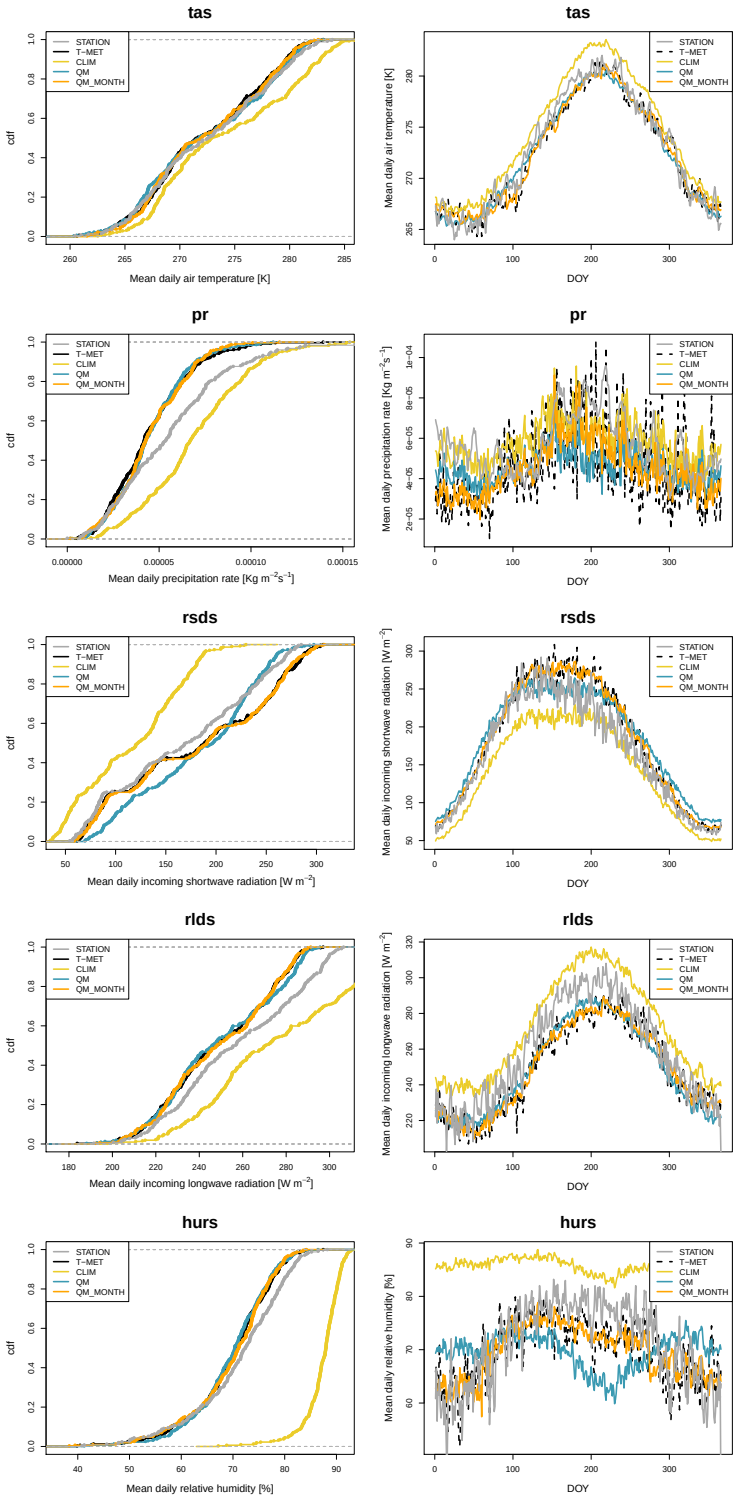

Figure 3. Evaluation of the quantile mapping routine at the Weissfluhjoch station in standard mode "QM" single parameter set and "QM_MONTH" seasonally varying parameter set. CLIM is the uncorrected CORDEX data. STATION is the Weissfluhjoch station measurements. T-MET is the downscaled ERA5 data obtained using TopoSCALE. Shown are the cumulative density function over the period 1981-2010 (left panel) and seasonal distribution given as average by day of year (DOY) for the same period (right panel).

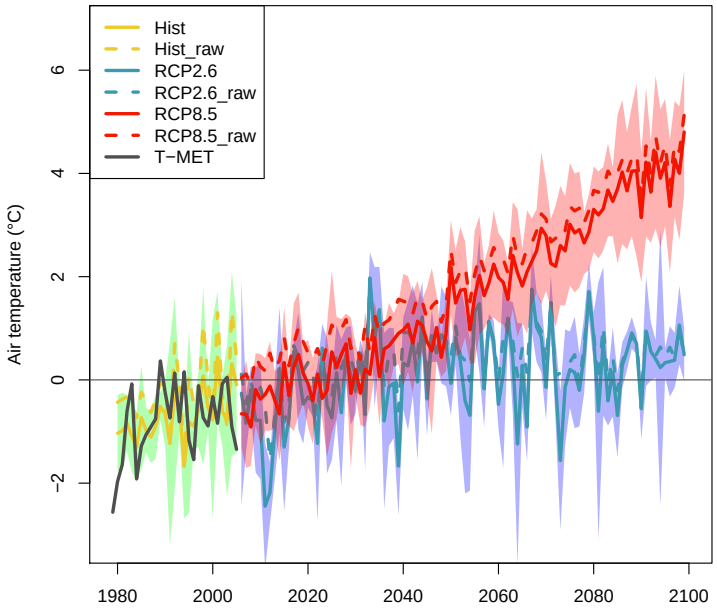

Figure 4. A point-scale TopoCLIM product: Mean annual near surface air temperature at the Weissfluhjoch (2540 masl) showing corrected historical, RCP2.6 and RCP8.5 time series. T-MET and uncorrected CORDEX data are also shown for comparison. The coloured envelopes indicate the model spread and ensemble mean is given by the bold line. The zero degree isotherm is given by the horizontal line.

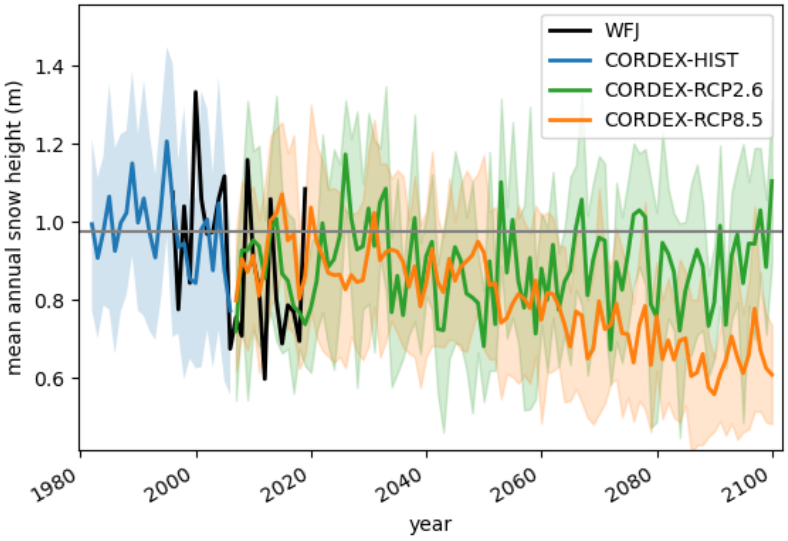

Figure 5. An example point-scale TopoCLIM product: Mean annual snow depth (m) ensemble at the Weissfluhjoch station (2540m asl) for historical and future RCP scenarios. Observations from Weissfluhjoch are given in black for a qualitative comparison. Ensemble means are indicated by solid lines and long term average over the plotted historical period (1981-2010) is given by the horizontal line. Observations lie within the ensemble spread, indicating that the bias correction method works satisfactorily. It should be stressed that we do not attempt a quantitative comparison here as CORDEX (and climate models more generally) variability is not expected to be perfectly synchronised with observed variability. Nonetheless, this representation usefully shows the inter-annual variability that is present in both the model and the observations instead of simply showing decadal means.

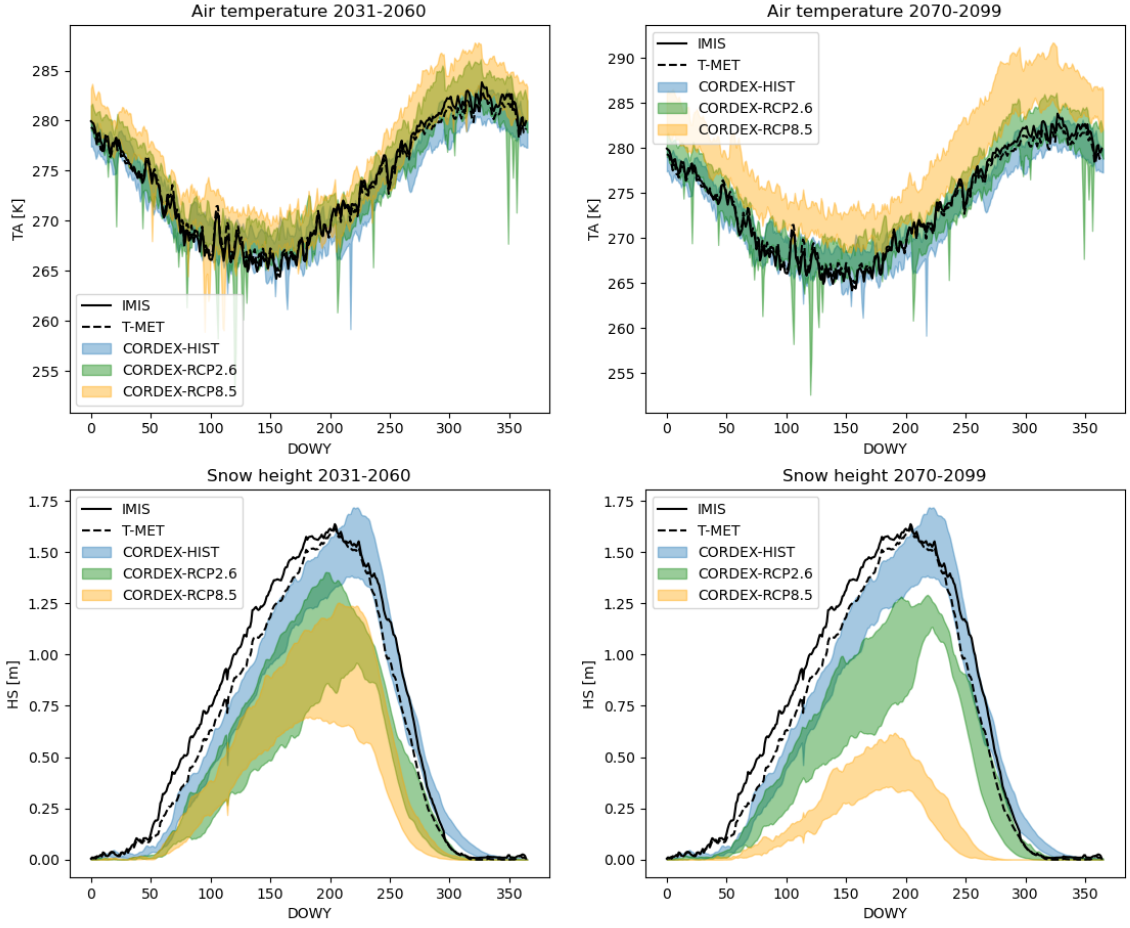

Figure 6. Large scale evaluation of the scheme at all IMIS station locations: Network mean air temperature and snow depth across all IMIS stations (IMIS) is compared to the same data generated by the TopoSCALE downscaling of ERA5 (T-MET) and TopoCLIM downscaled CORDEX data for the historical period (1981-2010) and scenarios RCP2.6/8.5 for near (2031-2060) and far future periods (2070-2099) for each day of water year (DOWY), starting at September 1. The full width of the model ensemble is represented by the shading. Note T-MET and IMIS have the same time-frame (variable by station between 1996-2020) and are therefore directly comparable (see Table 5), but this only partially overlaps with the CORDEX historical period and therefore is intended for visual comparison only.

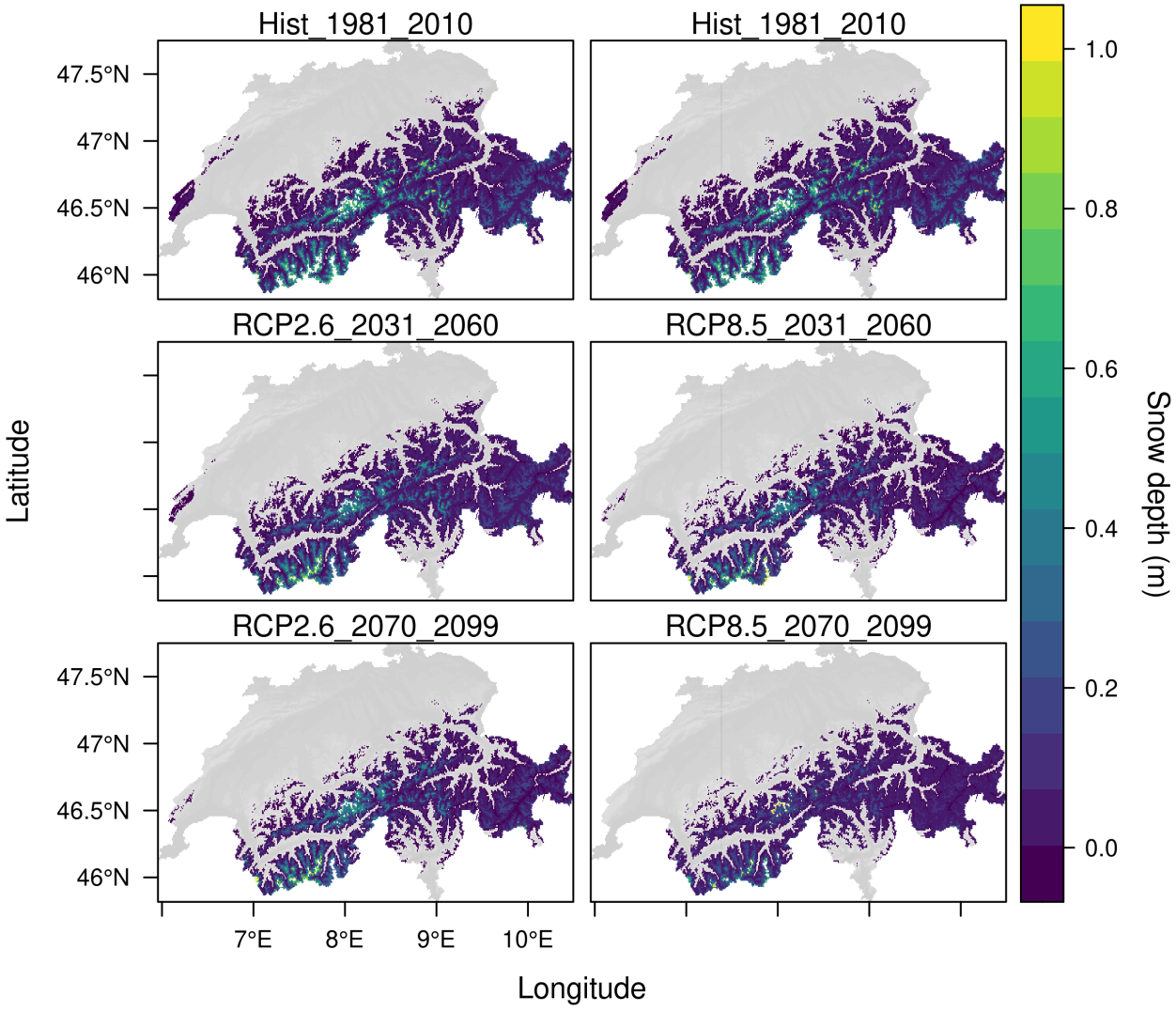

Figure 7. Example TopoCLIM forced climate change maps: Mean snow depth (m) for RCP2.6 and RCP8.5 (columns) and time periods 1981-2010, 2031-2060 and 2070-2099 (rows). Perennial snow-pack (does not melt in an annual cycle) is masked out (white regions) and assumed to represent climatological glacier accumulation zones.

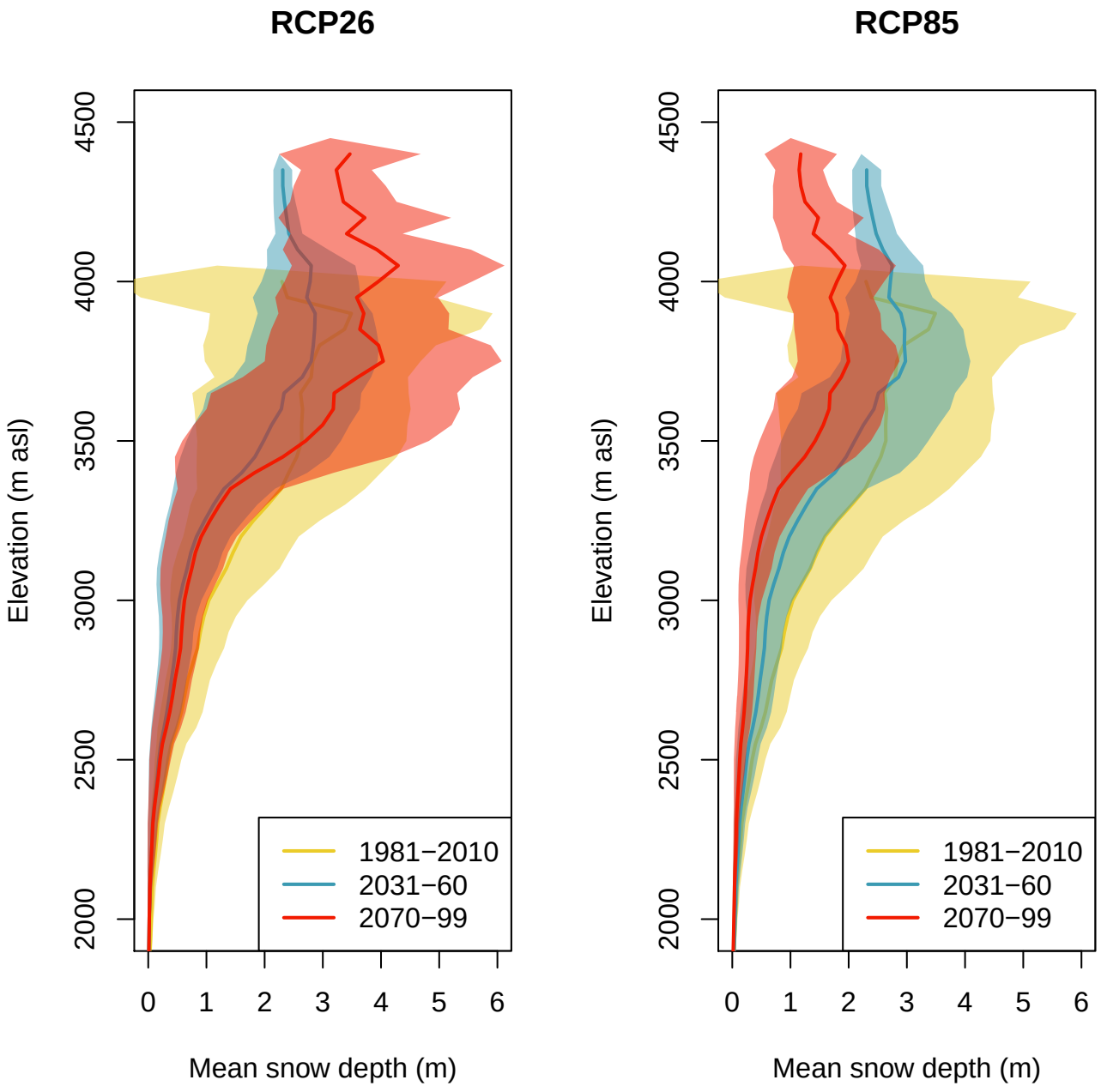

Figure 8. Hypsometry of snow depth over the Swiss Alps for RCP2.6, RCP8.5 for periods 1981-2010, 2031-2060 and 2070-2099. This data is derived from the ensemble mean, with shading showing +/- 1 standard deviation of values in each elevation band. The elevation limit in each time period corresponds to the threshold between seasonal and non-seasonal snow, or glacier accumulation zones. This demonstrates the rising equilibrium-line altitude of glaciers over the 21st century in all RCP's and extension of seasonal snow into the former glacier zones.

Table 1. CORDEX variables used in this study, together with the Climate and Forecast Conventions (CF) standard name.

Variable name	Units	Timestep	CF long name	Melodist method
tas	K	daily	Near-Surface Air Temperature	sine mean
pr	kg m^{-2} s^{-1}	daily	Precipitation	equal
ps	Pa	daily	Surface Air Pressure	-
hurs	%	daily	Near-Surface Relative Humidity	equal
rsds	W m^{-2}	daily	Surface Downwelling Shortwave Radiation	pot rad
rlds	W m^{-2}	daily	Surface Downwelling Longwave Radiation	-
uas	m s^{-1}	daily	Eastward Near-Surface Wind	random
vas	m s^{-1}	daily	Northward Near-Surface Wind	random

Table 2. CORDEX model chains used in this study.

GCM	RCM	Scenario	Ensemble	Version
CNRM-CERFACS-CNRM-CM5	CLMcom-CCLM5-0-6	Hist/RCP2.6/8.5	r1i1p1	v1
CNRM-CERFACS-CNRM-CM5	SMHI-RCA4	Hist/RCP2.6/8.5	r1i1p1	v1
ICHEC-EC-EARTH	CLMcom-CCLM5-0-6	Hist/RCP2.6/8.5	r12i1p1	v1
ICHEC-EC-EARTH	KNMI-RACMO22E	Hist/RCP2.6/8.5	r12i1p1	v1
ICHEC-EC-EARTH	SMHI-RCA4	Hist/RCP2.6/8.5	r12i1p1	v1
MIROC-MIROC5	CLMcom-CCLM5-0-6	Hist/RCP2.6/8.5	r12i1p1	v1
MPI-M-MPI-ESM-LR	SMHI-RCA4	Hist/RCP2.6/8.5	r12i1p1	v1
NCC-NorESM1-M	SMHI-RCA4	Hist/RCP2.6/8.5	r1i1p1	v1
NOAA-GFDL-GFDL-ESM2M	SMHI-RCA4	Hist/RCP2.6/8.5	r1i1p1	v1

Table 3. Statistics from the evaluation in Figure 3. Correlation (R), root mean squared error (RMSE) and percentage bias (PBIAS) are given relative to downscaled ERA5 data T-MET.

Statistic	Scheme	tas [K]	pr [kg m^{-2} s^{-1}]	rsds [W m^{-2}]	rlds [W m^{-2}]	hurs [%]
R [-]	CLIM	0.98	0.39	0.99	0.95	0.02
R [-]	QM	0.98	0.35	0.98	0.95	-0.11
R [-]	QM_MONTH	0.99	0.55	0.99	0.97	0.74
RMSE	CLIM	2.27	2.17e-05	43.00	24.19	17.29
RMSE	QM	1.04	1.75e-05	19.09	7.41	6.89
RMSE	QM_MONTH	0.84	1.58e-05	10.46	5.67	3.94
PBIAS [%]	CLIM	0.7	28.8	-21.1	9.1	23.3
PBIAS [%]	QM	0	0.3	0.1	0.1	-0.1
PBIAS [%]	QM_MONTH	0	0.2	0	0.1	-0.1

Table 4. Statistics from the evaluation in Figure 4. Correlation (R), root mean squared error (RMSE) and percentage bias (PBIAS) are given relative to station measurements at Weissfluhjoch (STATION).

Statistic	Scheme	tas [K]	pr [kg m^{-2} s^{-1}]	rsds [W m^{-2}]	rlds [W m^{-2}]	hurs [%]
R [-]	CLIM	0.99	0.22	0.97	0.97	-0.16
R [-]	QM	0.98	0.20	0.97	0.96	-0.30
R [-]	QM_MONTH	0.98	0.32	0.97	0.94	0.76
R [-]	T-MET	0.98	0.35	0.97	0.94	0.76
RMSE	CLIM	1.86	2.59e-05	29.41	15.61	16.94
RMSE	QM	1.01	2.82e-05	22.89	11.59	9.31
RMSE	QM_MONTH	1.15	2.76e-05	23.78	12.94	5.34
RMSE	T-MET	1.15	2.85e-05	24.27	13.45	5.24
PBIAS [%]	CLIM	0.6	2.6	-13.9	5.4	21
PBIAS [%]	QM	-0.1	-20.1	9.1	-3.3	-2
PBIAS [%]	QM_MONTH	-0.1	-20.2	9.1	-3.3	-2
PBIAS [%]	T-MET	-0.1	-20.3	9	-3.4	-1.9

Table 5. Statistics from Figure 6 evaluation comparing T-MET results against IMIS station data for the historical period.

Variable	Mean IMIS	Mean T-MET	R	Bias	RMSE
TA [K]	273.8	273.7	0.998	-0.12	0.63
HS [m]	0.70	0.63	0.995	-0.07	0.095