Observed snow depth trends in the European Alps 1971 to 2019

Michael Matiu[1], Alice Crespi[1], Giacomo Bertoldi[2], Carlo Maria Carmagnola[3], Christoph Marty[4], Samuel Morin[3], Wolfgang Schöner[5], Daniele Cat Berro[6], Gabriele Chiogna[7,8], Ludovica De Gregorio[1], Sven Kotlarski[9], Bruno Majone[10], Gernot Resch[5], Silvia Terzago[11], Mauro Valt[12], Walter Beozzo[13], Paola Cianfarra[14], Isabelle Gouttevin[3], Giorgia Marcolini[8], Claudia Notarnicola[1], Marcello Petitta[1,15], Simon C. Scherrer[9], Ulrich Strasser[8], Michael Winkler[16], Marc Zebisch[1], Andrea Cicogna[17], Roberto Cremonini[18], Andrea Debernardi[19], Mattia Faletto[18], Mauro Gaddo[13], Lorenzo Giovannini[10], Luca Mercalli[6], Jean-Michel Soubeyroux[20], Andrea Sušnik[21], Alberto Trenti[13], Stefano Urbani[22], Viktor Weilguni[23]

[1]Institute for Earth Observation, Eurac Research, Bolzano, 39100, Italy
[2]Institute for Alpine Environment, Eurac Research, Bolzano, 39100, Italy
[3]Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Etudes de la Neige, Grenoble, 38000, France
[4]WSL Institute for Snow and Avalanche Research SLF, Davos, 7260, Switzerland
[5]Department of Geography and Regional Sciences, University of Graz, Graz, 8010, Austria
[6]Società Meteorologica Italiana, Moncalieri, 10024, Italy
[7]Chair of Hydrology and River Basin Management, Technical University Munich, Munich, 80333, Germany
[8]Department of Geography, University of Innsbruck, Innsbruck, 6020, Austria
[9]Federal Office of Meteorology and Climatology MeteoSwiss, Zurich-Airport, 8058, Switzerland
[10]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, 38123, Italy
[11]Institute of Atmospheric Sciences and Climate, National Research Council, (CNR-ISAC), Turin, 10133, Italy
[12]Centro Valanghe di Arabba, Arabba, 32020, Italy
[13]Meteotrentino, Provincia Autonoma di Trento, Trento, 38122, Italy
[14]Dipartimento di Scienze della Terra, dell'Ambiente e della Vita - DISTAV, Università degli Studi di Genova, Genova, 16132, Italy
[15]SSPT-MET-CLIM, ENEA, Rome, 00123, Italy
[16]ZAMG, Innsbruck, 6020, Austria
[17]ARPA Friuli Venezia Giulia, Palmanova, 33057, Italy
[18]ARPA Piemonte, Torino, 10135, Italy
[19]Assetto idrogeologico dei bacini montani, Region Valle d'Aosta, Aosta, 11100 Italy / Fondazione Montagna sicura, Courmayeur, 11013, Italy
[20]Météo-France, Direction de la Climatologie et des Services Climatiques, Toulouse, 31057, France
[21]Meteorology Office, Slovenian Environment Agency, Ljubljana, 1000, Slovenia
[22]Centro Nivometeorologico, ARPA Lombardia, Bormio, 23032, Italy
[23]Abteilung I/3 - Wasserhaushalt (HZB), BMLRT, Vienna, 1010, Austria

Correspondence to: Michael Matiu (michael.matiu@eurac.edu)

Abstract

The European Alps stretch over a range of climate zones, which affect the spatial distribution of snow. Previous analyses of station observations of snow were confined to regional analyses. Here, we present an Alpine wide analysis of snow depth from six Alpine countries: Austria, France, Germany, Italy, Slovenia, and Switzerland; including altogether more than 2000 stations, of which more than 800 were used for the trend assessment. Using a principal component analysis and k-means clustering, we identified five main modes of variability and five regions, which match the climatic forcing zones: north & high Alpine, northeast, northwest, southeast, and south & high Alpine. Linear trends of monthly mean snow depth between 1971 and 2019 showed decreases in snow depth for most stations for November to May. The average trend among all stations for seasonal (November to May) mean snow depth was -8.4 % per decade, for seasonal maximum snow depth -5.6 % per decade, and for seasonal snow cover duration -5.6 % per decade. Stronger and more significant trends were observed at periods and elevations where the transition from snow to snow-free occurs, consistent with an enhanced albedo feedback. Additionally, regional trends differed substantially at the same elevation, which challenges the notion of generalizing results from one region to another or to the whole Alps. This study presents an analysis of station snow depth series with the most comprehensive spatial coverage in the European Alps to date.

1 Introduction

In the European Alps, snow is pervasive throughout nature and human society. Snow is a major driver of Alpine hydrology by storing water during the winter season, which gets released in spring and summer and which is used for water supply, agriculture, and hydropower generation. Water stored in the snow cover also feeds alpine aquifers through the network of fault and fracture systems. Ecologically, the mountain flora and fauna depend on the timing and abundance of snow cover (Esposito et al., 2016; Keller et al., 2005; Lencioni et al., 2011). Snow is tightly linked to human culture in the European Alps and has brought economic wealth to previously remote regions through tourism (Beniston, 2012a; Steiger and Stötter, 2013). Since snow cover depends on temperature and precipitation, ongoing climate change in the Alps and especially rising temperatures and changing precipitation patterns affect the abundance of snow (Beniston and Stoffel, 2014; Gobiet et al., 2014; Steger et al., 2013). Snow cover extent decreased globally, while for snow mass some regions experienced increases (Pulliainen et al., 2020). Decreases are expected for the future, especially at low elevations, with more uncertain trends in observations and future projections at higher elevation (Beniston et al., 2018; Hock et al., 2019; IPCC, 2019).

Observations are needed to assess ongoing changes in snow cover. The most widespread snow cover measurements are snow depth (HS), depth of snowfall (HN, also denoted fresh snow or snowfall), snow water equivalent (SWE), snow cover area (SCA), and snow cover duration (SCD). Snow depth and depth of snowfall measurements have been scientifically documented in the European Alps since the late 18th century (Leporati and Mercalli, 1994). Such measurements indicate the height of the snow cover relative to the ground (snow depth) or a reference surface, usually a board (depth of snowfall), are performed each morning by observers, and only require a graduated stake or rod and a meter stick . While automatic sensors have been

developed in the recent decades, most European weather and hydrological services continue with manual observations. Although there is a trend towards automatization, missing standards on the processing of the data (even at national level) impedes their uptake (Haberkorn, 2019; Nitu et al., 2018). The main limitation of snow depth and depth of snowfall measurements is that their number decreases sharply with elevation, with few stations available above 3000 m in the European Alps. SWE is the mass of snow per unit surface area, which corresponds to the amount of water stored in the snow cover and thus is a hydrological key variable. However, its measurement is far more complicated and available with lower temporal frequency than snow depth, and thus not as widely observed. SCA and SCD identify the spatial extent and temporal duration of snow on the ground. SCD can be inferred from snow depth measurements using a threshold, or more recently from satellite observations, which also allow SCA retrieval at different spatial scales from tens of metres to several kilometres. The main benefit of satellite observations is that they cover the whole elevational gradient and are also available in data-scarcer regions. Satellite observations can identify SCA and SCD at high spatial resolutions (1 to 5 km for decadal length time periods), and less accurately SWE at coarser resolution (~25km) (Schwaizer et al., 2020). However, they typically cover a relatively short time period and are hampered by cloud cover and rugged topography (Bormann et al., 2018), and the satellite orbit might not provide a worldwide cover. An application of global satellite imagery for 2000-2018 has shown SCD declines for 78% of global mountain areas and only a few regions with increasing SCD (Notarnicola, 2020), although the short time span of 19 years is a limiting factor in interpreting these trends.

The European Alps are densely populated and have a long history of manual snow depth and depth of snowfall observations, which makes them ideal to study long-term trends over a large spatial domain with complex topography and strong climate gradients. Not surprisingly, much literature on the topic exists (see Table B1 in Appendix B for an overview). However, most studies are limited in their spatial extent to regions or nations, restricted by a lack of data sharing, harmonized data portals, and joint projects or initiatives fostering such analyses (Beniston et al., 2018).

The most relevant findings of the latest literature on snow cover trends (Table B1) can be summarized as follows. Snow variables exhibited a strong temporal and spatial variability (e.g. Beniston, 2012b; Schöner et al., 2019). Long-term analyses identified periods of high snow cover in the 1940s/50s, as well as in the 1960s/70s, followed by absolute minima in the 1980s and early 1990s, with some recovery afterwards, but not to the pre-1980s values (Marty, 2008; Micheletti, 2008; Scherrer et al., 2013; Schöner et al., 2009; Valt and Cianfarra, 2010). Trends were strongly related to elevation (Laternser and Schneebeli, 2003; Marcolini et al., 2017b; Valt et al., 2008) and were mostly negative at low elevations (Bach et al., 2018), while higher elevations showed no change or even increases (Marty et al., 2017; Terzago et al., 2010). Snow melt was identified as the main contribution to the decreasing trends (Klein et al., 2016), which explains the pronounced trends at low elevations and in spring (Marty et al., 2017). Finally, after accounting for elevation, regional differences between trends were observed (Beniston, 2012b; Laternser and Schneebeli, 2003; Schöner et al., 2019; Terzago et al., 2013).

Quantitatively synthesizing all these studies into a common Alpine view is challenging and thus the provision of quality-ensured information on snow cover climatology and trends at larger extents, such as the whole Alpine mountain range, is hampered (Hock et al., 2019). The challenge starts from the different definitions of the studied seasons, which range from

December–February to October–May, and thus sometimes include start, middle, and end of the season. Difficulties also arise in the selection of existing snow variables and indices, such as mean snow depth, maximum snow depth, snow days (based on thresholds from 1 to 50 cm), 3-day cumulative values, etc. Naturally, the station series are of different lengths, and the studied periods get longer for the more recent studies. And finally, the statistical methods differ from one study to another: linear regressions, Mann-Kendall tests, Sen slopes, moving window approaches (windows ranging from 5 to 20 years), breakpoint analysis, principal component analysis / empirical orthogonal function analysis (PCA / EOF) and more.

To overcome these limitations, we embarked on the effort to collect and analyse an Alpine wide data set of snow measurements from stations covering Austria, France, Germany, Italy, Slovenia, and Switzerland. The main aim is to understand how changes in snow cover vary over space and time by applying the same methods to an as homogenous as possible Alpine wide data set. This approach avoids sub-regional perspectives, inconsistencies from single data sources and different methods, and influences of artificial boundaries such as national borders. Since we wanted the data collection effort to be of use for the scientific community, we provide as much as possible of the data openly accessible (as far as data policies allow us to). The remainder of the paper is structured as follows: Section 2 introduces the data and the statistical methods, Section 3 presents results and discusses them, while Section 4 provides conclusions.

2 Data and methods

2.1 Study region

The European Alps extend with their arc-shaped structure over more than 1000 km from the French and Italian Mediterranean coasts to the lowlands east of Vienna, covering south-eastern France, Switzerland, northern Italy, southern Germany, Austria, and Slovenia (see Fig. 1 (a)). The Alpine region is characterized by a very complex orography with large elevation gradients and deep valleys of different orientation intersecting the ridge and shaping numerous mountain massifs.

Regarding their climatic setting, the European Alps are located in a transitional area influenced by the intersection of three main climates: The zone impacted by the Atlantic Ocean with moderate wet climate, the zone linked to the Mediterranean Sea characterized by dry summers and wet and mild winters, and the zone characterized by European continental climate with dry and cold winters and warm summers. Elevational effects and very small-scale climatic features originating from the complex Alpine topography are superimposed on this large-scale climatic setting (Auer et al., 2005; Isotta et al., 2014).

The interaction of the three climate forcing zones together with the topography of the Alps results in climatic gradients along the north-south and west-east directions. The intersection of these two gradients can be characterized by four main climate regions, as shown by Auer et al. (2007). The first and sharpest climatic border is along the central main ridge, separating the temperate westerly from the Mediterranean subtropical climate. The second climatic border separates the western oceanic form the eastern continental influences.

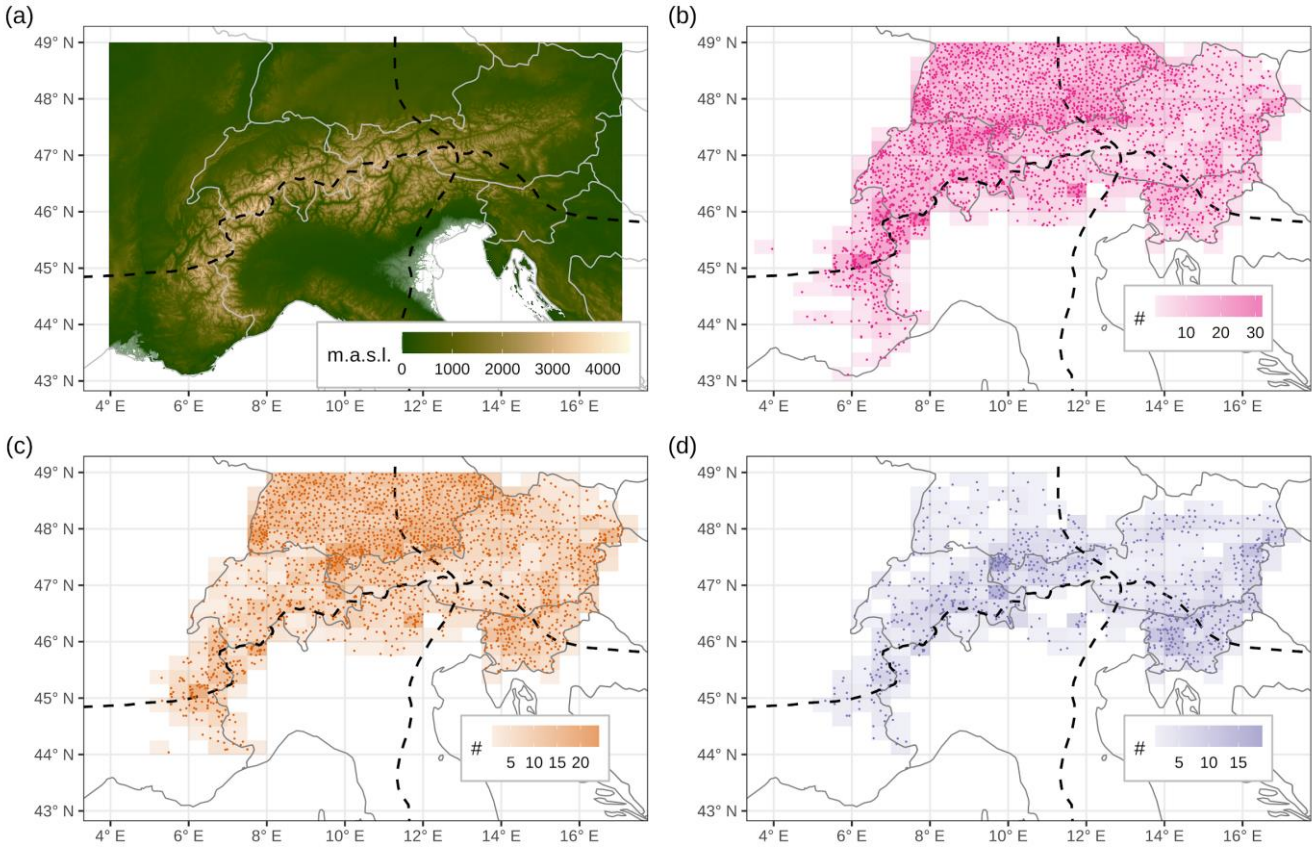


Figure 1: Topography of the European Alps (a) and overview of station locations (b-d). (a) shows the SRTM30 DEM (Shuttle Radar Topography Mission Digital Elevation Model) with ~1 km resolution. (b) shows the location of snow depth measurement locations that were available (provided). (c) shows the locations of stations used in the regionalization analysis. (d) shows the stations used for the long-term trend analysis. The station density for a 0.5x0.25 degree grid is shown underneath the points in (b)-(d). The main climatic divides from Auer et al. (2007) are shown as dashed lines in (a)-(d). See also Appendix A, Sec. 2.4 and Sec 2.5 for selection criteria.

2.2 Data sources

Acquisition of snow observation data was performed by using open data portals and by directly contacting data providers (see Table 1 for an overview). For Austria, the Austrian Hydrographical Service (HZB, Hydrographisches Zentralbüro) offers free download of their data for the recent decades, and additional historical data at the seasonal scale was kindly provided by the HZB. For France, data was kindly provided by the national weather service Météo-France. This includes data collected as part of the collaborative network (réseau nivo-météorologique) between Météo-France and mountain stakeholders (in particular

Domaines Skiables de France, Association Nationale des Maires de Stations de Montagne, Association Nationale des Directeurs de Pistes et de la Sécurité de Stations de Sports d'Hiver). For Germany, data was downloaded from the national weather service's (DWD, Deutscher Wetterdienst) open data portal using the R-package rdwd. For Germany, only stations below 49° N were downloaded. For Italy the data was kindly provided by many regional authorities:

- for the province of Bolzano from the hydrographical office of Bolzano (BZ)
- for Friuli Venezia Giulia (FVG) from the regional weather observatory (OSMER, Osservatorio meteorologico regionale), which is part of the ARPA (Agenzia regionale per la protezione dell'ambiente) FVG, and where the data was collected and cleaned by the Servizio foreste e corpo forestale struttura stabile centrale per l'attivita' di prevenzione del rischio da valanga
- for Lombardy from the ARPA Lombardia
- for Piedmont from the ARPA Piemonte
- for the province of Trento from Meteotrentino (TN), with some additional long-term series previously analysed (TN_TUM, Marcolini et al., 2017a)
- for the Aosta Valley (VDA) from the civil protection office (CF: Centro funzionale, Regione Valle d'Aosta) and from the avalanche office (AIBM: Assetto idrogeologico dei bacini montani, Regione Valle d'Aosta)
- for Veneto from the avalanche office (Centro valanghe di Arabba), which is part of the ARPA Veneto
- finally, additional data for Piedmont and Aosta Valley was provided by the Italian meteorological society (SMI, Società Meteorologica Italiana)

For Slovenia, data was kindly provided by the Slovenian Environmental Agency (ARSO, Agencija Republike Slovenije za okolje). For Switzerland, data was downloaded from the IDAWEB portal of the national weather service MeteoSwiss, and additional data was kindly provided by the WSL Institute for Snow and Avalanche Research SLF. This dataset comprises the entire geographical range of the European Alps, yet we are aware of the existence of additional data sets (such as in the private sector, or public but not yet digitized), which unfortunately were not included in this analysis, and whose inclusion would be beneficial for even more robust results.

The data consists of daily measurements of snow depth (HS) and depth of snowfall (HN). The largest part of the data are manual measurements. Some automatic measurements were included in the dataset provided for France. For a few sites in the Aosta Valley in Italy, manual series were merged with automatic series. This was done in order to extend up to the present some records that were dismissed at the beginning of the last decade, and this was performed in close communication with the operating office. While the observers follow slightly different guidelines in each country or network, the observation modalities are remarkably similar, thus allowing a combination of the different sources. For more detailed information on the measuring modalities, we refer to the European Snow Booklet (Haberkorn, 2019). Values of HS and HN were rounded to full centimetres. The further processing, quality checking, and gap filling are described in Appendix A. For all the following statistical analyses the quality checked and gap filled data were used.

The fraction of stations used from the MeteoSwiss data is very low compared to the other networks. The MeteoSwiss data contains a large number of stations from the manual precipitation network which is not dedicated to snow. Many stations contain an important data gap for the 1981–1997 period that rendered a large fraction of the stations unusable for this study.

The homogenization of series, which is the removal of non-climatic parts in the time series, such as e.g. caused by instrumentation changes or station relocations, is a standard practice in long-term temperature and precipitation records (Auer et al., 2007). Applying the same tools to snow depth is not straightforward. There is discussion ongoing on the appropriate homogeneity tests and suitable observation frequency, such as daily, monthly, or seasonal (Marcolini et al., 2017a, 2019; Schöner et al., 2019). An analysis of a data set with parallel snow measurements indicates that snow cover duration and maximum snow depth are amongst indicators least affected by inhomogeneities (Buchmann et al., 2021). Current research tries extending existing approaches with new innovations (Resch et al., 2020). Homogenization could improve the robustness of estimated trends, and be especially useful for areas with sparse observations, such as for elevations above 2000 m. Given the large extent of our dataset, it was not possible to apply a common homogenization framework for our study and we leave this for future studies.

Table 1: Overview of the number of stations with daily data provided by the different data sources. The data source consists of a country abbreviation, followed by the data source. Country abbreviations are AT for Austria, CH for Switzerland, DE for Germany, FR for France, IT for Italy, and SI for Slovenia, respectively. For source abbreviations, please see Sec. 2.2. Station numbers are shown for depth of snowfall (HN) and snow depth (HS) time series. See Appendix A, Sec 2.4 and Sec 2.5 for more details on station selection procedures associated with the different types of analyses. HN was not analysed but used for checking HS.

Data source	HN	HS	HS used (regionalization)	HS used (trend analysis)
AT_HZB	653	652	588	335
CH_METEOSWISS	505	501	142	79
CH_SLF	96	96	94	84
DE_DWD	956	964	830	104
FR_METEOFRANCE	239	286	145	45
IT_BZ	60	64	48	0
IT_FVG	30	30	18	8
IT_LOMBARDIA	11	11	11	0
IT_PIEMONTE	34	34	24	15
IT_SMI	6	8	8	7
IT_TN	52	52	29	8

IT_TN_TUM	0	5	1	0
IT_VDA_AIBM	57	57	17	5
IT_VDA_CF	0	17	11	3
IT_VENETO	10	11	11	9
SI_ARSO	130	172	172	152
Total sum	2,839	2,960	2,149	854

2.3 Data overview

The locations of the stations are shown in Fig. 1 (b-d), the availability of stations in time in Fig. 2 (a), the elevational distribution in absolute terms in Fig. 2 (b) and in relative terms in Fig. 2 (c). The stations cover the whole Alpine arc, but they are distributed with different station densities arising from the different national and regional networks. As expected, most stations were found at lower elevations, the maximum number was at ~500 m, and sharply declining for higher elevations. Above 2000 m, the number was low, and no stations above 3200 m were available for this study. The longest series dates back to the late 19th century for HS (Passau_Maierhof in Germany, starting 1879). The total number of available HS stations depended on the availability of digitized data. It slowly started increasing around ~1900, with significant jumps in the 1960s and 1970s, when the French, Slovenian and Austrian series started, and in the 1980s, when Germany had a large network increase. The highest number of stations was available after the 1980s, with approximately 2000 stations. The total number of stations dropped significantly after 2017, because the data for Austria was only available until 2016, due to the delays induced from performing quality checks by the data provider. Moreover, the data collection was performed between 2019 and 2020, thus some sources ended in between. We used two different periods for the two analyses that we performed. For the regionalization we aimed to have the largest possible spatial extent and density of the stations, so the period 1981 to 2010 was chosen, because it is the period with the highest number of stations. For the trend analysis, we aimed to have as long as possible trends that sample the whole region, so the period 1971 to 2019 was chosen, because it offered the best tradeoff between station coverage and period length.

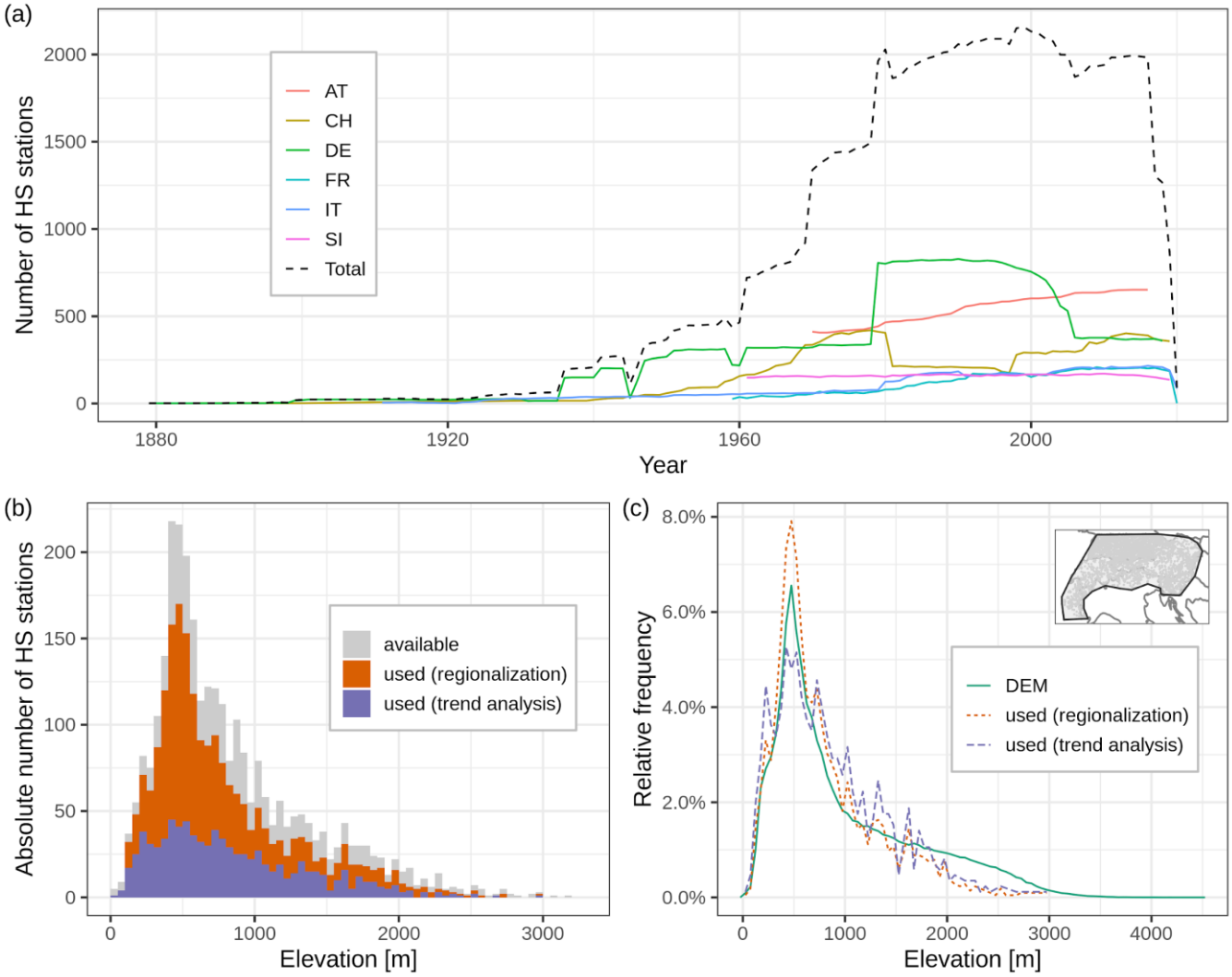


Figure 2: Overview of temporal data availability and station elevation. (a) The number of stations with daily data (before gap filling) is shown per year and country, and a total sum for the whole Alpine region. Stations are included in the count, if they have at least one non-missing observation in the respective calendar year. This simple threshold was chosen, because the aim of this figure is to show the availability and network abundance. Country abbreviations are as in Table 1. (b) The elevational distribution of snow depth (HS) stations in absolute numbers. For the histogram 50 m bins were used. (c) Comparison of the relative elevational distribution of the station locations versus a digital elevation model (DEM). The distribution of the stations is shown in relative terms, using the same bin width (50m) as in the histogram in (b), but normalized to show the relative frequency instead of absolute numbers, and displayed as lines instead of bars. This is compared to the elevation for the whole area spanned by the stations (see polygon in inset map; area was outlined manually along the stations), which is extracted from the SRTM30 DEM (Shuttle Radar Topography Mission, ~1 km resolution).

2.4 Regionalization

An empirical orthogonal function (EOF) analysis, also called principal component analysis (PCA), was conducted to determine the common modes of spatial variability. PCAs are widely employed in climatological studies to evaluate spatial modes of variability (Storch and Zwiers, 1999). They have been employed for meteorological records in the European Alps (Auer et al., 2007) and also for snow variables (López-Moreno et al., 2020; Scherrer and Appenzeller, 2006; Schöner et al., 2019; Valt and Cianfarra, 2010). For the PCA we used daily quality checked and gap filled data. However, the gap filling was only employed, when enough confidence in the filled value could be expected (see Appendix A for a detailed description). So some of the series still had gaps. Because the aim of this regionalization was to have a large spatial coverage, we did not want to exclude series with only few missing values. Consequently, we used a modification of the PCA algorithm that allows using data with gaps to estimate the principal components (Taylor et al., 2013).

The PCA was applied to the daily data from December to April for the hydrological years 1981 to 2010. The period was chosen, because it is long enough to provide a climatological reference (30 years), and it is the period that has the largest number of stations available. A hydrological year is defined here as starting in October, and it is designated as the calendar year of the ending month (e.g. December 1998 to April 1999 belong to the hydrological year 1999). Only those stations were selected that had at least 70% of daily data available in this period. Each series was scaled to zero mean and unit variance before applying the PCA.

In order to identify spatially homogeneous regions within the Alpine domain, we performed a k-means clustering on the estimated PCA matrix. We tested configurations with 2 to 8 clusters with the PCA matrix and with 2 to 8 PCs as input. We also applied k-means clustering directly on scaled daily observations of snow depth for comparison. To identify the best number of clusters, we used the "elbow-method", average silhouette coefficients, and visual interpretation. For the "elbow-method", the fraction of explained variance is plotted against the number of clusters, and the "elbow" of this curve is the point where the increase in explained variance becomes marginal. This is a semi-objective method, because an elbow cannot always be clearly identified. The silhouette is a measure of how well an observation fits into its own cluster versus the others. For an observation i in cluster C_i, the silhouette coefficient is $1 - a(i)/b(i)$ if $a(i) < b(i)$, $b(i)/a(i) - 1$ if $a(i) > b(i)$, and 0 if $a(i) = b(i)$, where $a(i)$ is the mean distance between i and all other points in the same cluster, and $b(i)$is the smallest mean distance of observation i to all other clusters. Specifically, $a(i) = \frac{1}{|C_i|-1}\sum_{j \in C_i, i \neq j} d(i,j)$ and $b(i) = min_{k \neq i} \frac{1}{|C_k|}\sum_{j \in C_k} d(i,j)$, where $d(i,j)$ is the Euclidean distance between observations.

The optimal number of clusters varied between 2 and 5 depending on the input (observations or PCA matrix) and depending on the metric (elbow in variance explained or average silhouette coefficients). Additional PCs explained only less than 2.6% of variance. After looking at the clustering results on maps (see Fig. S1), all 2 to 5 clusters are meaningful. They simply highlight different aspects of the snow depth spatial variability, such as the gradients along elevation, north-south and west-east. Finally, five clusters based on the PCA matrix were chosen, because they provide the best trade-off between the semi-objective metrics and the patterns expected from the climatic drivers.

2.5 Trend analysis

For the trend analysis monthly and seasonal indices were used, which are indicative for different aspects and times of the snow season: monthly mean HS for November to May, mean winter HS (December to February, DJF), mean spring HS (March to May, MAM), mean seasonal HS (November to May), maximum HS from November to May (maxHS), early season snow cover duration (SCD, November to February), late season SCD (March to May), and full season SCD (November to May). SCD was the number of days with HS above 1 cm (Brown and Petkova, 2007). Indices were calculated from the quality checked and gap filled daily snow depth observations, if more than 90% of the daily values in the respective period were available. Trends of all indices were calculated for the period 1971 to 2019 for stations with complete data in the period. For the monthly mean HS analysis only, April and May series displaying mean HS less than 1 cm in all years were discarded, because these are insignificant snow amounts and divert attention from the other sites; series of the other months at the site were still included. The number of series available for each snow variable differs: the largest number of series is available for the monthly mean HS, less for the half-seasonal (three to four months) and the fewest number for the full-season indices.

Trend analysis was performed using two generalized least squares (GLS) regression. GLS was used because it allows accounting for changes in the variance (Pinheiro and Bates, 2000). This was employed because the monthly snow depth series exhibited a change in the inter-annual variability, especially at the end of the season, where monthly snow depths approached zero. The regression formula was $y_t = \beta_0 + \beta_1 t + \epsilon_t$, where y_t is the value of the respective snow variable in year t (centred such that year 1971 becomes year 0), β_0 and β_1 are the estimated regression coefficients and ϵ_t are the normally distributed errors with mean zero. GLS allows the variance to depend on the year t with $Var(\epsilon_t) = \sigma^2 * exp(2 * \gamma * t)$, where γ is a coefficient to be estimated in the inference procedure that indicates the change in variance associated to t. The GLS regressions for monthly mean HS showed a significantly improved goodness-of-fit ($p<0.05$, likelihood ratio test) for 40% of all cases, and, specifically for November, April and May, even for more than 60% when compared to ordinary least squares (OLS) that assumes a constant error variance. The significance of trends was assessed using a 95% confidence level. For the fraction of variance explained by the trend, we used the R squared statistic. To determine the magnitude of the interannual variability after accounting for the trend, we used the standard deviation of the model residuals.

An alternative for dealing with such heteroscedastic data is to use the robust nonparametric Theil-Sen trend estimator with the Mann-Kendall test for significance assessment. We systematically evaluated the differences in the estimated trend magnitudes and trend significances of the Theil-Sen approach versus the GLS model, and found only negligible differences (Fig. S13 and Table S10): The mean difference between trend estimates was 0.02 cm per decade, the correlation between trend estimates was 0.96, and the agreement of significance based on a p-value threshold of 0.05 was 86%.

The SCD variables are bounded counts, which can pose problems to the assumption of standard linear regression with normally distributed errors. This was only problematic for very low and very high elevation sites, which display many SCD values at the minimum or maximum. For MAM this concerns series below 500 m and above 2000 m, while for NDJF and NDJFMAM

this is problematic below 250 m and above 2500 m. Instead, for such count data a probability distribution such as Negative Binomial would be more appropriate (Venables and Ripley, 2002). Compared to the Poisson distribution, the Negative Binomial family accounts for overdispersion. We evaluated the differences in trend estimates and trend significance between the Negative Binomial linear model and the GLS model. Since the Negative Binomial linear model gives relative estimates of trends, these were transformed to absolute decadal trends for comparison. Again, differences were negligible on average (Fig. S13 and Table S10). Consequently, we applied the GLS model for all snow variables.

2.6 Air temperature and precipitation data

In order to study the relationship of snow depth with temperature and precipitation, we extracted temperature and precipitation series for each station from available gridded products. While gridded data sets clearly have some shortcomings, e.g. comparisons to point observations need a cautious interpretation (Salzmann and Mearns, 2011), their strength is the spatial and temporal coverage.

Two types of products were considered, the first is a reanalysis and the second is an observation-based spatial analysis. For the reanalysis, we used temperature and precipitation from the MESCAN-SURFEX data set (Bazile et al., 2017), which was produced during the UERRA (Uncertainties in ensembles of regional reanalyses) project and which is available via the Copernicus data store (CDS). It covers the period from January 1961 to July 2019 on a 5.5 km grid. Precipitation is available as total daily sum and temperature at 6-hour intervals (00, 06, 12, 18 UTC). For the observational based data, we chose E-OBS v20.0e for mean daily temperature (Cornes et al., 2018), and the Alpine precipitation grid dataset (EURO4M-APGD) for total daily precipitation (Isotta et al., 2014; Isotta and Frei, 2013). E-OBS v20.0e spans the period from January 1950 to July 2019 on a 0.1° grid. APGD covers the period January 1971 to December 2008 on a 5 km grid. It should be noted that the observation-based precipitation grids do not account for undercatch, which can lead to uncertainties at high elevations and in winter (Prein and Gobiet, 2017).

In order to assign grid cells to stations for temperature and precipitation, we selected those grid cells which contain the stations. Consequently, some nearby stations could have the same series of temperature and precipitation. The daily (or 6-hour for temperature MESCAN-SURFEX) series were aggregated to monthly means for temperature and monthly sums for precipitation.

The gridded products have a reference orography that, in complex mountain terrain, can differ significantly from the elevation of the point observation, thus e.g. introducing biases in temperature. Thus, temperatures were adjusted using a constant lapse rate of 6.5 °C km^{-1}.

Monthly temperature and precipitation can be considered largely independent from one month to the next, while snow cover is a cumulative process across the snow season. Because of this, seasonal comparisons were performed with average seasonal temperature and precipitation for winter (December to February), spring (March to May), or the whole snow season (November to May). The time period 1981 to 2010 was used, which had the densest station coverage. Climatological averages were

computed for all seasons using the quality-checked and gap filled snow depth data. Since EURO4M-APGD ends in 2008, the time period 1981 to 2008 was used for the observation-based products. The manuscript contains results from the comparison with the reanalysis product (MESCAN-SURFEX), and the results from the observation-based products are shown in the supplementary material as sensitivity analysis.

3 Results and discussion

3.1 Regionalization of daily snow depths 1981 to 2010

The PCA of daily snow depth series yielded five main modes of spatial variability, which explained in total 84% of the variance in the period December to April from 1981 to 2010 (Fig. 3). The first PC explained 54.3% of variance and distinguished between high to middle and low elevation stations (approximate threshold 500–1000 m, Fig. 4). It explained the variability in snow depth for stations above 1000 m, and was probably also partly linked to the permanence (or permanent absence) of snow cover, which is why also some low elevation sites presented similar loading to the high sites (a PC loading can be considered the correlation of the original series with the principal component). The second PC explained 11.9% of variance and was also linked to elevation, but captured the variability below 1000–1500 m (Fig. 4). Consequently, PC1 and PC2 together captured the variability across the whole elevation range. The third PC explained 8.1% of variance and separated the stations into north and south of the main ridge. The fourth PC explained 6.0% of variance and separated east from west. The fifth PC explained 3.7% of variance and separated the south–eastern and north–western stations from the rest.

Some gradients in the PC loadings map (Fig. 3) could give the impression that data artefacts between the different data providers exist, such as at the Austrian-German border in PC2 and PC5, or at the French-Italian border for PC3-5. However, this is caused by the fact that the administrative borders in the Alps are tied to topography, and thus closely located near elevational borders (Fig. 1(a)). A version of Fig. 3 subdivided by data provider highlights clearly that the gradients were not associated with the administrative borders (Fig. S2 in the supplementary material).

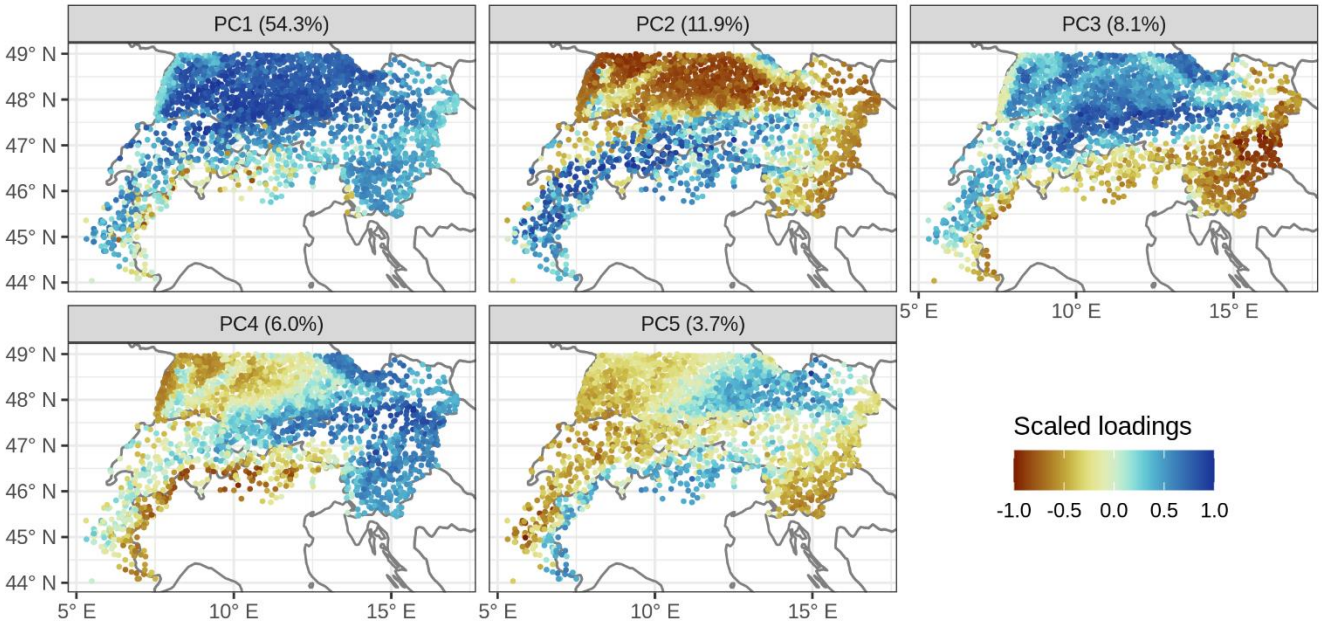


Figure 3: Main modes of variability in daily snow depth series. The plots show scaled loadings for the first five principal components (PCs), which can be considered the correlation of the original series with the respective PC. The title in each panel contains the amount of variance explained by the respective PC. The principal component analysis was applied on daily snow depth data from December to April for the hydrological years 1981 to 2010, for stations that had at least 70% of available data.

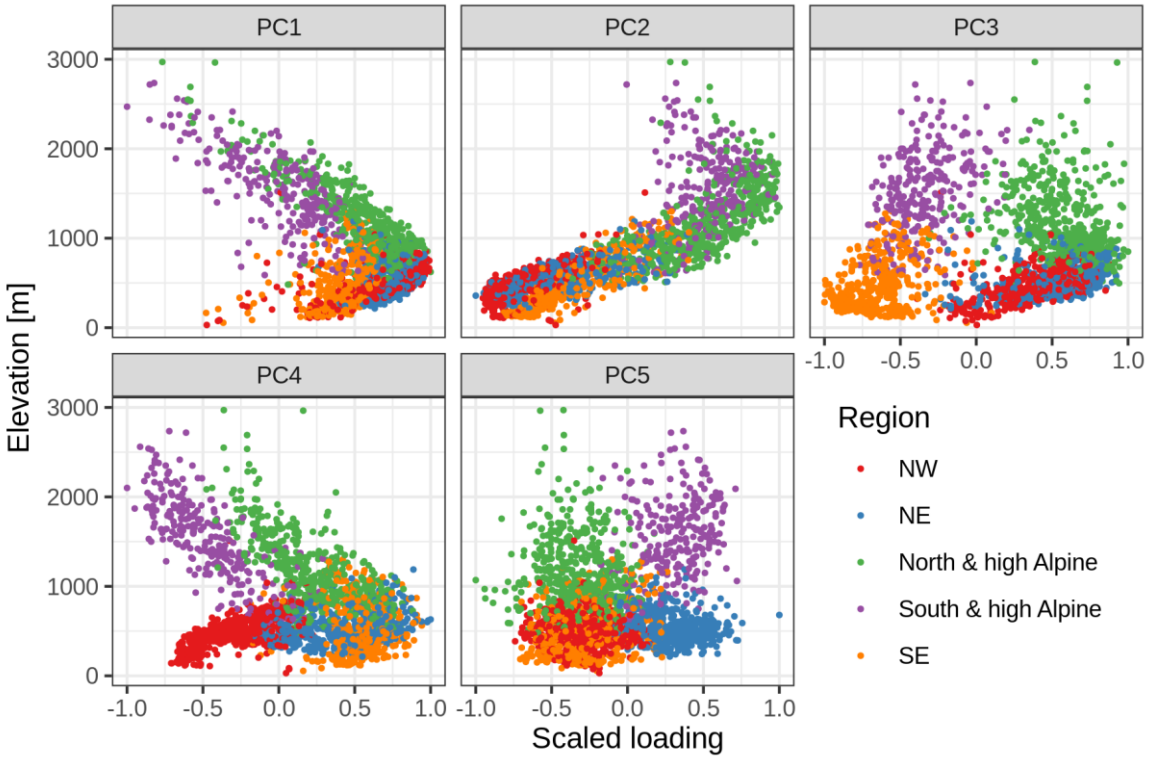


Figure 4: Scatterplots of principal component (PC) loading versus elevation and region. The PC loading can be considered the correlation of the original series with the respective PC. See Fig. 3 for a map of the PC loadings, and Fig. 5 for a map of the regions.

The PCA loadings from the five PCs were used as input for a clustering algorithm (k-means), which divided the stations into five clusters or regions (Fig. 5). This yielded three regions in the north: northwest (NW) with a median elevation of 472 m (min–max: 30–1510 m), which contained stations from southwest Germany, northwest Switzerland, few from France, and a few from eastern Austria; northeast (NE) with a median elevation of 515 m (215–1188 m), which contained stations from southeast Germany and north Austria; North & high Alpine with a median elevation of 1050 m (482–2970 m), which contained stations mainly located in France, Switzerland, and Austria, but also includes the high-elevation sites in Germany, such as in the Black forest and Bavarian forest. Two regions emerged south of the main ridge: South & high Alpine with a median elevation of 1530 m (588–2735 m), which contained stations from the southern French Alps, almost all of Italy, few of southern Switzerland, and some of south Austria and east Slovenia; and finally southeast (SE) with a median elevation of 420 m (55–1300 m), which contained almost all stations from Slovenia and parts of eastern Austria.

Consequently, clusters NW, NE, and SE contained lower elevation sites, while North & high Alpine and South & high Alpine contained the higher elevations. The spatial coverage of the stations in this study included low elevations sites for Switzerland, Germany, Austria, and Slovenia, but not in France and Italy, where the available stations were mostly high elevation sites. For a future analysis, it would be interesting to include more low elevation sites from France and Italy, and see whether a third

cluster would emerge (as in the north), because the division into South & high Alpine and SE is surely also caused by the different station elevations.

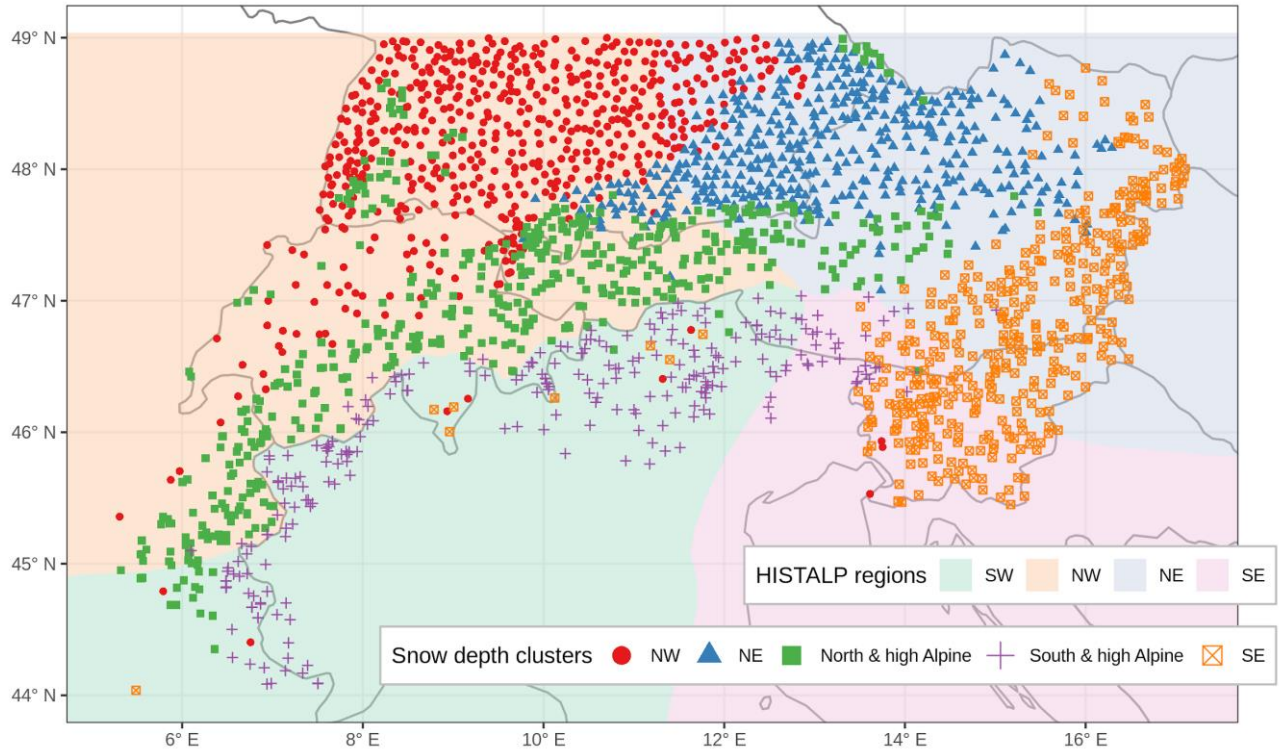


Figure 5: Clustering of stations based on daily snow depth data. Map of regions from applying a k-means clustering on the first five principal components. Underlaid are the HISTALP coarse resolution subregions (Auer et al., 2007), which were derived using a semi–automatic principal component analysis of climate variables (temperature, precipitation, air pressure, sunshine, and cloudiness).

The results from the clustering were obtained automatically and no manual post-processing or modification of the cluster assignments was performed. Additionally, the only input into the clustering algorithm was daily snow depth series and no information on location or elevation was included. Given this absence of location information in the clustering process, the estimated modes of variability and the resulting regions were very homogenous in space. However, in the clustering, some stations seemed off, such as the few "northwest" stations around Lugano in Switzerland, northern Italy, and at the Adriatic coast in Slovenia, as well as the SE stations in France, Switzerland, and northern Italy. This was not related to the used PCA algorithm that allowed gaps in data, since the results looked almost identical to a standard PCA (see Fig. S3 and S4), where the clustering agreed in 98.5% of the stations, and the same stations seemed mis-clustered. Instead, this might be related to special local climatic conditions affecting snow cover or to the fact that these stations did not have any similar neighbours in the estimated clusters. For example, the five stations in Ticino, located in Switzerland south of the main ridge, are low elevation

stations, which had no correspondence in the South & high Alpine cluster, which contained middle to high elevations. Thus, the next best clusters were SE and NW, which, however, did not fit well: these sites and all other seemingly mis-clustered stations had low silhouette values (Fig. B1), which is a measure of how well a point matches its cluster compared to the others. Low silhouette values were also found along the borders of the different clusters, especially between NW and NE, which implies a smoother transition between NW and NE compared to the north–south boundary.

The estimated modes of variability of snow are similar to previous estimates on climatic subregions in the Alps, as identified in the HISTALP project (Auer et al., 2007), and which are underlaid in Figure 5. The HISTALP regions were based on temperature, precipitation, air pressure, sunshine and cloudiness, and the division into north, south, east and west matches what we found for snow depth. Since the four regions were a compromise between all variables, they do not match perfectly to what we found for snow depth, because the individual atmospheric variables exert different controls on surface snow cover. While the north–south boundary is almost identical in the central–western part, the eastern part has large mismatches. However, if the single element boundary for precipitation were considered as main factor (cf. Fig. 8 from Auer et al., 2007), then the agreement with snow depth would be almost perfect. This finding confirms a consistent picture of the Alpine climate, in which snow depth is strongly related to precipitation and air temperature patterns.

The amount of variance explained in the PCA with five PCs (84%) might seem surprisingly high, given that snow cover is hypothesized to have a high spatial and temporal variability. The value is higher than recent estimates for the Swiss Alps, where the first three PCs explained 78% (Scherrer and Appenzeller, 2006), or for Austria and Switzerland, where the first three PCs explained 70% (Schöner et al., 2019). However, since here we included more stations and also stations from regions with different climatic influences, such as south of the main ridge, an increase in the amount of explained variance could be expected.

3.2 Snow depth climatology 1981 to 2010 and links to temperature and precipitation

Besides differences in the patterns of daily variability of the snow depth series, the regions also demonstrated different snow depth climatologies (Fig. 6). Looking at average winter (December to February) snow depth 1981–2010, the northern regions had higher snow depths than their southern counterparts. These differences became larger with increasing elevation: While below 750 m no substantial differences were observed, southern stations had ≈30% less snow than northern stations until 1750 m, and ≈20% less until 2250 m; above the number of stations is too low to obtain robust results (Table S1).

Average winter temperatures were higher in NW compared to NE and SE, and the latter two were similar. In North & high Alpine and South & high Alpine temperatures were also comparable, although northern sites were colder at 1500–2000 m. However, precipitation amounts were significantly lower south than north, and South & high Alpine sites received ≈100 mm less winter precipitation than North & high Alpine sites up to 2000 m, which amounts to ~1/3 of the precipitation north. These results suggest that the difference in December to February snow amounts north versus south are predominantly driven by precipitation differences and not temperature.

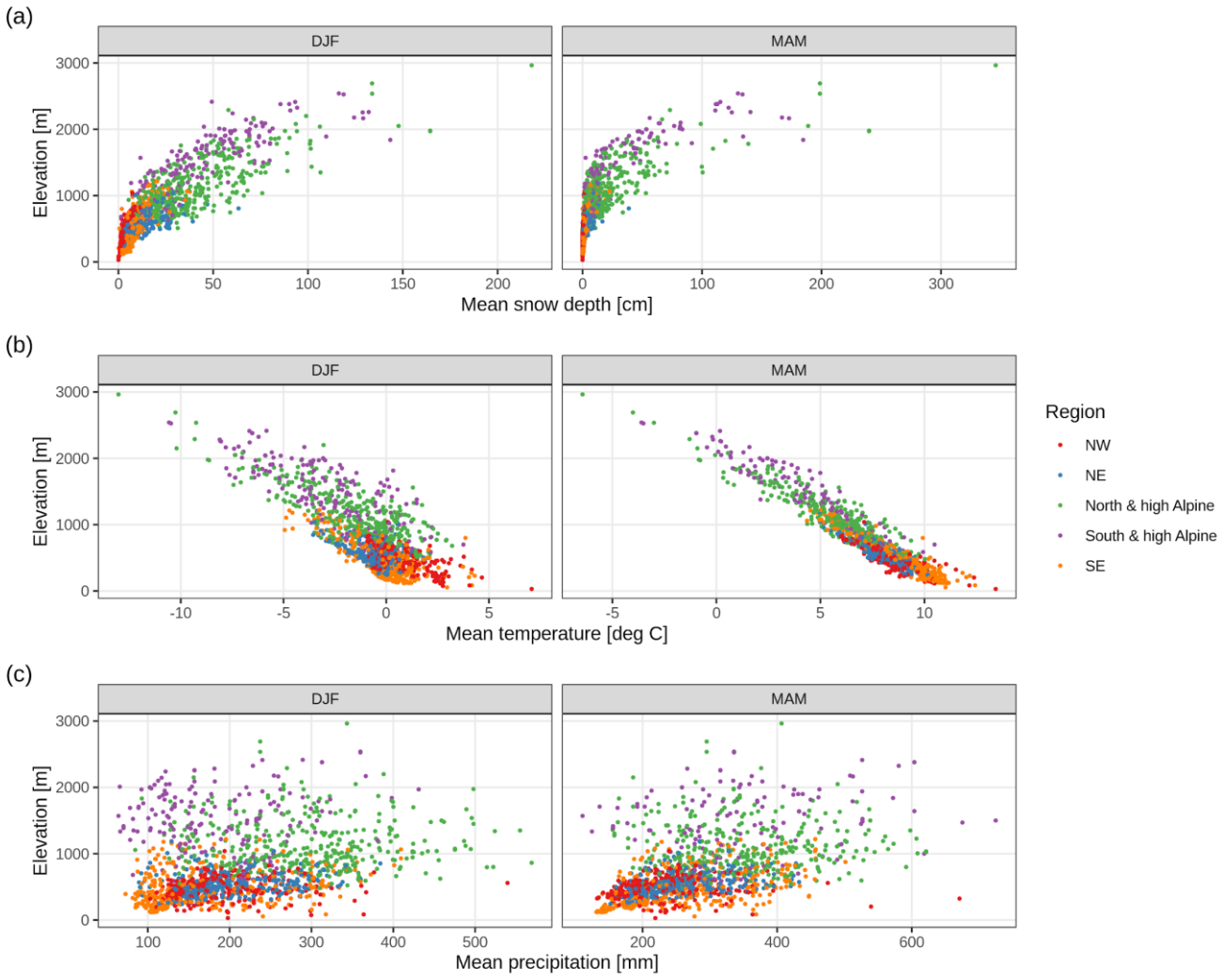


Figure 6: Climatology of (a) snow depth, (b) temperature, and (c) precipitation across regions and elevations for the winter season (December to February, DJF) and spring season (March to May, MAM). Average values are for the period 1981–2010. Each point represents one station. The temperature and precipitation values were extracted from MESCAN–SURFEX reanalysis, while the snow depths are based on station data. See also Table S1 and S2 for summary values.

Seasonal snow depth was correlated to temperature and precipitation extracted from a gridded reanalysis (MESCAN–SURFEX). Results indicated negative correlations with temperature, decreasing strongly with elevation, and positive correlations with precipitation, mildly increasing with elevation (Fig. S5). The magnitude of temperature correlations was between -0.8 and -0.5 below 1000 m, and the correlation decreased to about -0.2 up to 2000 m. For precipitation, correlations were between -0.2 and 0.7, with much higher variability than temperature. Correlations of snow depth with temperature did not differ by region. However, the stations in SE exhibited stronger (more positive) correlations with precipitation than the NE and NW regions.

The findings on the correlations agree with previous estimates for Swiss and Austrian stations (Schöner et al., 2019) in terms of signs and elevation patterns. However, our estimates are of higher magnitude for both temperature and precipitation. As sensitivity analysis, we repeated the climatology and correlational analysis using observation based spatial analyses instead of reanalysis for extracting temperature and precipitation (Fig. S6, Fig. S7, Table S3, Table S4), but results did not differ substantially from above.

3.3 Long-term trends for the period 1971 to 2019

Trends of monthly mean snow depth from November to May were mainly negative with some exceptions (Fig. 7 and Table 2). Over all stations and all months, 85% of the trends were negative and 15% positive; 23% significantly negative and 0% (only 4 station month combinations) significantly positive (for significance, p-values had to be less than 0.05). The percentage of significant negative trends was substantially higher in the spring months (March to May) and at lower elevations, irrespective of region, and it could reach 40–70% (see also Table 2).

In the low elevation regions (NE, NW, SE), snow depth was decreasing much stronger in SE than in NE or NW across all months. The mean trend of December snow depth below 1000 m in NE was -0.7 cm per decade (all further trends in the same unit) and -0.8 in SE, while in January it was -0.5 in NW, -0.6 in NE, but -1.6 in SE (Table 2). In February, NE stations even had increasing snow depth with +0.8, while NW and SE decreased. In the middle elevation (1000 to 2000 m), differences between north and south were even stronger and variable in amplitude during the snow season: in December the mean trend in North & high Alpine (N&hA) stations was stronger negative (-1.9) than compared to South & high Alpine (S&hA) stations (-0.8), but for January and February we observe the opposite behaviour, with a less pronounced negative trend in N&hA (-1.6 and -2.2) compared to S&hA (-3.9 and -5.1).

In the spring months March and April, trends in snow depth were again more negative south than north. For example, in the middle elevations (1000 to 2000 m), the mean March snow depth trend was -3.9 in N&hA compared to -7.0 in S&hA, in April -5.7 compared to -6.6, and in May -1.4 compared to -2.7. Notably, stations in S&hA above 2000 m exhibited strong variability in trends, and there were stations with increasing snow depth in all months (November to May). While mean trends were positive until January (November 2.7, December 4.0, January 0.0), mean trends were negative otherwise (February -1.9, March -2.6, April -8.3, May -9.5).

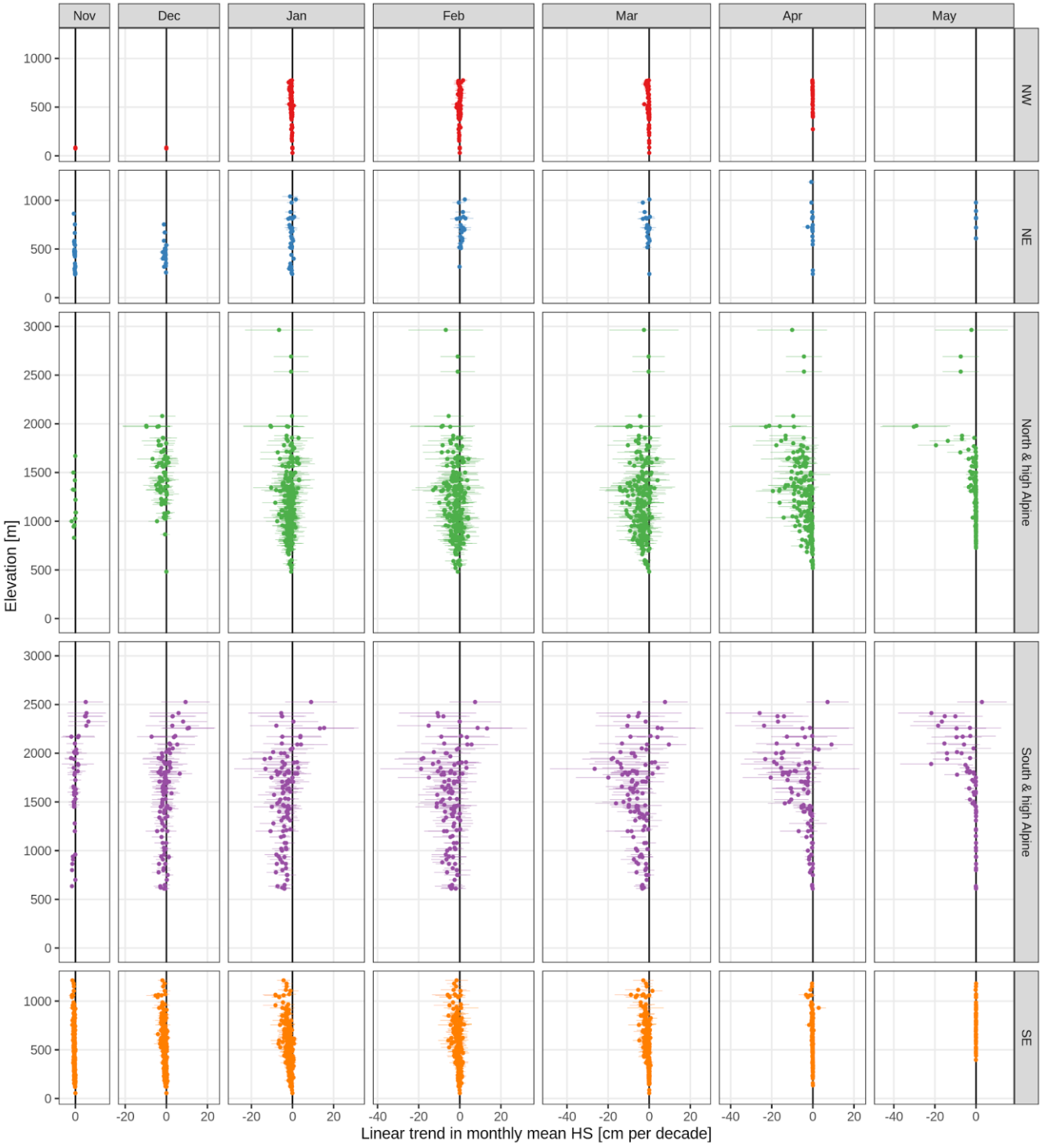

Nov
Dec
Jan
Feb
Mar
Apr
May
NW
NE
North & high Alpine
South & high Alpine
SE
Elevation [m]
Linear trend in monthly mean HS [cm per decade]

Figure 7: Long-term (1971 to 2019) linear trends in mean monthly snow depth (HS). Trends are shown separately by month (columns) and region (rows). Each point is one station. The points indicate the trend and the lines the associated 95% confidence interval.

Table 2: Overview of long-term (1971 to 2019) trends in mean monthly snow depth. Summaries are shown by month, region, and 1000 m elevation bands (0 to 1000, 1000 to 2000, and 2000 to 3000 m). Cell values are the number of stations (#), the mean trend (mean, in cm per decade), and percentages of significant negative (sig-) and positive (sig+) trends; the remaining percentage (not shown) corresponds to the total of non-significant negative and positive trends. Empty cells denote no station available (for # and mean), and no stations with significant negative or positive trends (sig- and sig+). Trends were considered significant if $p < 0.05$. See also Fig. 7. A version of the table with 500 m bands instead of 1000 m is available in the supplementary material (Table S5).

Month	Region	Elevation: (0,1000] m				Elevation: (1000,2000] m				Elevation: (2000,3000] m			
		#	mean	sig-	sig+	#	mean	sig-	sig+	#	mean	sig-	sig+
Nov	NW	2	-0.01	50.0%									
	NE	34	-0.32	41.2%									
	N&hA	4	-0.93	50.0%		9	-0.31						
	S&hA	7	-1.02	71.4%		23	-0.22			12	2.68		
	SE	218	-0.50	52.3%		8	-1.21	50.0%					
Dec	NW	2	-0.01										
	NE	24	-0.68	29.2%									
	N&hA	3	-1.72	33.3%		67	-1.91	1.5%		1	-2.02		
	S&hA	17	-1.34	5.9%		67	-0.89	1.5%	1.5%	17	3.98		
	SE	221	-0.77	24.9%		9	-2.38	44.4%					
Jan	NW	81	-0.51	12.3%									
	NE	32	-0.55	3.1%		2	0.23						
	N&hA	83	-1.59	3.6%		154	-1.59	4.5%		4	-2.02		
	S&hA	19	-4.91	73.7%		76	-3.94	21.1%		17	0.50		
	SE	243	-1.59	29.6%		10	-4.32	70.0%					
Feb	NW	78	-0.09	11.5%									
	NE	24	0.75			1	2.44						
	N&hA	84	-1.36	4.8%		153	-2.24	7.2%		4	-3.56		
	S&hA	19	-4.10	10.5%		78	-5.09	15.4%		17	-1.91	5.9%	
	SE	228	-0.63	4.4%		12	-2.50						

Mar	NW	65	-0.33	4.6%								
	NE	20	-0.93	10.0%		1	0.10					
	N&hA	75	-3.10	30.7%		151	-3.94	21.9%	4	-1.91		
	S&hA	18	-3.52	33.3%		73	-7.00	46.6%	17	-2.55	11.8%	5.9%
	SE	212	-0.65	5.2%		12	-3.22	16.7%				
Apr	NW	34	-0.08	23.5%								
	NE	18	-0.33	38.9%		1	-0.73					
	N&hA	69	-1.48	68.1%		133	-5.70	65.4%	4	-7.07	25.0%	
	S&hA	14	-0.92	50.0%		65	-6.63	56.9%	17	-8.28	41.2%	
	SE	136	-0.13	38.2%	0.7%	7	-1.42	14.3%				
May	NE	7	-0.01									
	N&hA	36	-0.03	5.6%		114	-1.42	28.1%	3	-5.69		
	S&hA	9	-0.01	11.1%		41	-2.68	39.0%	15	-9.46	40.0%	
	SE	52	-0.02		1.9%	7	-0.02					

3.4 Interannual variability from 1971 to 2019

Complementing the trend analysis, this section presents an evaluation of the interannual variability of snow depth series. Figure 8 highlights that mean snow depth exhibited a strong interannual variability in the analysed period. Because of the large number of stations, only time series that average over all stations in 500 m elevation bands are shown; however, individual station behaviour was well represented by the 500 m averages, see also auxiliary plots at the repository (Matiu et al., 2020). In the 1970s and 1980s high snow depths were observed, followed by a period of extreme low snow depth in the 1990s. Since the 1990s, snow depths in winter have partly recovered, while in spring snow depths have continued to decline. At the end of the snow season and for lower elevations, average snow depths approached zero, such as in April for 500 to 1000 m or in May for 1000 to 1500 m. The different regions showed similar large-scale patterns, and, for example, the 1990s drop can be seen across the whole Alps. Particular years, especially extreme ones, show concurrent behaviour, for example February 1986 or 2009. Otherwise, there is mixed coherence across regions, as can be also seen from looking at standardized anomalies (Fig. B2) instead of raw snow depth.

These patterns are generally in line with those presented in previous studies, which showed high snow amounts in the 1960s and 1980s and negative anomalies in the 1970s and 1990s, i.e. snow scarce winters, regime shifts or breakpoints in that period in France, Switzerland, Italy, and the western and southern part of Austria, and a recovery afterwards (Durand et al., 2009;

Laternser and Schneebeli, 2003; Mallucci et al., 2019, 2019; Marcolini et al., 2017b; Marty, 2008; Micheletti, 2008; Scherrer et al., 2013; Schöner et al., 2019; Valt and Cianfarra, 2010). In an Alpine wide view, this temporal variability is also accompanied by a strong regional variability.

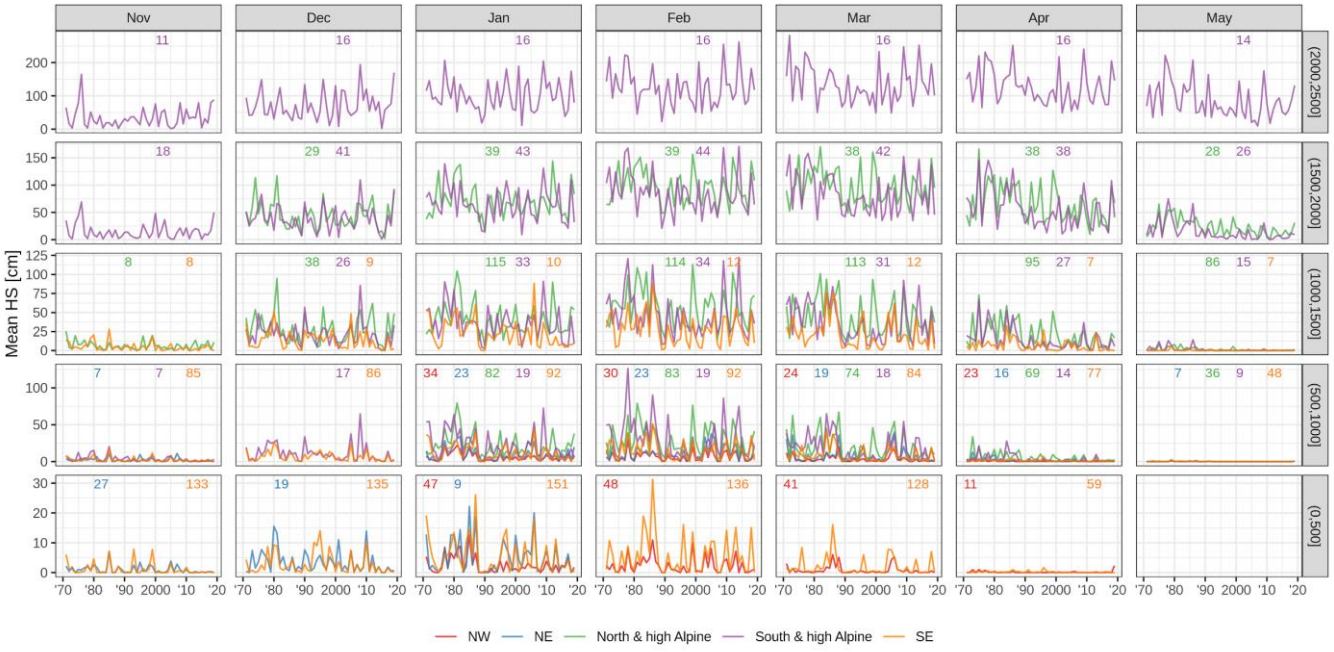


Figure 8: Time series of mean monthly snow depth averaged by 500 m elevation bands. The rows indicate elevation band and the columns the months. The small numbers at the top of each panel denote the number of stations included in the average. Lines are only shown if more than 5 stations were available. Time series of all single stations are available at the repository (Matiu et al., 2020).

In order to put the trends from Sec. 3.3 into context of interannual variability, we examined their relationship by looking at the ratio between the 1971 to 2019 trend and the standard deviation of residuals (Fig. B3(a)). This gives an indication of the relative contribution of the trend to interannual variability. The highest ratios were observed in November to January below 1000 m, in March between 500 and 2000 m, in April between 0 and 2500 m, and in May between 1500 and 2500 m.

As expected from the high temporal variability of the snow depth series, the fraction of explained variance from the linear trends was low. The average R^2 over models with significant trends ($p<0.05$) was 10%. However, R^2 increased with elevation and in the last months of the snow season, reaching up to 32%.

From Figure 8 a decrease in the variability of the snow depth series can be observed, especially at the end of the season and for lower elevations. This is confirmed by the large fraction of negative time coefficients for the error variance in April and May (Table B2), where approximately 40-80% of the stations presented significantly decreasing variability, depending on the region. Notable decreases in variability were also observed in November and in January for NE, NW, and SE. Considerable

significant increases of variability, on the other hand, were only observed in December for 27% of the South & high Alpine series.

3.5 Seasonal snow indices of snow depth and snow cover duration

In addition to the analysis of monthly mean snow depth from Sec. 3.3 and 3.4, this section gives a summary of trends in seasonal indices of mean and maximum snow depth as well as snow cover duration (Table 3, Appendix C). The results of seasonal mean HS agree with the monthly analysis and show generally decreasing snow depths in winter up to 2000 m and in spring for all elevations. Maximum snow depth across the whole season (November to May) decreased stronger than mean snow depth, e.g. the average trend of mean HS for stations in the north (N&hA, NE, NW) between 1000 and 2000 m was -2.8 cm per decade and -5.2 cm per decade for maximum HS, which corresponds to -6.2 and -4.2 % per decade, respectively. Again, stations south (S&hA, SE) had more negative trends: e.g. -4.1 cm per decade for mean HS and -9.8 cm per decade for maximum HS for the same elevations (1000 to 2000 m), which corresponds to -8.9 and -7.1 % per decade, respectively. Average relative trends below 1000 m were more negative than average trends between 1000 and 2000 m for meanHS (DJF and NDJFMAM) and all SCD indices, but not that obviously for meanHS in MAM and maxHS.

Seasonal SCD also decreased for almost all stations below 2000 m, while above no consistent or significant changes were observed. The average trend in November to May SCD over all stations below 1000 m was -4.5 days per decade in the north and -4.8 in the south, and over all stations between 1000 and 2000 m, -5.3 in the north and -7.0 in the south, respectively. The fact that above 2000 m no changes in SCD were observed might also be caused by our season definition (November to May), which is not always enough to capture the full season above 2000 m. In terms of relative changes, mean HS decreased stronger than maximum HS in our study, which is consistent with previous findings (Bach et al., 2018). However, in terms of absolute trends, the opposite was true for our study: mean HS decreased less than maximum HS. Another potential explanation for these differences might be the fact that the study period in Bach et al. (2018) starts 20 years earlier than in our study and their maxHS trends are influences by some extreme events at the start of their study period.

In addition, our changes per decade for maximum HS and SCD are clearly smaller than the ones found by Klein et al. (2016) for a similar time period but a small number of stations in Switzerland. We were able to reproduce the exact estimates from Klein et al. (2016) for the same sites, and hereby found that the differences were caused mostly by the different period (1970-2015 versus 1971-2019). Which makes sense, since 1970 was a snow abundant year, as were the years after 2015 compared to before 2015. In the case of SCD, the different season length (Klein et al. used the whole year, and we only November to May) also had an impact, especially for the higher elevation sites. This supports our introductory statement on the challenge of synthesizing different studies and on the requirement of a unified analysis.

Table 3: Summary of 1971 to 2019 trends in seasonal snow indices. The five regions were collapsed into two (north and south). The number of stations differs by season and the range of available series is indicated in the third column. Average trends (with minimum, maximum in parentheses) are given for seasonal indices of mean snow depth (meanHS), maximum snow depth (maxHS), and snow cover duration (SCD). The season is indicated in the second row with the first letter of the included months (e.g. NDFJ is November, December, January, and February). Absolute trends are in cm per decade for meanHS and maxHS, and in days decade^{-1} for SCD. Relative trends are expressed as % per decade (a few stations south below 1000m were removed, because their low and insignificant snow amounts caused unlikely high relative trends).

Elevation [m]	Region	# series	meanHS	meanHS	meanHS	maxHS	SCD	SCD	SCD
		(range)	DJF	MAM	NDJFMAM	NDJFMAM	NDJF	MAM	NDJFMAM
Absolute changes			*cm per decade*				*days per decade*		
(0,1000]	North	141-190	-0.9 (-5.3, 1.0)	-0.8 (-6.4, 0.1)	-0.8 (-4.7, 0.4)	-2.4 (-11.2, 3.1)	-2.8 (-11.5, 2.7)	-1.7 (-5.6, 0.1)	-4.5 (-13.6, 2.9)
	South	224-241	-1.2 (-6.0, 0.9)	-0.3 (-3.2, 0.3)	-0.7 (-3.6, 0.2)	-3.2 (-15.3, 3.1)	-3.7 (-10.7, 1.4)	-1.1 (-5.5, 1.2)	-4.8 (-14.6, 0.6)
(1000,2000]	North	122-155	-2.1 (-11.0, 3.1)	-3.7 (-21.9, 0.8)	-2.8 (-15.6, 1.6)	-5.2 (-19.9, 3.0)	-2.1 (-8.0, 5.0)	-3.0 (-7.5, 0.7)	-5.3 (-13.9, 0.7)
	South	61-84	-3.5 (-12.6, 2.3)	-4.9 (-18.7, -0.3)	-4.1 (-14.0, 1.6)	-9.8 (-29.2, 2.6)	-2.5 (-7.3, 1.7)	-4.1 (-8.3, 1.2)	-7.0 (-13.9, -0.2)
(2000,3000]	North	3-4	-4.3 (-9.9, -2.2)	-4.5 (-5.2, -4.1)	-5.0 (-8.2, -3.3)	-8.1 (-15.8, -4.2)	0.1 (-0.1, 0.2)		0.1 (-0.1, 0.2)
	South	16-17	-0.1 (-9.2, 11.3)	-6.7 (-18.2, 6.6)	-2.9 (-11.5, 6.8)	-9.4 (-29.2, 6.1)	-0.2 (-2.1, 1.8)	-0.6 (-4.5, 1.9)	-1.0 (-4.7, 1.7)
Relative changes			*% per decade*						
(0,1000]	North	141-190	-7.2 (-20.4, 12.1)	-11.2 (-20.6, 18.0)	-8.7 (-20.4, 10.0)	-4.7 (-19.4, 8.6)	-5.2 (-18.1, 7.8)	-9.7 (-28.5, 9.7)	-6.1 (-16.6, 6.9)
	South	220-238	-8.7 (-18.6, 22.7)	-7.5 (-21.7, 28.0)	-10.0 (-19.0, 10.6)	-6.8 (-16.7, 10.4)	-6.8 (-14.7, 8.2)	-8.0 (-19.5, 15.0)	-7.3 (-14.3, 4.4)
(1000,2000]	North	122-155	-3.6 (-17.6, 23.3)	-9.5 (-20.0, 3.1)	-6.2 (-18.3, 5.1)	-4.2 (-13.8, 3.9)	-2.0 (-8.5, 10.6)	-6.0 (-18.3, 4.0)	-3.5 (-11.8, 0.7)
	South	61-84	-6.5 (-14.1, 5.1)	-11.4 (-17.2, -1.0)	-8.9 (-14.8, 4.9)	-7.1 (-12.0, 3.0)	-2.6 (-8.8, 1.9)	-7.8 (-16.8, 4.2)	-4.7 (-10.7, -0.1)
(2000,3000]	North	3-4	-2.5 (-4.0, -1.5)	-1.8 (-2.0, -1.4)	-2.4 (-2.9, -2.1)	-2.1 (-3.1, -1.6)	0.1 (-0.1, 0.1)		0.0 (-0.0, 0.1)
	South	16-17	0.2 (-8.0, 11.7)	-4.2 (-11.9, 13.2)	-2.4 (-10.1, 6.7)	-3.4 (-9.1, 4.6)	-0.2 (-1.8, 1.6)	-0.7 (-5.3, 2.4)	-0.6 (-2.4, 0.8)

3.6 Representativeness of the stations in an Alpine wide context

Since we aimed to give an Alpine wide assessment, the horizontal and elevational coverage of the station observations is crucial in determining the confidence in the results. For this, we compared the elevation distribution of our station set with a digital elevation model (DEM) at 1 km resolution for the area spanned by the stations (Fig. 2(c)). In relative terms, the elevations of the stations used in this study oversampled the elevations up to 1000 m, were similar from 1000 to 2000 m, significantly underrepresented 2000 to 3000 m, and did not cover elevations above 3000 m.

If the absolute number of stations used in this study is deemed sufficient to describe the spatial coverage, then the confidence of statements would be high for elevations up to 2000 m, while between 2000 and 3000 m, the results should be taken more cautiously. While the elevations above 3000 m only cover a minimal area (0.7% of the area studied here, see Fig. 2(c)), they store large amounts of snow: Figure 6(a) gives an indication of the expected increase in HS with elevation. Long-term monitoring is extremely challenging at elevations above 3000 m, and the snow cover at these elevations is relevant for hydrology, mountain ecosystems, glacier dynamics and mountain (ski) tourism.

Spatial variability of snow increases with elevation (see also Fig. 6), and thus the absolute number of stations required for comparative assessments would be even higher for high elevations compared to low elevations. This limitation could be tackled with automatic snow depth sensors, which better sample high elevations; however, their historical time series are yet too short for assessing long-term trends, besides their issue of harmonized data processing (see also Sec. 1).

An alternative method to derive spatially representative results is to transform the station point observations into a gridded product, by e.g. deterministic or geostatistical interpolation (i.e., kriging). However, in the complex topography of the European Alps with strong elevation gradients, it is challenging to determine an appropriate horizontal resolution that represents elevation well. Moreover, a high enough station density would be needed to perform interpolation. An observation-based grid of snow depth for the Alps would have many potential use cases, from hydrological applications to evaluation of remote sensing and climate models, but it is beyond the scope of this study.

3.7 Outlook

The scope of this study was primarily the detection of snow depth trends, thereby contributing to better understanding and quantification of the state and evolution of the mountain cryosphere in the European Alps (Beniston et al., 2018; Hock et al., 2019). The formal attribution of the trends to climatic drivers, such as temperature and precipitation, as well as the influence of anthropogenic climate change on snow trends (Najafi et al., 2017; Pierce et al., 2008) is not explicitly addressed, although the collation of this unique dataset allows the scientific community developing such studies in the future. The correlational analysis from Sec. 3.2 suggests that temperature and precipitation are important drivers of temporal and spatial variability of snow depth across the whole Alps.

Besides snow depth, also observations of the depth of snowfall (HN) were collected for this study, which were used partly for quality checking the snow depth series. However, analysing the HN series and comparing results to those obtained for snow depth would have exceeded the scope of this study. In the future, we plan to continue with the analysis of HN series, for which we are also aware of other data sources, in particular for low elevation sites in Italy (Pifferetti et al., 2017).

4 Conclusions

We presented the first Alpine wide assessment of snow depth trends based on in-situ measurements in the European Alps. This enabled the identification of five distinct snow regions, whose spatial gradients are related to the known diverse climatic influences for the Alps.

The trend analysis, based on measurements from 1971 to 2019, highlighted the overall reduction in snow cover. Decreases in monthly mean snow depth from November to May were observed for 85% of the station-month combinations (of which 26% significant, $p < 0.05$), while only 15% showed increases (of which <1% significant). Stronger negative trends with higher significance were observed in spring, and in the case of low elevations during the whole season (Table 2). These are the times and elevations, where the transition from snow to snow-free occurs. The observed changes are thus consistent with the

expectations from the snow-albedo-feedback (Thackeray et al., 2019) and highlight its importance for mountain climates (Pepin et al., 2015). Seasonal maximum snow depth decreased stronger than seasonal mean snow depth in absolute terms, while in relative terms the opposite was true (Table 3). Snow cover duration decreased below 2000 m, while above no consistent change was observed, partly due to our choice of snow season (November to May).

The different regions showed good agreement of the inter-annual variability for snow cover duration indices (Fig. C6) and less for snow depth variables (Fig. C4). The magnitude of trends differed by region and the decreases in the south were on average stronger than in the north (Table 2, Table 3). Combined with the lower snow depths south than north (Fig. 6, Table S1 and S2), this resulted in an even stronger relative decrease south than north. The number of stations analysed here gives high confidence to the changes up to 2000 m, while above this elevation the changes have to be interpreted more carefully, especially in the north, where only few stations were available.

The orography of the Alps clearly manifests as the main impact on the snow climatology. It defines boundaries for subregions in north versus south, followed by west versus east. The location of a station with respect to the climatic forcing zones defines the snow depth climatology and impacts the variability of snow depth at a daily scale. Additionally, it can result in different trend magnitudes and also trend signs. Besides these larger scale features, substantial variability exists at higher elevations within the estimated snow depth regions. In summary, the assumption that results from one region are valid in another or for the whole European Alps needs to be evaluated cautiously.

This study provides a clear and harmonized picture for the detection of observed snow depth trends across the European Alps. Thereby it contributes to bridge a scientific gap, which exists for many mountain areas in the world (Hock et al., 2019). We anticipate that the dataset developed for this study, from which a large part is made available to the broader scientific community, will provide support for further studies. In particular to formal attribution studies, which quantify the anthropogenic component in the physical drivers of change, and which remain extremely limited regarding snow cover trends (Najafi et al., 2017; Pierce et al., 2008).

A large community effort and open data sharing for research purposes has made this study possible. We have shown the benefits of a data set that spans many nations and institutions. We expect this dataset to be used for further studies addressing various sectoral applications or for the evaluation of remote sensing or reanalysis products. Perhaps it might be expanded in the future thanks to additional contributing organizations. However, we currently lack the opportunity to have a continuously updated version. With ECA&D (European Climate Assessment & Dataset), a harmonized station data collection portal exists at the European scale for many meteorological variables. But while the coverage of e.g. temperature and precipitation is balanced across Europe, snow depth is only limited or not at all available for many European mountain regions, such as the European Alps, Carpathians, Balkan Mountains, or Dinaric Alps. It would be desirable to have an updated harmonized station dataset for snow cover, given its importance in mountains and further downstream. This would enable a better monitoring of the changes, their consequences and impacts, and contribute more quantitatively to climate change, ecosystems and environmental assessments than is possible at the moment. However, such an endeavour requires a more formal umbrella and long-term commitment, e.g. in the framework of the Copernicus Earth monitoring programme of the European Commission.

650

Appendices

Appendix A: Data Processing

After collecting the data, the series from different data providers were harmonized and put into a common data format. This included converting all station coordinates into latitude and longitude. In a few cases, where only station name and elevation were available but no coordinates, the missing coordinates were extracted from Google Maps using the approximate location (with correct elevation) based on the station name. Most data providers used station identifiers along with station names. We chose to have unique identifiers for all stations based on the station name. Station names were standardized by replacing blanks and apostrophes with underscores, and by removing accents. If multiple stations had the same name within one data source i.e. by data provider, the names were suffixed with the station identifier from the data provider. If multiple stations had the same name across data providers, the names were suffixed with the data provider identifier.

A.1 Merging of records

The final database included several cases in which snow measurements for the same location were stored as separate records since they covered different periods and/or a slight relocation of the same station site occurred. In some cases, different records were available at very close locations where snow data were collected at the same time or over partially overlapping periods for different operative or research purposes. In order to maximize the temporal continuity and extent of available HS and HN series, the records referring to the same site or to very close locations were merged: one series was created from the multiple series by replacing missing values or missing periods. In particular, the merging was performed only if the sites were closer than 3 km and their vertical distance was less than 200 m. In the case of overlapping periods, the data from the series with the fewest gaps was retained. The merging was evaluated and performed on HS series first. In the case that HN series for the same sites were also available, the data were merged by following the same criteria used for HS, in order to preserve consistency between HS and HN measurements. The metadata of the most recent series included in the merging was assigned to the resulting record. About 60 merged series were obtained in total and the duplicated records for the same site were discarded.

A.2 Quality control

The series were quality checked in order to remove recording errors. First, below zero HS or HN values were replaced with missing values. Then a temporal consistency check was applied to HS to identify recording errors. Series were screened for jumps larger than 50 cm (up and down in two consecutive days; or vice versa). This criterion identified 680 values from the daily observations from all series, which were checked manually, and recording errors were replaced with missing values. Another issue with HS series is that missing observations might falsely be recorded as 0 cm. To identify suspicious series, mean winter (December to February) HS and the fraction of 0 cm values were calculated per station. Then, looking at a surrounding elevation band per station (200 to 500 m, depending on the elevation and station availability), series were marked

if the mean HS was less than the 5th percentile or the fraction of 0 cm values was higher than the 95th percentile of all stations in the elevation band. Given the climatological nature of this pre-screening and the stronger dependence on elevation, we did not consider horizontal distance for this step. This resulted in 181 suspicious series, which were checked manually. For 32 stations, there were periods where 0 cm were obviously missing values, and in these periods the 0 cm values were replaced with missing values; the remaining 149 stations had no missing values denoted as 0 cm. Finally, during all previous manual checks, series that showed "dubious" behaviour were marked, which were in total 48 series. Dubious behaviour was e.g. inconsistency between HN and HS, unlikely values, improbable temporal variability, multiple seasons with no snow, or excessive gaps. From these 48 series, 29 were considered usable, 11 had some periods removed, and 8 were completely removed.

These procedures could identify some errors, but definitely not all. Because of the large number of series, it was not feasible to manually quality check all of them, and fully automatic checks are often not feasible. Instead a spatial consistency check was applied (see Appendix A.4), and the rest of remaining errors could be considered noise given the large amount of data.

A.3 Gap filling

Most series contained gaps ranging from some days up to whole seasons. In order to conduct climatological or trend analyses, gaps in the series needed to be filled. For this we employed a spatial interpolation approach, similar to the one used for temperature and precipitation records (see e.g. Brunetti et al., 2006; Crespi et al., 2018; Golzio et al., 2018). The approach is based on correlations between the series, and because snow strongly depends on elevation, we first performed a spatial analysis to identify which correlations can be expected depending on horizontal and vertical distances between stations. For this, pairwise correlations (Pearson) between the daily HS series were performed for December to April from 1981 to 2010, only if the series had at least 70% valid data, and only if each pair had at least 50% of data in common. As expected, correlations decreased with both horizontal and vertical distance (Fig. A1). But correlations remained high even for large distances, e.g. correlations higher than 0.7 were found up to vertical distances of 500 m (with less than 100 km horizontal distance) or up to horizontal distances of 200 km (with less than 250 m vertical distance). It should be noted that correlations can be high even if there are large differences in amounts or ratios between the series, as long as the differences and ratios are constant across the range of values.

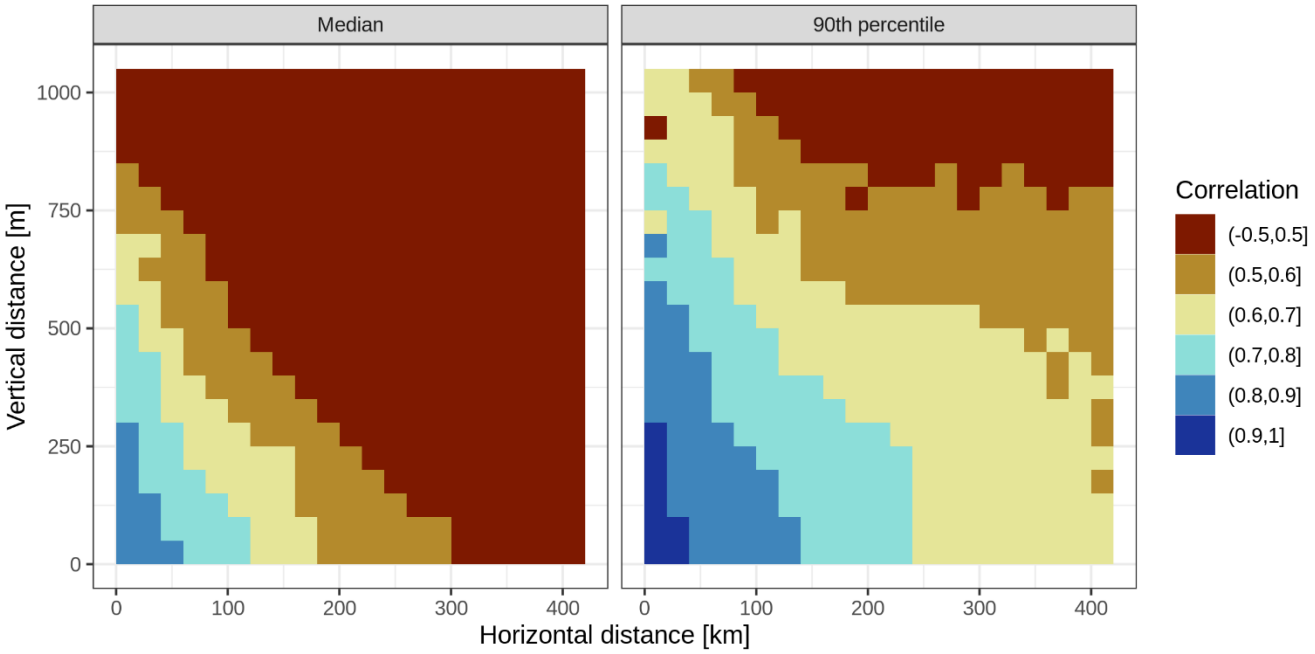


Figure A1: Summary of pairwise correlations between HS series for December to April, 1981 to 2010. Shown is the average (median, left) and 90th percentile (right) of all pairwise correlations in bins of 20 km horizontal distance by bins of 50 m vertical distance. The correlations were only calculated if each series had at least 70% valid data in the period and if each pair had at least 50% of data in common.

The chosen approach fills a gap based on finding highly correlated neighbouring series to the one with gaps. The gap filling algorithm works as follows. For each gap:

1. Find temporally surrounding non-missing values in the gap series around the gap date ("window data"), see also Fig A2 (a).
    1.1. Take 15 days before and after the gap. This results in 31 days of the year: e.g. for Jan 15, this would be Jan 01 to Jan 31; for Jan 01, this would be Dec 16 to Jan 16.
    1.2. Repeat step 1.1. for 10 years before and after the gap. This results in 21 years. E.g. for 1996, this would be 1986 to 2006.
    1.3. This window data potentially contains 651 values (21*31), but likely has missing values. If there are more than 150 non-missing values continue to step 2. If there are less than 150 non-missing values, increase the day window by 5 days in both directions, and repeat from 1.1. If the day window has reached 45 in one direction (i.e. total 91 days), and still there are less than 150 non-missing dates, stop. Note: only the day window is increased, the year window from 1.2. stays constant at 10 years before and after.
2. Pre-select potential reference series (Fig. A2 (b)) based on the following criteria: vertical distance to gap series below 500 m, horizontal distance below 200 km, and the value at the date of the gap is not missing.

3. For each potential reference series:
    3.1. Identify dates with values available for both gap and reference series in the window identified in step 1. (Fig. A2 (c)). Continue only if more than 80% of the minimum 150 non-missing values (i.e. 120) are available in common.
    3.2. For the common dates: calculate mean of gap series and mean of reference series; calculate correlation between gap series and reference series. If all values of gap and reference series are zero, set the correlation to the minimum threshold (see step 4.) plus 0.001 (in order to be able to fill also zero periods). If only one of the series has all zero values, i.e. either gap or reference but not both, set the correlation to zero.
    3.3. Calculate ratio between mean of gap series divided by mean of reference series. If the mean of the reference series (divisor) is zero, set the ratio to zero (in order to be able to fill also zero periods).
4. Sort potential reference series by correlation with gap series (from step 3.1.). Remove all candidates with a correlation below 0.7. This threshold was chosen as it is used e.g. in the homogenization of snow depth (Marcolini et al., 2017a).
5. Select the first 5 best correlated reference series, or up to 5, depending on how many available.
6. Calculate weights based on vertical distance. The weights are based on exponential decay with a halving distance of 250 m ("half-time" transformation of decay constant). This implies that the weights are halved every 250 m.
7. Fill the gap value with a weighted (step 6.) average of the reference series values adjusted by the ratios between gap and reference series (step 3.3.): $HS_t^{gap} = \frac{1}{n}\sum_{i=1}^{n} w_i * HS_t^{ref_i} * \frac{HS_{mean}^{gap}}{HS_{mean}^{ref_i}}$, where t is the date of gap, i is the index of reference series, n is the number of reference series 1...5, and w_i are the weights with $\sum_i w_i = 1$.

The filled value was rounded to the nearest integer value in cm. Since the method requires finding suitable reference stations, it was only performed for the period 1961 to 2020, because the station density was too low before. The gap filling was applied to all gaps in all series considering all available data; afterwards thresholds were applied to select usable series (see end of this section).

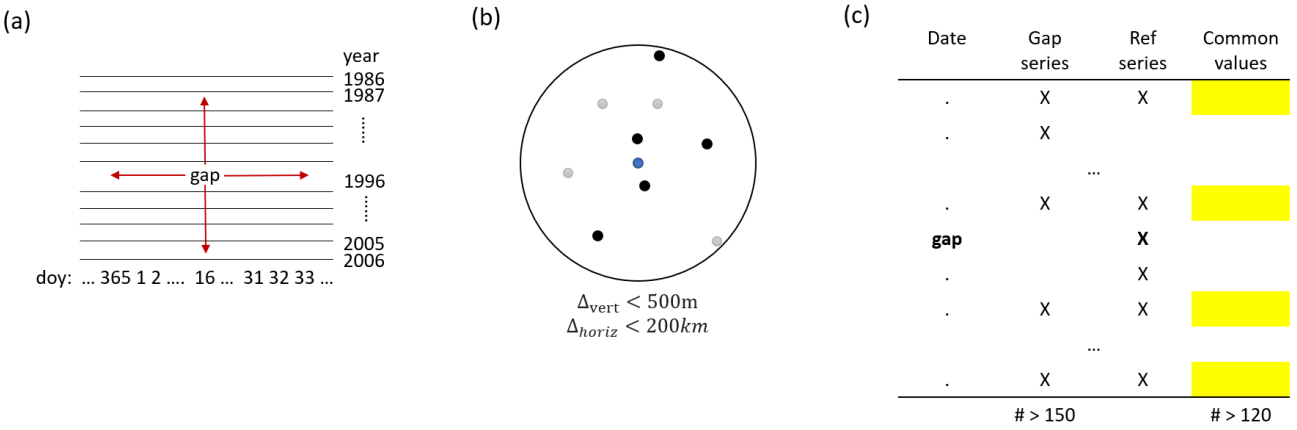


Figure A2: Visualization of some steps of the gap filling algorithm. (a) shows how the window data in the gap series around the gap is determined (step 1.); doy is day-of-the-year. (b) shows the selection of potential reference series by horizontal and vertical distance (step 2.). (c) shows how common dates for gap and reference series are identified (step 3.); the dates come from the window in (a).

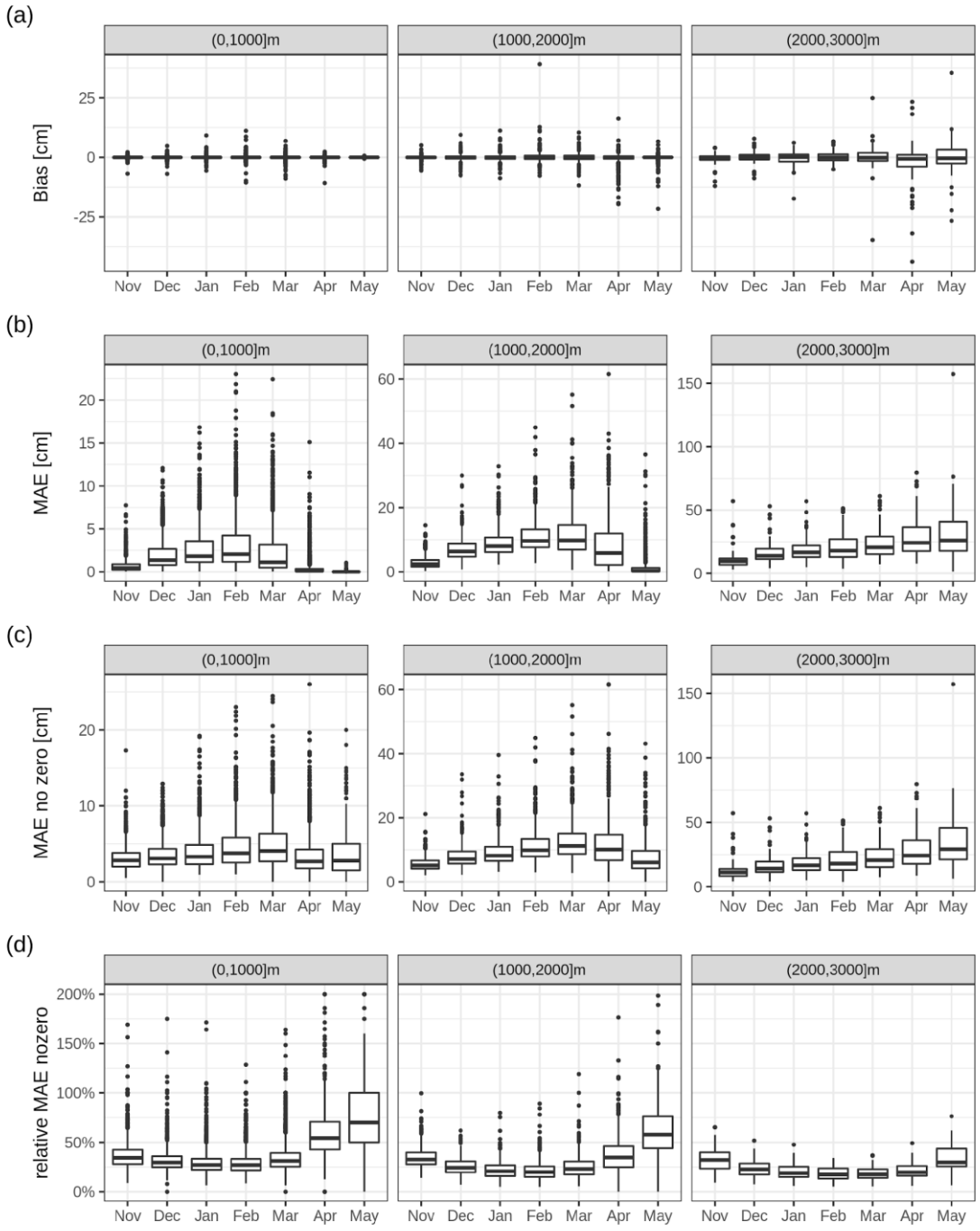
The chosen limits of 200 km horizontal distance and 500 m vertical distance might seem very high in the Alpine context with the complex topography. Since we were interested in larger scale snow patterns and not local snow peculiarities, such large distances are justified. Moreover, the correlation threshold should exert control on selecting only stations that share the same snow cover evolution, and high correlations were found up to these horizontal and vertical limits (Fig. A1). On the other hand, a nearby station might also be a worse predictor than a more distant one, if, e.g., it differs in its local climate.

Since this gap filling approach has not yet been used for snow depth, we performed a cross-validation analysis to identify the gap filling errors. For this, we used data from November to May in the period 1981 to 2010. For each station and each year, one month at a time was held out, but only if at least 10 days were available. Thus, for each month, a maximum potential of ~900 values were cross-validated; however, the effective number was lower, because of missing values, and because not all gaps could be filled, if no suitable reference stations were available. In order to test the effect of shorter period gaps, we also applied the cross-validation on subsets (to reduce computation time): 1) 100 random samples of 1 day and 2) 20 random samples of 5 consecutive days. Then, the held-out values were filled using the above approach, and metrics calculated based on the filled and held-out values. Metrics include the bias, the MAE (mean absolute error), the MAE for non-zero held-out values only, and a modified version of relative MAE. The relative MAE is based on the MAE for non-zero values only, and this non-zero MAE is divided by the average of the held-out non-zero values. This is then not a "true" relative error, which would divide each error by the true value, i.e. $\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - x_i}{x_i}\right|$, but our modification is $\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - x_i|}{|\underline{x}|}$, where $\underline{x}$ is the average of all x_i. This was done to remove the large influence of errors close to zero, which are not that relevant in this case. The metrics were only calculated if more than 50 values were available per month and station (out of potentially ~900 for the month-long gaps, and 100 for the 1 and 5 day gaps), in order to provide robust estimates.

The cross-validation showed that the gap filling has extremely little bias (Table A1) with the overall average daily bias for the month long gaps being -0.04 cm. Average daily MAE for filling whole months was 1.6 cm (averaged over stations located in 0–1000 m), 7.7 cm (1000–2000 m), and 22.0 cm (2000–3000 m). MAE was lower for 1 and 5 day long gaps compared to month long gaps, but almost no differences were observed comparing 1 day or 5 days, e.g. for the 1000–2000 m band, MAE for 1 day gaps was 6.2 cm, for 5 day gaps 6.4 cm, compared to 7.7 cm for 1 month gaps. The relative MAE of month-long gaps decreased with elevation from 39.4% (0–1000 m) to 32.7% (1000–2000 m) to 22.8% (2000–3000 m). Additionally, there was also a seasonal dependence of MAE, while bias remained largely constant across the season (Fig. A3). MAE below 2000 m peaked in February, while above 2000 m MAE increased throughout the season. Relative MAE decreased with higher snow depths, both temporally and with elevation, that is, relative MAE was lowest in February and at high elevations. It is to be expected that errors at the end of the season are related to the ablation scheme (i.e. local climatic and topographic characteristics that influence ablation) of the different stations; however, at this stage we did not check this issue further.

Table A1: Cross-validation (CV) metrics for the gap filling approach: Bias (the difference between gap filled and observed values), the mean absolute error (MAE), mean absolute error only for non–zero observed values (MAE no zero), and MAE no zero divided by the average of all true non–zero values (Rel. MAE no zero).

Elevation band [m]	CV period	Bias [cm]	MAE [cm]	MAE no zero [cm]	Rel. MAE no zero
(0,1000]	1 day	-0.0	1.3	3.1	30.1%
	5 days	-0.0	1.4	3.3	34.0%
	1 month	-0.0	1.6	3.9	39.4%
(1000,2000]	1 day	-0.1	6.2	7.9	26.1%
	5 days	-0.1	6.4	8.2	28.5%
	1 month	-0.1	7.7	9.7	32.7%
(2000,3000]	1 day	-0.6	18.2	18.6	18.9%
	5 days	-0.8	18.3	18.7	19.2%
	1 month	-0.4	22.0	22.5	22.8%



800 **Figure A3: Cross-validation metrics for the gap filling approach: (a) bias, (b) mean absolute error (MAE), (c) mean absolute error for non-zero values (MAE no zero), (d) non-zero MAE divided by the true non-zero mean (relative MAE no zero). Panels show the 1000 m elevation bands indicated in the title. The boxplots represent statistical quantities: the box indicates the first and third quartile; the bold line inside the box is the median; the vertical lines outside the box extend up to the most extreme point but at most**

1.5 times the interquartile-range (IQR; height of the box); finally, points below/above 1.5*IQR of the first/third quartile are shown as separate points.

Moreover, we compared our proposed gap filling approach to results from gap filling snow depth series using simulations of the Crocus snow model for the French Alps. The Crocus simulations with meteorological forcing were performed independently of this study, but we found it useful to compare the two approaches - albeit only exploratively. The observed snow depths with gaps were assimilated into the Crocus modelling scheme, using SAFRAN reanalysis data as forcing (López-Moreno et al., 2020). The two gap filling approaches were compared only for existing gaps in the French Alps. This was intended as a preliminary companion evaluation, and no cross-validation was performed. Thus, there was no ground truth to evaluate the two gap filling approaches with formal metrics, and we only performed a visual assessment (figures for comparison available at Matiu et al., 2020). Time series of both gap filling procedures looked remarkably similar, even for reconstructions of complete missing seasons: the different snowfall events were visible in both and snow depths averaged over multiple days were comparable. Differences emerged in the snow settling behaviour and for the spring snow melting periods. More information on this exercise is available from the authors on request.

For Switzerland, a comparison of gap filling methods for HS was performed, which aimed at reconstructing complete missing seasons, and which included regression based methods and snow models (Aschauer et al., 2020). While our proposed method was not explicitly used in that comparison, it can be assumed to be similar to the regression-based and distance weighted methods used there. The errors reported in their study (root mean squared error less than 20 cm) are in the same order of magnitude as those found in our cross-validation.

Altogether, the abovementioned (the cross-validation results, the comparison to Crocus, and the preliminary findings of the Swiss study) convinced us that the gap filling procedure is also suitable for reconstructing whole seasons, and not only some intermediate gaps, considering the fact that we only used it to derive monthly means (see below) and did not use the daily values directly. Further research would be required to check the suitability of the daily reconstructions, in our opinion, also considering the temporal distance to the last existing observations. For the final analysis, all gap filled data within the recording period was used, and we also allowed extending the period up to five years before the start or after the end of the recordings – but only if the total number of gap filled observations was less than the number of observations without gap filling. The main reason for this extension was to have series covering the complete period until 2019, because some series stopped just a few years earlier. As sensitivity analysis, we repeated most of the statistical analysis also for the original data without gap filling and provided results in the supplementary material: the estimated modes of variability matched (Fig. S12); the magnitude and variability of monthly trends was similar, although a lot less stations were available (Fig. S10); and finally the time series of 500 m average HS also showed similar behaviour (Fig. S11). The gap filling was able to significantly increase the temporal availability, but its aim was not to fill all gaps. Gaps were not filled, for example, if no suitable reference station was found or if not enough common data was available.

A.4 Aggregation and spatial consistency

The daily snow depth (HS) values were aggregated to mean monthly HS, if at least 90% of the daily values were available in the respective months after the gap filling (monthly time series plots available at Matiu et al., 2020).

Based on the monthly series, a consistency check was performed (Crespi et al., 2018), which identifies dubious values/series, (but can also identify series with strong local influences on snow depth). Each monthly HS series of the tested station was reconstructed from up to five reference stations by a spatial interpolation approach. The reference series were selected if the monthly record was available and if at least 10 monthly records were in common with the tested station. If more than five neighbours were available, the ones with the highest weights were selected with weights being derived from the horizontal distance and elevation difference, similar to the gap filling procedure described above. Each reference station value was rescaled by the ratio between tested and reference mean HS for the month under reconstruction. Finally, the monthly simulation of the tested series was defined as the median of the up to five rescaled neighbouring values. The comparison between simulated and observed monthly HS series for each station was evaluated by computing bias, MAE, and R^2 (squared correlation) from December to February, in order to avoid unreliable low error values due to zeros in HS records outside of winter.

The mean bias over all stations was -0.3 (min, max: -8.0, 10.9) cm, average MAE was 4.8 (0.1, 61.3) cm, and average R^2 was 0.83 (0.0, 0.98). However, there was a strong elevational dependency, and station metrics deteriorated with elevation (Fig. A4). A semi-automatic approach was considered to look for suspicious series. The following criteria were used to screen stations: bias outside the 95% confidence interval per elevation bands (250 m bands up to 1500 m, then 1500 to 2000 m, and 2000 to 3000 m) or MAE above a manually defined threshold line (see Fig. A4 (b)) or R^2 below 0.5 or simulation not successful because of too many gaps. This yielded 225 stations, which were checked manually by looking at monthly simulated and observed series, and daily series. Only 14 stations were found suspicious and 18 partly suspicious; all 32 series were removed from the statistical analyses. More detailed results and time series comparing simulated with observed snow depths are available as auxiliary material (Matiu et al., 2020).

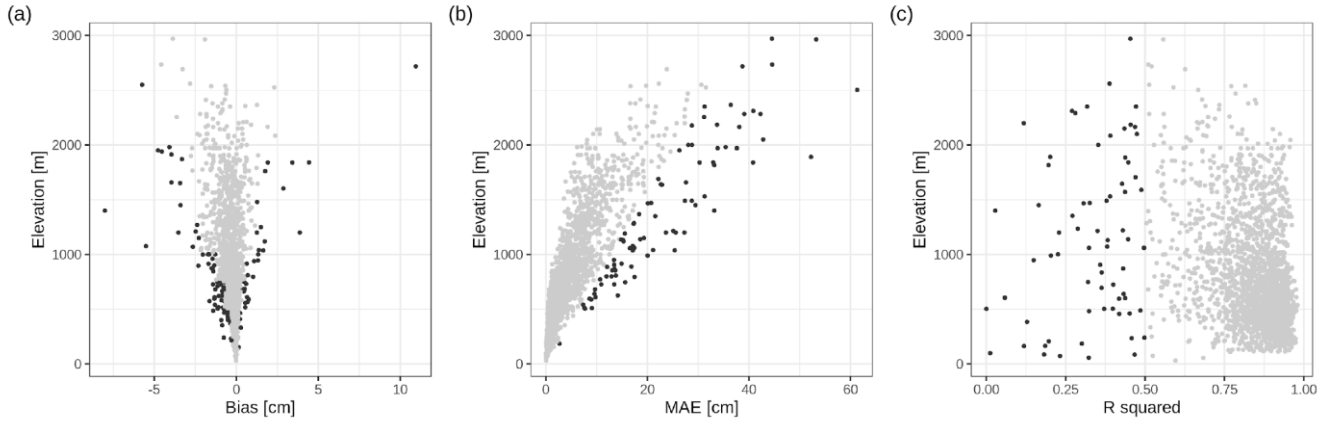


Figure A4: Metrics for spatial consistency: (a) bias, (b) mean absolute error (MAE), and (c) R squared (squared correlation). Metrics were derived from statistical simulations of the monthly series from December to February using spatial neighbours. Black points indicate stations which were further analysed with manual checks.

865

Appendix B: Additional figures and tables

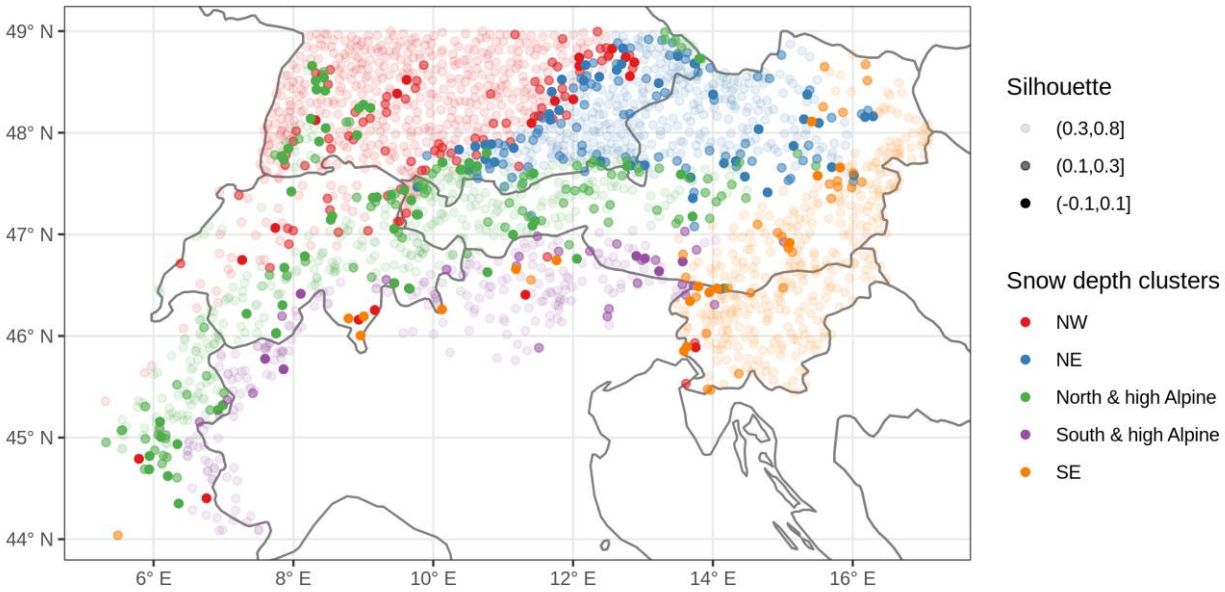
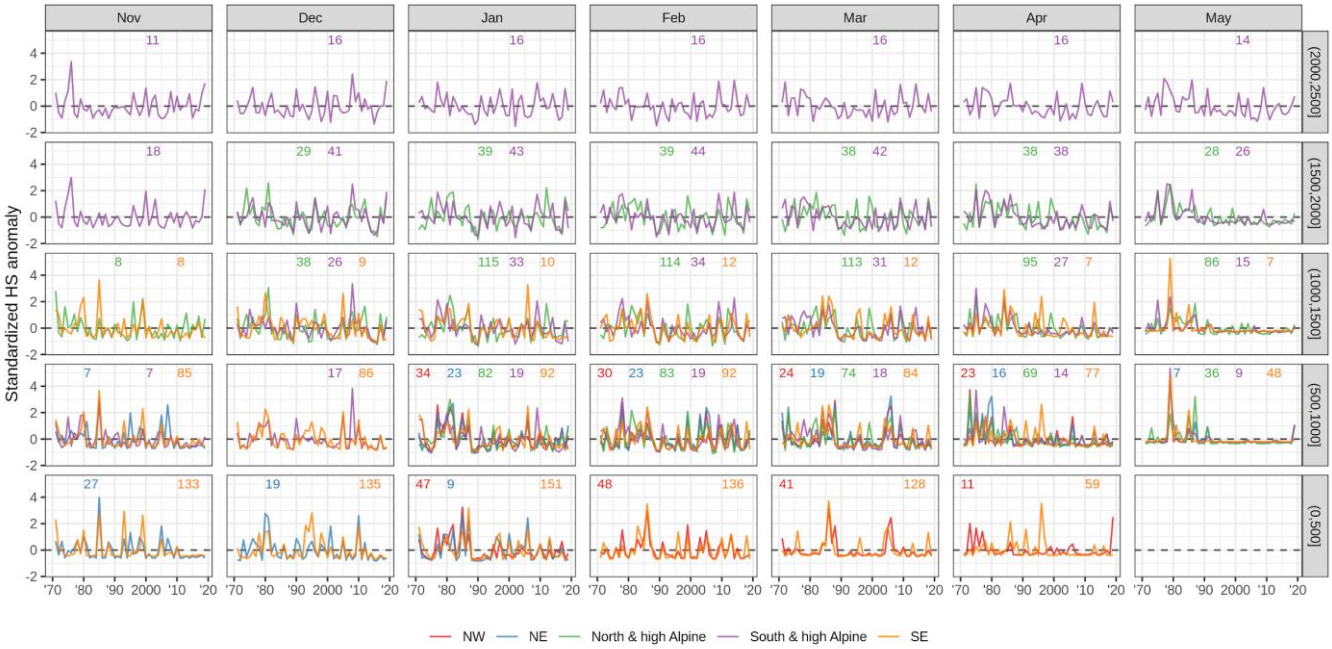


Figure B1: Silhouette values of the stations, which show consistency of clustering. The silhouette is a measure of how similar the station is to its own cluster compared to the other clusters (see methods for formula). High values indicate a good match, while low and negative values indicate a poor match.



875

Figure B2: Same as Figure 8, but using standardized anomalies.

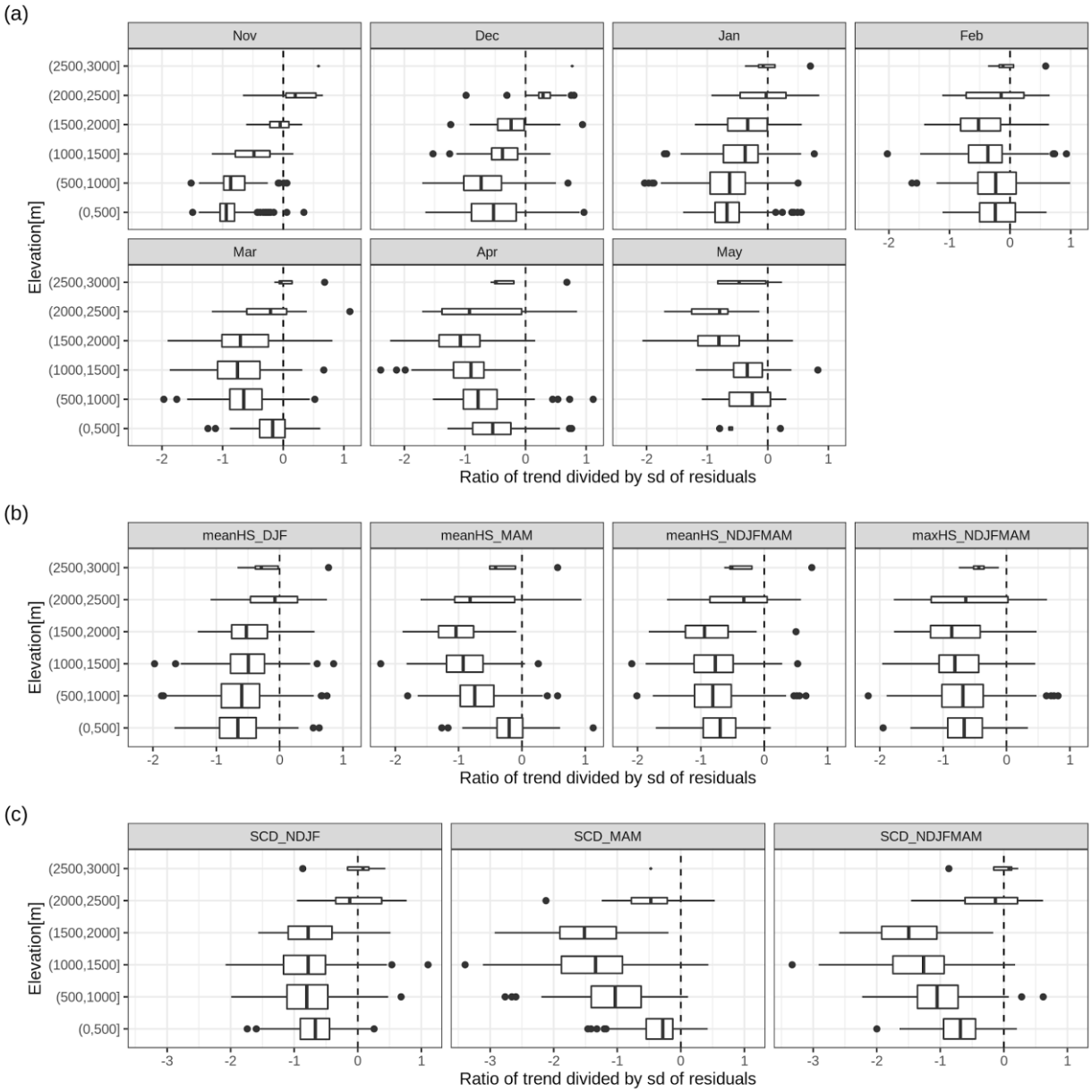


Figure B3: Ratio between the trend over the full period (1971 to 2019) and interannual variability (standard deviation of the residuals). (a) shows the values for monthly mean HS (snow depth), (b) for seasonal indices of HS, and (c) for seasonal indices of SCD (snow cover duration). The boxplots represent statistical quantities: the box indicates the first and third quartile; the bold line inside the box is the median; the vertical lines outside the box extend up to the most extreme point but at most 1.5 times the interquartile-range (IQR; width of the box); finally, points below/above 1.5*IQR of the first/third quartile are shown as separate points. The height of the box is proportional to the number of observations in each group.

Table B1: Overview of literature on snow cover trends in the European Alps.

Area	Number stations	Time period	Season	Snow variable	Methods
(Bach et al., 2018)					
Mean HS -12.2%/10y (40% stn p<0.05); max HS -11.4%/10y (36% stn p<0.05); except coldest climates.					
pan-Europe: mostly Germany, Benelux, AT-Tirol, Czech Republic, Slovakia, Finland; partly UK, Balkan, part E of Baltic Sea	(not specified)	1951-2017	DJF	Mean HS; Max HS (95pctl)	OLS if trend pos.; OLS (exp) if trend neg.; Significance: Mann-Kendall
(Beniston, 2012b)					
10-50% decline in DJF HS (less decline moist north vs. dry south)					
Switzerland	10	1930-2010	DJF; NDJFMA	mean HS, snow days (10cm)	visual; 5yr moving window
(Durand et al., 2009)					
HS no trend at 2700m, decreases below; n0 negative trends					
French Alps	Modelling: ERA-40, SAFRAN, Crocus	1959-2005	DJF	Mean HS n0 (Number of days with snow) HS100d (minimum 100day snow depth)	300m elevation steps (1500-2700); Spearman correlation (year-n0); Step-year (n0); Linear trend (n0)
(Klein et al., 2016)					
SCD shorter 8.9days/decade; more because of earlier snow melt (5.8days/decade); decrease in maxHS, and earlier date of maxHS					
Switzerland	11	1970-2015	Sep-Aug	maxHS; date of snow onset, snowmelt, maxHS; SCD; snow days (1,20,50,100)	Theil-Sen, Mann-Kendall; stepwise regression
(Kreyling and Henry, 2011)					
150 stations showed decrease (p<0.05 for 69), 22 positive (p<0.05 for 1); decrease accelerated over the last 15yr, -0.48 to -0.89d/yr					
Germany	177	1950-2000	Aug-July	SD (1cm)	OLS; random effects with stations
(Laternser and Schneebeli, 2003)					
All variables show increase until 1980, followed by significant decrease. Trends more pronounced at mid and low elevation. South != North. Shorter SCD because earlier melt in spring.					
Swiss alps	140 (HS) 120 (HN)	1931-1999	NDJFMA; two-month splits	Mean seasonal HS; SCD (start, end, length); Days with HN > 0, 10, …; HN3max (max 3 day HN)	Trend analysis; relative to long-term mean; Trend short period equal long period
(Lejeune et al., 2019)					
39cm less 1990-2017 vs. 1960-1990					

France	1 (Col de Porte)	1960-2017	DJFMA	mean HS	moving window (15yr) and comparison 30yr
(Marcolini et al., 2017b) *different dynamics above and below 1650m; larger reductions at lower elevation; strong change late 1980s*					
Italy (BZ+TN)	37	1980-2009	NDJFMA	SCD (>30cm); seasonal HS	homogenization; hovmöller plots; wavelet analysis
(Marty, 2008) *Regime shift at end of 1980s, no clear trend since then.*					
Switzerland	34	~1931-2008	DJFM	Snow days (5, 30, 50cm)	Mann-Kendall; shift detection
(Marty and Blanchet, 2012) *44% of stations show sig decrease in HSmax, 32% for HN3max; decrease in spread of HSmax*					
Switzerland	18 (HSmax) 25 (HN3max)	1931-2010	annual	HSmax (annual max HS); HN3max (annual max sum HN 3day)	GEV with time-dependent location and shape
(Marty et al., 2017) *SWE decline (independent of lat or long); stronger and more significant decrease in spring (-80% to -10% low to high elevation / 60 years) than winter; winter: some pos non-significant at high elevation.*					
alpine wide (AT, FR, DE, IT, CH)	54	1968-2012	index values (spring and winter)	SWE (not continuously measured)	Mann-Kendall; Theil-Sen
(Micheletti, 2008) *pos anomalies until end 80s, then shift to low snow amounts until beginning 2000; some recovery, but still below level of 80s*					
Italy (FVG)	8	1972-2007	seasonal	sumHN, max of monthly meanHS	timeseries (only descriptive); % anomalies w.r.t. 1972-2007
(Scherrer et al., 2013) *strong decadal variability; high values 1900-1920 and 1960-1970/80; lowest values end 1980/1990; increases/plateau 2000s linked to temperature evolution*					
Switzerland	9	1864-2009	annual	MAXNS (max annual HN); NSS (sum annual HN); DWSF (days with snowfall)	plots; 20yr smooth; comparison to 71 other stations
(Schöner et al., 2009) *largest HS in 1940s/50s; summer snow decreasing; interannual variability of winter precipitation closely related to HS (highest in 40s50s, strong decreases since -> less extremes)*					
Austria	1 (Sonnblick)	1928-2005	monthly	HS	(visual)
(Schöner et al., 2019) *EOF groups AT-CH in 7 regions; trend analysis based on first PC; strong trends south ~2000m: up to -12cm/10y; strongest trends at highest elevations; regional dependence of trends*					

Switzerland and Austria	196 (139 passed QC)	1961-2012	NDJFMA	seasonal HS and HN	MK-test with lag1 pre-whitening; running trend approach; Sen slope; EOF for regionalization
(Terzago et al., 2010) *more snow Nov-Dec, less Jan-Apr, disappeared in May*					
Italy (Piemonte)	3	1971-2009	monthly	HN, HS, snowy days (HN>=1cm)	1971-2000 vs 2000-2009
(Terzago et al., 2013) *some maxima 1940,50,60, absolute 1970, absolute minima 1990, then recovery; significant decrease seasonal HS 2-14cm/decade; stronger decreases in North (considering elevation); changes not driven by precip changes; snowfall anticorrelated to NAO*					
Italy (West)	6	1926/1951 - 2010	DJF, MAM, NDJFMAM	precip, days with precip, solid precip fraction, HN, snowy days (HN>0), HS	trend analysis; Mann-kendall; spectral analysis
(Valt et al., 2008) *snow cover decreased 14 days (1991-2007 vs 1960-1990), stronger <1600m (16d) vs >1600m (11d); fresh snow decreased 1990-2000, then stationary (for all altitudes and months)*					
Italy (east and west)	5 (west); 6 (east)	~1920/1960-2007	Oct-May	snow days (>=1cm); sumHN	(visual)
(Valt and Cianfarra, 2010) *NDJFMA CSF shows -3 to -40 cm/10yr for all 18 stations 1960-2009, SCD also all negative; breakpoint ~ 1990, before decrease, after increase; strongest negative trend in spring and below 1500m; neg trend related to precipitation decrease; PCA shows long-term negative trend*					
Italy (east and west)	18	1950-2009	DJFMA; DJF; MA	SCD (>1cm); CSF (sum of new snow)	split by 1500m alt; OLS, Mann-Kendall; changepoint; PCA

Table B2: Fraction of models with significantly positive or negative changes in the error variance by time. The remaining percentage (not shown) corresponds to the total of non-significant negative and positive changes. Empty cells indicate no stations with significant negative or positive trends (sig- and sig+). Changes were considered significant if the GLS model with a time coefficient for the error variance showed significantly improved goodness-of-fit compared to the OLS model with constant error variance ($p < 0.05$).

Region	Nov		Dec		Jan		Feb		Mar		Apr		May	
	sig-	sig+	sig-	sig+	sig-	sig+	sig-	sig+	sig-	sig+	sig-	sig+	sig-	sig+
NW					86.4%		24.4%	5.1%	30.8%	3.1%	76.5%	11.8%		
NE	47.1%	2.9%	16.7%		47.1%	2.9%	4.0%	8.0%	28.6%	4.8%	78.9%		80.0%	
N&hA	53.8%		4.3%		28.9%	0.4%	5.5%	0.4%	22.4%		72.8%		75.3%	4.7%
S&hA	43.9%	4.9%		26.7%	8.0%	8.0%	9.1%	6.4%	14.0%	6.5%	41.1%	1.1%	76.6%	1.6%
SE	72.4%	0.4%	22.6%	2.2%	40.5%	2.4%	18.4%	2.9%	29.1%	2.7%	55.2%	6.3%	100.0%	

Table B3: Overview of shareable data. Column daily indicates if the original daily data can be shared, and monthly if the derived monthly data can be shared.

Code	Country	Data provider	Daily	Monthly
AT_HZB	Austria	HZB	no	yes
CH_METEOSWISS	Switzerland	MeteoSwiss	no	yes
CH_SLF	Switzerland	SLF	no	yes
DE_DWD	Germany	DWD	yes	yes
FR_METEOFRANCE	France	MeteoFrance	yes	yes
IT_BZ	Italy	Bolzano	yes	yes
IT_FVG	Italy	Friuli Venezia Giulia	yes	yes
IT_LOMBARDIA	Italy	Lombardia	yes	yes
IT_PIEMONTE	Italy	Piemonte	no	no
IT_SMI	Italy	SMI	no	no
IT_TN	Italy	Trentino	yes	yes
IT_TN_TUM	Italy	Trentino (TUM)	no	no
IT_VDA_AIBM	Italy	Valle D'Aosta (AIBM)	no	no
IT_VDA_CF	Italy	Valle D'Aosta (CF)	yes	yes
IT_VENETO	Italy	Veneto	no	yes
SI_ARSO	Slovenia	ARSO	no	yes

Appendix C: Seasonal snow indices

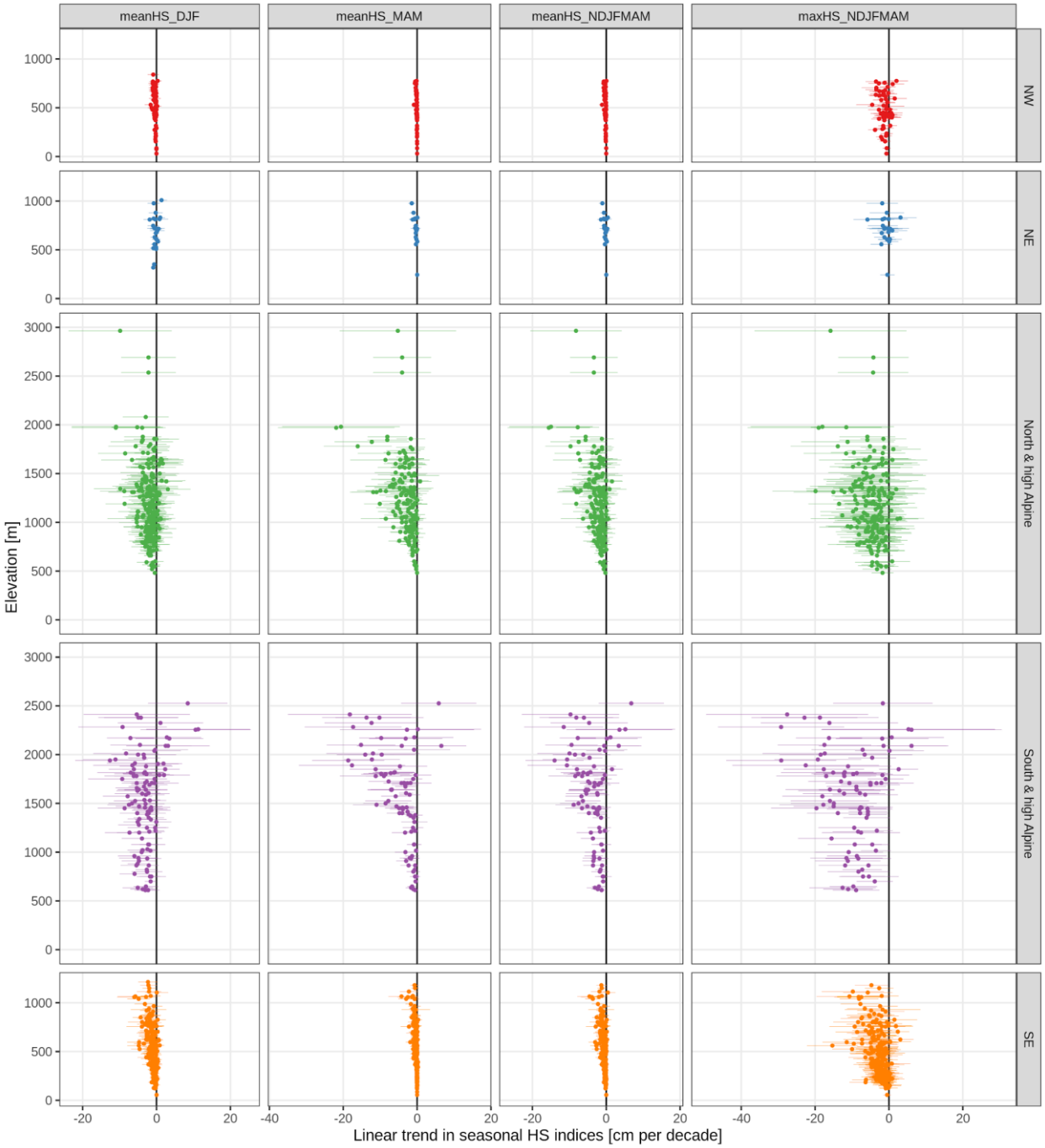


Figure C1: Long-term (1971 to 2019) linear trends in seasonal snow depth (HS) indices. Trends are shown separately by index (columns) and region (rows). The season is indicated in the columns with the first letter of the included months (e.g. DFJ is December, January, and February). Each point is one station. The points indicate the trend and the lines the associated 95% confidence interval.

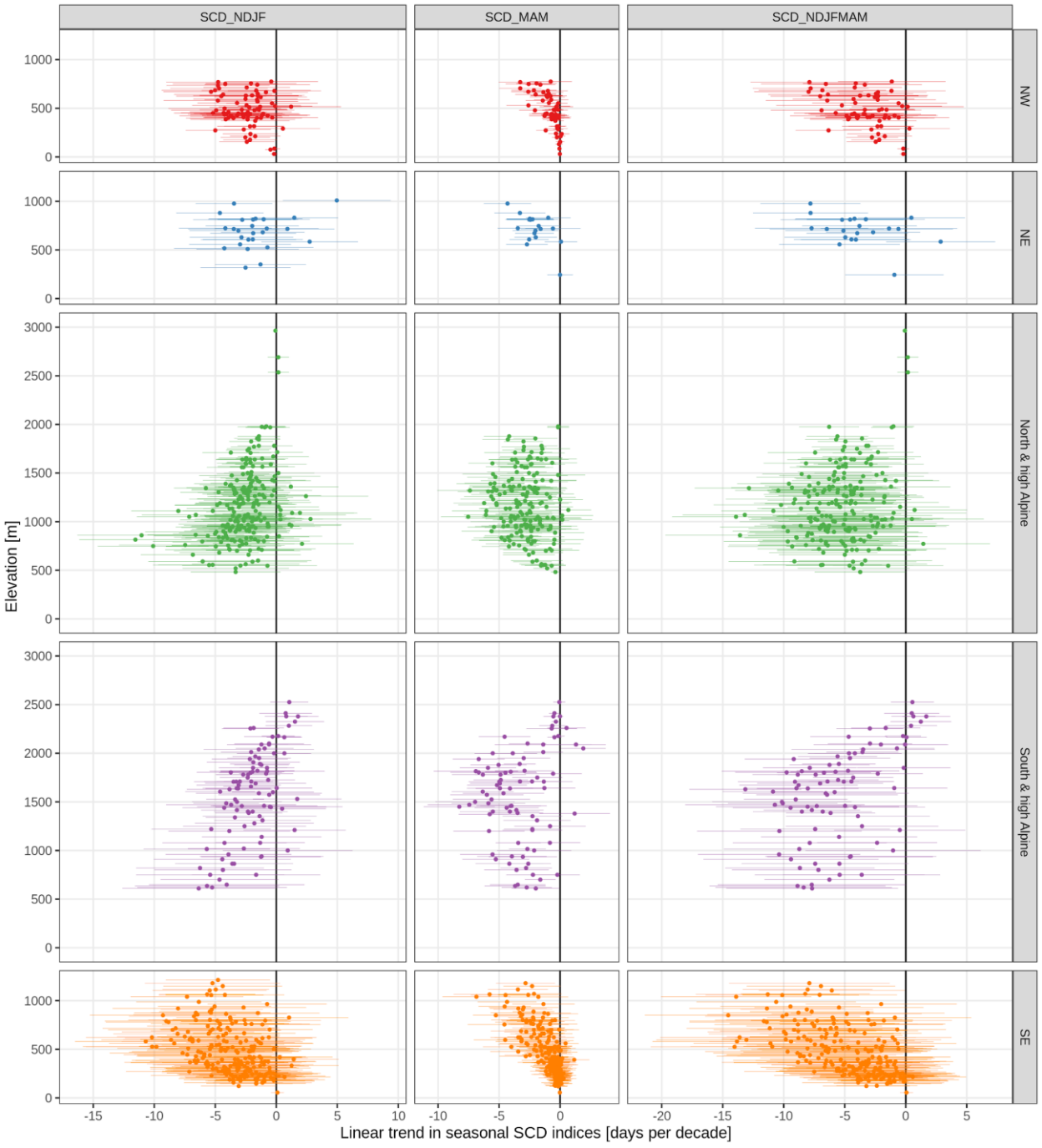


Figure C2: Long-term (1971 to 2019) linear trends in seasonal snow cover duration (SCD) indices. Trends are shown separately by index (columns) and region (rows). The season is indicated in the columns with the first letter of the included months (e.g. NDFJ is November, December, January, and February). Each point is one station. The points indicate the trend and the lines the associated 95% confidence interval.

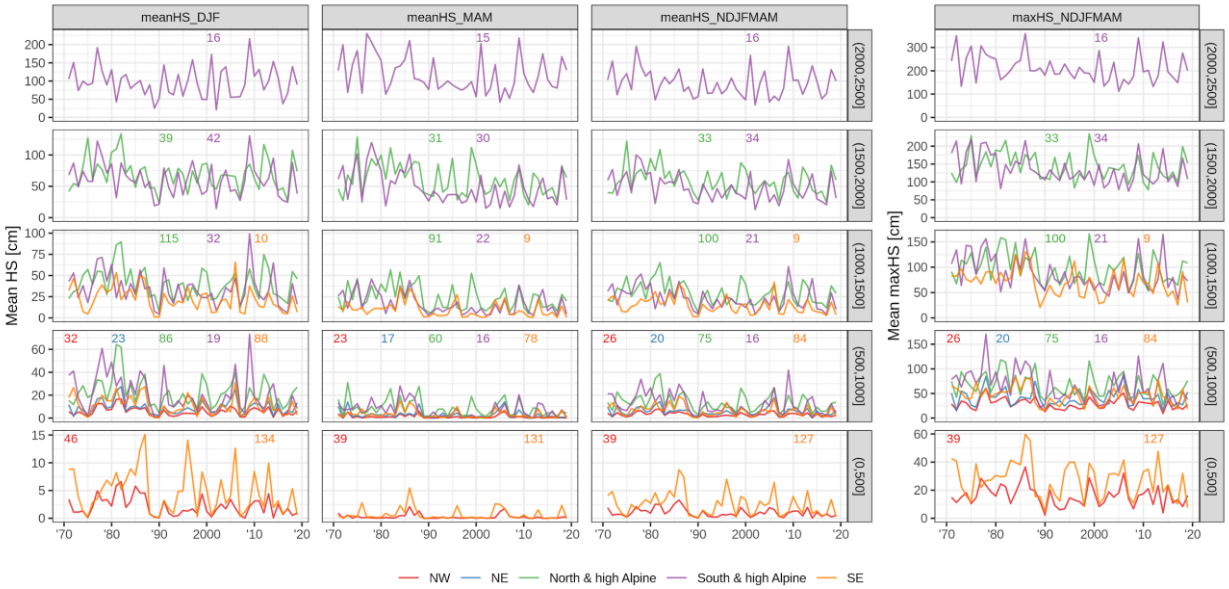


Figure C3: Time series of mean seasonal snow depth (HS) indices averaged by 500 m elevation bands. The rows indicate elevation band and the columns the index. The season is indicated in the columns with the first letter of the included months (e.g. DFJ is December, January, and February). The small numbers at the top of each panel denote the number of stations included in the average. Lines are only shown if more than 5 stations were available.

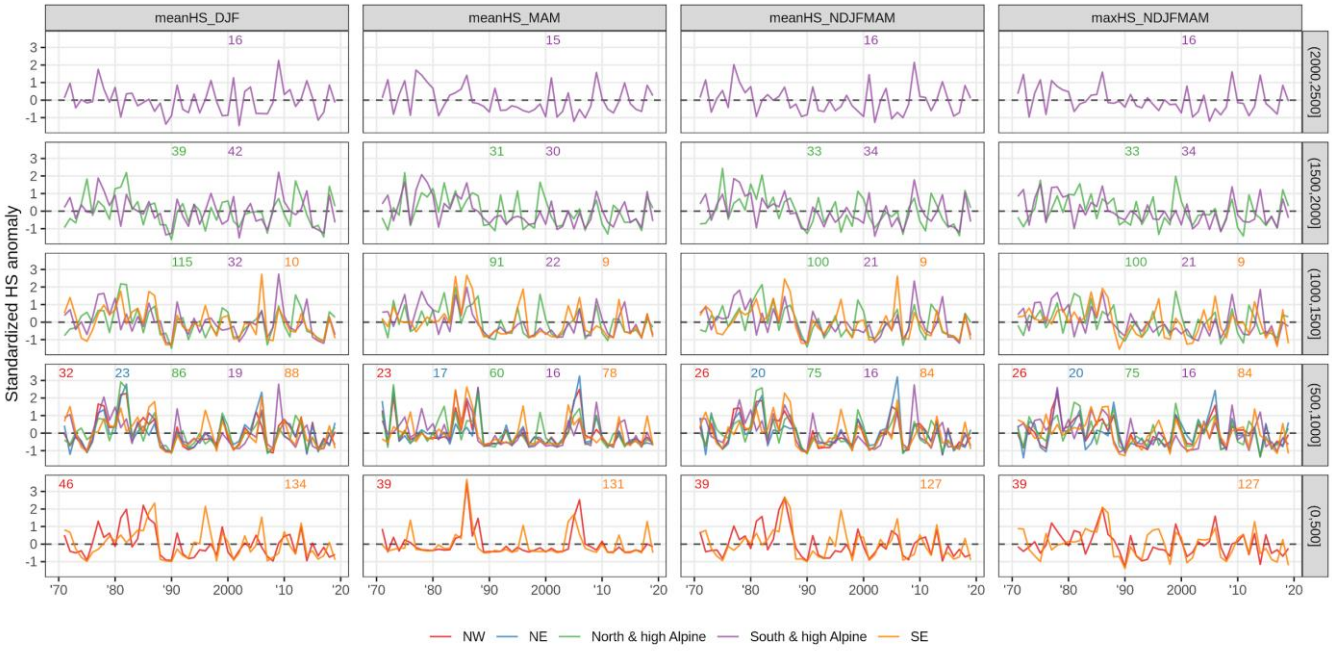


Figure C4: Same as Figure C3 but for standardized anomalies.

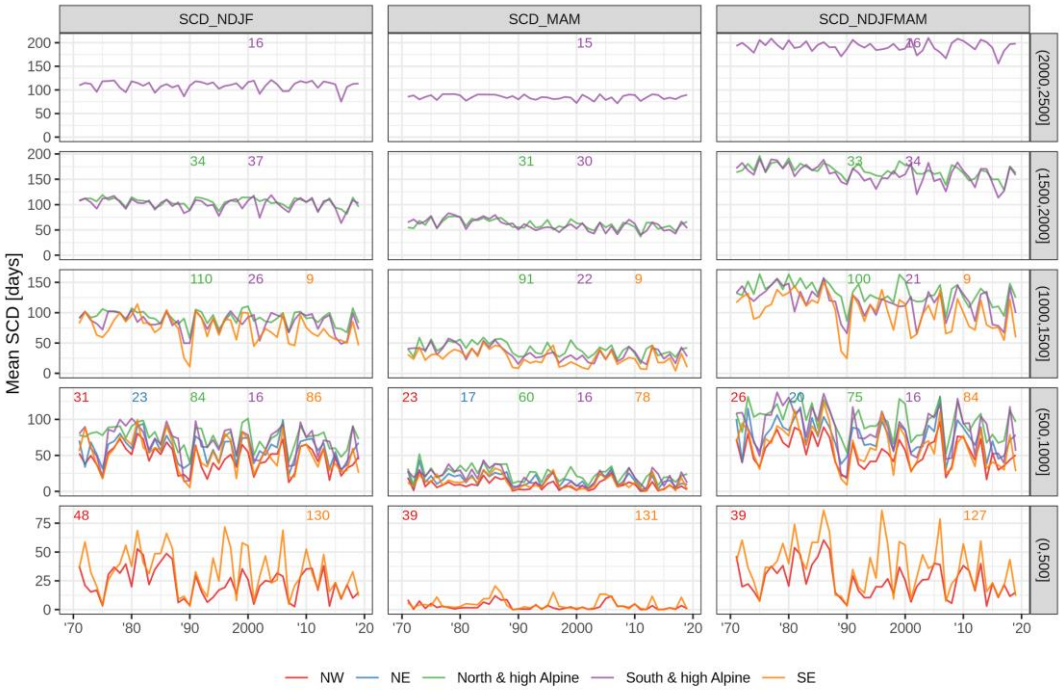


Figure C5: Time series of mean seasonal snow cover duration (SCD) indices averaged by 500 m elevation bands. The rows indicate elevation band and the columns the index. The season is indicated in the columns with the first letter of the included months (e.g. NDFJ is November, December, January, and February). The small numbers at the top of each panel denote the number of stations included in the average. Lines are only shown if more than 5 stations were available.

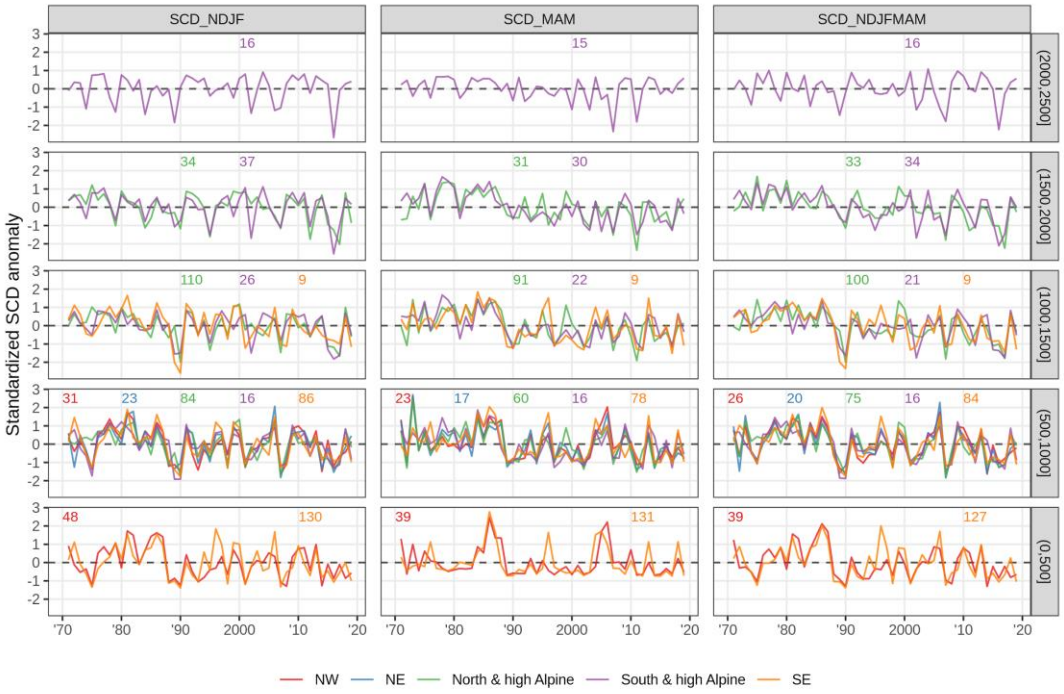


Figure C6: Same as Figure C5 but for standardized anomalies.

Table C1: Overview of long-term (1971 to 2019) trends in mean seasonal snow depth indices. Summaries are shown by index, region, and 1000 m elevation bands (0 to 1000, 1000 to 2000, and 2000 to 3000 m). Cell values are the number of stations (#), the mean trend (mean, in cm decade^{-1}), and percentages of significant negative (sig-) and positive (sig+) trends; the remaining percentage (not shown) corresponds to the total of non-significant negative and positive trends. Empty cells denote no station available (for # and mean), and no stations with significant negative or positive trends (sig- and sig+). Trends were considered significant if $p < 0.05$. See also Fig. C1. A version of the table with 500 m bands instead of 1000 m is available in the supplementary material (Table S6).

Index	Region	Elevation: (0,1000] m				Elevation: (1000,2000] m				Elevation: (2000,3000] m			
		#	mean	sig-	sig+	#	mean	sig-	sig+	#	mean	sig-	sig+
meanHS_DJF	NW	78	-0.38	26.9%									
	NE	25	-0.26	8.0%		1	1.36						
	N&hA	87	-1.64	14.9%		154	-2.09	7.8%		4	-4.28		
	S&hA	19	-3.57	42.1%		74	-3.56	17.6%		17	-0.07	5.9%	
	SE	222	-0.95	22.1%		10	-2.94	50.0%					
meanHS_MAM	NW	62	-0.12	9.7%									
	NE	18	-0.45	11.1%									
	N&hA	61	-1.56	47.5%		122	-3.74	42.6%		3	-4.45		
	S&hA	16	-1.34	43.8%		52	-5.38	69.2%		16	-6.73	31.2%	
	SE	209	-0.24	7.2%	0.5%	9	-1.82	33.3%					
meanHS_NDJFMAM	NW	65	-0.23	41.5%									
	NE	21	-0.31	9.5%									
	N&hA	76	-1.44	32.9%		133	-2.77	27.1%		3	-4.96		
	S&hA	16	-2.15	56.2%		55	-4.38	50.9%		17	-2.91	23.5%	
	SE	211	-0.60	27.0%		9	-2.13	55.6%					
maxHS_NDJFMAM	NW	65	-1.15	16.9%									
	NE	21	-0.82	4.8%									
	N&hA	76	-3.99	19.7%		133	-5.19	20.3%		3	-8.11		
	S&hA	16	-8.87	75.0%		55	-10.33	56.4%		17	-9.37	41.2%	
	SE	211	-2.78	27.0%		9	-6.59	55.6%					

Table C2: Overview of long-term (1971 to 2019) trends in mean seasonal snow cover duration indices. Summaries are shown by index, region, and 1000 m elevation bands (0 to 1000, 1000 to 2000, and 2000 to 3000 m). Cell values are the number of stations (#), the mean trend (mean, in days decade^{-1}), and percentages of significant negative (sig-) and positive (sig+) trends; the remaining percentage (not shown) corresponds to the total of non-significant negative and positive trends. Empty cells denote no station available (for # and mean), and no stations with significant negative or positive trends (sig- and sig+). Trends were considered significant if $p < 0.05$. See also Fig. C2. A version of the table with 500 m bands instead of 1000 m is available in the supplementary material (Table S7).

Index	Region	Elevation: (0,1000] m				Elevation: (1000,2000] m				Elevation: (2000,3000] m			
		#	mean	sig-	sig+	#	mean	sig-	sig+	#	mean	sig-	sig+
SCD_NDJF	NW	79	-2.47	30.4%									
	NE	25	-1.92	16.0%		1	4.96		100.0%				
	N&hA	85	-3.33	38.8%		144	-2.14	36.1%		3	0.09		
	S&hA	16	-3.79	6.2%		63	-2.08	28.6%		17	-0.20		5.9%
	SE	216	-3.67	28.2%		9	-5.26	66.7%					
SCD_MAM	NW	62	-0.82	22.6%									
	NE	18	-2.05	66.7%									
	N&hA	61	-2.56	59.0%		122	-3.03	66.4%					
	S&hA	16	-3.06	75.0%		52	-4.16	78.8%		16	-0.60	18.8%	6.2%
	SE	208	-0.99	20.7%		9	-3.58	66.7%					
SCD_NDJFMAM	NW	65	-3.33	40.0%									
	NE	21	-3.86	33.3%									
	N&hA	76	-5.58	57.9%		133	-5.28	73.7%		3	0.09		
	S&hA	16	-6.66	50.0%		55	-6.67	80.0%		17	-1.01	17.6%	
	SE	211	-4.70	34.6%		9	-8.84	88.9%					

Code and data availability

All computations were performed with R statistical software version 4.0.2 (RCoreTeam, 2008). Colors for the figures were taken from scientific color scales (Crameri, 2019) and colorbrewer. Code is available at a repository (Matiu et al., 2020), which includes scripts for the following tasks: reading in the different data sources, performing all data pre-processing, quality checking, gap filling, and statistical analyses.

Most data providers agreed to share their data: see Table B3 for the availability of daily and monthly values. For the full data set, please contact the main authors (MM or AC); the usage is generally free for research purposes, though explicit consent is required from some data providers, which want to keep track of the usage of the data. The shareable data is available at an open repository (Matiu et al., 2020).

Author contribution

Conceptualization: MM, AC, GB, CM, SM, WS; Data curation: MM, AC, GB, CMC, CM, DCB, GC, MV, WB, PC, GM, SCS, AC, RC, AD, MF, MG, LM, JMS, AS, AT, SU, VW; Formal analysis: MM, AC; Funding acquisition: MM; Investigation: MM; Methodology: MM, AC, CM, SM, WS; Resources: MZ; Software: MM, AC, CMC; Supervision: MM, MZ; Validation: MM, AC; Visualization: MM; Writing – original draft preparation: MM, AC; Writing – review & editing: MM, AC, GB, CMC, CM, SM, WS, GC, LDG, SK, BM, GR, ST, MV, PC, IG, GM, CN, SCS, US, MW, LG.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the reviewers (two anonymous and Ross Brown) for their comments, which have greatly improved the manuscript. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 795310. This work has benefited from funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730203. CNRM/CEN is a member of LabEX OSUG@2020. G.C. acknowledges the support of the Stiftungsfonds für Umweltökonomie und Nachhaltigkeit GmbH (SUN) and likewise the support from the DFG (Deutsche Forschungsgemeinschaft) Research Group FOR2793/1 "Sensitivity of High Alpine Geosystems to Climate Change since 1850 (SEHAG)" grant CH981/3.

We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (https://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu). For providing us with station data, we are grateful to Günther Geier from the meteorological office and avalanche warning from the province of Bolzano, to Sara Ratto from the Centro Funzionale della Regione Autonoma Valle d'Aosta, and to Gregor Vertačnik from the meteorological office of the Slovenian Environmental Agency.

References

Aschauer, J., Bavay, M., Begert, M. and Marty, C.: Comparing methods for gap filling in historical snow depth time series, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-17211, https://doi.org/10.5194/egusphere-egu2020-17211, 2020.

Auer, I., Böhm, R., Jurković, A., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Brunetti, M., Nanni, T., Maugeri, M., Briffa, K., Jones, P., Efthymiadis, D., Mestre, O., Moisselin, J.-M., Begert, M., Brazdil, R., Bochnicek, O., Cegnar, T., Gajić-Čapka, M., Zaninović, K., Majstorović, Ž., Szalai, S., Szentimrey, T. and Mercalli, L.: A new instrumental precipitation dataset for the greater alpine region for the period 1800–2002, International Journal of Climatology, 25(2), 139–166, https://doi.org/10.1002/joc.1135, 2005.

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z. and Nieplova, E.: HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region, International Journal of Climatology, 27(1), 17–46, https://doi.org/10.1002/joc.1377, 2007.

Bach, A. F., Schrier, G. van der, Melsen, L. A., Tank, A. M. G. K. and Teuling, A. J.: Widespread and Accelerated Decrease of Observed Mean and Extreme Snow Depth Over Europe, Geophysical Research Letters, 45(22), 12,312-12,319, https://doi.org/10.1029/2018GL079799, 2018.

Bazile, E., Abida, R., Verelle, A., Le Moigne, P. and Szczypta, C.: MESCAN-SURFEX surface analysis, deliverable D2.8 of the UERRA project. Technical Report European Commission. http://uerra.eu/publications/deliverable-reports.html, 2017.

Beniston, M.: Impacts of climatic change on water and associated economic activities in the Swiss Alps, Journal of Hydrology, 412–413, 291–296, https://doi.org/10.1016/j.jhydrol.2010.06.046, 2012a.

Beniston, M.: Is snow in the Alps receding or disappearing?, WIREs Climate Change, 3(4), 349–358, https://doi.org/10.1002/wcc.179, 2012b.

Beniston, M. and Stoffel, M.: Assessing the impacts of climatic change on mountain water resources, Science of The Total Environment, 493, 1129–1137, https://doi.org/10.1016/j.scitotenv.2013.11.122, 2014.

Beniston, M., Farinotti, D., Stoffel, M., Andreassen, L. M., Coppola, E., Eckert, N., Fantini, A., Giacona, F., Hauck, C., Huss, M., Huwald, H., Lehning, M., López-Moreno, J.-I., Magnusson, J., Marty, C., Morán-Tejéda, E., Morin, S., Naaim, M., Provenzale, A., Rabatel, A., Six, D., Stötter, J., Strasser, U., Terzago, S. and Vincent, C.: The European mountain cryosphere: a review of its current state, trends, and future challenges, The Cryosphere, 12(2), 759–794, https://doi.org/10.5194/tc-12-759-2018, 2018.

Bormann, K. J., Brown, R. D., Derksen, C. and Painter, T. H.: Estimating snow-cover trends from space, Nature Climate Change, 8(11), 924–928, https://doi.org/10.1038/s41558-018-0318-3, 2018.

Brown, R. D. and Petkova, N.: Snow cover variability in Bulgarian mountainous regions, 1931–2000, International Journal of Climatology, 27(9), 1215–1229, https://doi.org/10.1002/joc.1468, 2007.

Brunetti, M., Maugeri, M., Monti, F. and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, International Journal of Climatology, 26(3), 345–381, https://doi.org/10.1002/joc.1251, 2006.

Buchmann, M., Begert, M., Brönnimann, S. and Marty, C.: Evaluating the robustness of snow climate indicators using a unique set of parallel snow measurement series, International Journal of Climatology, 41(S1), E2553–E2563, https://doi.org/10.1002/joc.6863, 2021.

Cornes, R. C., Schrier, G. van der, Besselaar, E. J. M. van den and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, Journal of Geophysical Research: Atmospheres, 123(17), 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Crameri, F.: Scientific Colour Maps, Zenodo, DOI: 10.5281/zenodo.3596401., 2019.

Crespi, A., Brunetti, M., Lentini, G. and Maugeri, M.: 1961–1990 high-resolution monthly precipitation climatologies for Italy, International Journal of Climatology, 38(2), 878–895, https://doi.org/10.1002/joc.5217, 2018.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L. and Lesaffre, B.: Reanalysis of 47 Years of Climate in the French Alps (1958–2005): Climatology and Trends for Snow Cover, J. Appl. Meteor. Climatol., 48(12), 2487–2512, https://doi.org/10.1175/2009JAMC1810.1, 2009.

Esposito, A., Engel, M., Ciccazzo, S., Daprà, L., Penna, D., Comiti, F., Zerbe, S. and Brusetti, L.: Spatial and temporal variability of bacterial communities in high alpine water spring sediments, Research in Microbiology, 167(4), 325–333, https://doi.org/10.1016/j.resmic.2015.12.006, 2016.

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J. and Stoffel, M.: 21st century climate change in the European Alps—A review, Science of The Total Environment, 493, 1138–1151, https://doi.org/10.1016/j.scitotenv.2013.07.050, 2014.

Golzio, A., Crespi, A., Bollati, I. M., Senese, A., Guglielmina, A. D., Pelfini, M. and Maugeri, M.: High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps), Advances in Meteorology, 2018, e9123814, https://doi.org/10.1155/2018/9123814, 2018.

Haberkorn, A.: European Snow Booklet – an Inventory of Snow Measurements in Europe, EnviDat, https://doi.org/10.16904/envidat.59, 2019.

Hock, R., Rasul, G., Adler, C., Cáceres, B., Gruber, S., Hirabayashi, Y., Jackson, M., Kääb, A., Kang, S., Kutuzov, S., Milner, Al., Molau, U., Morin, S., Orlove, B. and Steltzer, H.: High Mountain Areas, in IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, edited by H.-O. Pörtner, D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. M. Weyer, In press, , 2019.

IPCC: Summary for Policymakers, in IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, edited by H.-O. Pörtner, D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. M. Weyer, In press, , 2019.

Isotta, F. and Frei, C.: APGD: Alpine precipitation grid dataset, , https://doi.org/10.18751/CLIMATE/GRIDDATA/APGD/1.0, 2013.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G. and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34(5), 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Keller, F., Goyette, S. and Beniston, M.: Sensitivity Analysis of Snow Cover to Climate Change Scenarios and Their Impact on Plant Habitats in Alpine Terrain, Climatic Change, 72(3), 299–319, https://doi.org/10.1007/s10584-005-5360-2, 2005.

Klein, G., Vitasse, Y., Rixen, C., Marty, C. and Rebetez, M.: Shorter snow cover duration since 1970 in the Swiss Alps due to earlier snowmelt more than to later snow onset, Climatic Change, 139(3), 637–649, https://doi.org/10.1007/s10584-016-1806-y, 2016.

Kreyling, J. and Henry, H. A. L.: Vanishing winters in Germany: soil frost dynamics and snow cover trends, and ecological implications, Climate Research, 46(3), 269–276, https://doi.org/10.3354/cr00996, 2011.

Laternser, M. and Schneebeli, M.: Long-term snow climate trends of the Swiss Alps (1931–99), International Journal of Climatology, 23(7), 733–750, https://doi.org/10.1002/joc.912, 2003.

Lejeune, Y., Dumont, M., Panel, J.-M., Lafaysse, M., Lapalus, P., Gac, E. L., Lesaffre, B. and Morin, S.: 57 years (1960–2017) of snow and meteorological observations from a mid-altitude mountain site (Col de Porte, France, 1325 m of altitude), Earth System Science Data, 11(1), 71–88, https://doi.org/10.5194/essd-11-71-2019, 2019.

Lencioni, V., Marziali, L. and Rossaro, B.: Diversity and distribution of chironomids (Diptera, Chironomidae) in pristine Alpine and pre-Alpine springs (Northern Italy), Journal of Limnology, 70(s1), 106–121, https://doi.org/10.4081/jlimnol.2011.s1.106, 2011.

Leporati, E. and Mercalli, L.: Snowfall series of Turin, 1784–1992: climatological analysis and action on structures, Annals of Glaciology, 19, 77–84, https://doi.org/10.3189/S0260305500011010, 1994.

López-Moreno, J. I., Soubeyroux, J. M., Gascoin, S., Alonso-Gonzalez, E., Durán-Gómez, N., Lafaysse, M., Vernay, M., Carmagnola, C. and Morin, S.: Long-term trends (1958–2017) in snow cover duration and depth in the Pyrenees, International Journal of Climatology, n/a(n/a), https://doi.org/10.1002/joc.6571, 2020.

Mallucci, S., Majone, B. and Bellin, A.: Detection and attribution of hydrological changes in a large Alpine river basin, Journal of Hydrology, 575, 1214–1229, https://doi.org/10.1016/j.jhydrol.2019.06.020, 2019.

Marcolini, G., Bellin, A. and Chiogna, G.: Performance of the Standard Normal Homogeneity Test for the homogenization of mean seasonal snow depth time series, International Journal of Climatology, 37(S1), 1267–1277, https://doi.org/10.1002/joc.4977, 2017a.

Marcolini, G., Bellin, A., Disse, M. and Chiogna, G.: Variability in snow depth time series in the Adige catchment, Journal of Hydrology: Regional Studies, 13, 240–254, https://doi.org/10.1016/j.ejrh.2017.08.007, 2017b.

Marcolini, G., Koch, R., Chimani, B., Schöner, W., Bellin, A., Disse, M. and Chiogna, G.: Evaluation of homogenization methods for seasonal snow depth data in the Austrian Alps, 1930–2010, International Journal of Climatology, 39(11), 4514–4530, https://doi.org/10.1002/joc.6095, 2019.

Marty, C.: Regime shift of snow days in Switzerland, Geophysical Research Letters, 35(12), https://doi.org/10.1029/2008GL033998, 2008.

Marty, C. and Blanchet, J.: Long-term changes in annual maximum snow depth and snowfall in Switzerland based on extreme value statistics, Climatic Change, 111(3), 705–721, https://doi.org/10.1007/s10584-011-0159-9, 2012.

Marty, C., Tilg, A.-M. and Jonas, T.: Recent Evidence of Large-Scale Receding Snow Water Equivalents in the European Alps, J. Hydrometeor., 18(4), 1021–1031, https://doi.org/10.1175/JHM-D-16-0188.1, 2017.

Matiu, M., Crespi, A., Bertoldi, G., Carmagnola, C. M., Marty, C., Morin, S., Schöner, W., Cat Berro, D., Chiogna, G., De Gregorio, L., Kotlarski, S., Majone, B., Resch, G., Terzago, S., Valt, M., Beozzo, W., Cianfarra, P., Gouttevin, I., Marcolini,

G., Notarnicola, C., Petitta, M., Scherrer, S. C., Strasser, U., Winkler, M., Zebisch, M., Cicogna, A., Cremonini, R., Debernardi, A., Faletto, M., Gaddo, M., Giovannini, L., Mercalli, L., Soubeyroux, J.-M., Sušnik, A., Trenti, A., Urbani, S. and Weilguni, V.: Snow cover in the European Alps: Station observations of snow depth and depth of snowfall (Version v1.1) [Data set], Zenodo, https://doi.org/10.5281/zenodo.4064128, 2020.

Micheletti, S.: Cambiamenti Climatici in Friuli-Venezia-Giulia, Neve e Valanghe, 63 https://issuu.com/aineva7/docs/nv63, 2008.

Najafi, M. R., Zwiers, F. and Gillett, N.: Attribution of the Observed Spring Snowpack Decline in British Columbia to Anthropogenic Climate Change, J. Climate, 30(11), 4113–4130, https://doi.org/10.1175/JCLI-D-16-0189.1, 2017.

Nitu, R., Roulet, Y.-A., Wolff, M., Earle, M., Reverdin, A., Smith, C., Kochendorfer, J., Morin, S., Rasmussen, R., Wong, K., Alastrué, J., Arnold, L., Baker, B., Buisán, S., Collado, J. L., Colli, M., Collins, B., Gaydos, A., Hannula, H.-R., Hoover, J., Joe, P., Kontu, A., Laine, T., Lanza, L., Lanzinger, E., Lee, G. W., Lejeune, Y., Leppänen, L., Mekis, E., Panel, J.-M., Poikonen, A., Ryu, S., Sabatini, F., Theriault, J., Yang, D., Genthon, C., Heuvel, F. van den, Hirasawa, N., Konishi, H., Motoyoshi, H., Nakai, S., Nishimura, K., Senese, A. and Amashita, K.: WMO Solid Precipitation Intercomparison Experiment (SPICE) (2012 - 2015), World Meteorological Organization (WMO). https://www.wmo.int/pages/prog/www/IMOP/publications-IOM-series.html, 2018.

Notarnicola, C.: Hotspots of snow cover changes in global mountain regions over 2000–2018, Remote Sensing of Environment, 243, 111781, https://doi.org/10.1016/j.rse.2020.111781, 2020.

Pepin, N., Bradley, R. S., Diaz, H. F., Baraer, M., Caceres, E. B., Forsythe, N., Fowler, H., Greenwood, G., Hashmi, M. Z., Liu, X. D., Miller, J. R., Ning, L., Ohmura, A., Palazzi, E., Rangwala, I., Schöner, W., Severskiy, I., Shahgedanova, M., Wang, M. B., Williamson, S. N., Yang, D. Q., and Mountain Research Initiative EDW Working Group: Elevation-dependent warming in mountain regions of the world, Nature Climate Change, 5(5), 424–430, https://doi.org/10.1038/nclimate2563, 2015.

Pierce, D. W., Barnett, T. P., Hidalgo, H. G., Das, T., Bonfils, C., Santer, B. D., Bala, G., Dettinger, M. D., Cayan, D. R., Mirin, A., Wood, A. W. and Nozawa, T.: Attribution of Declining Western U.S. Snowpack to Human Effects, J. Climate, 21(23), 6425–6444, https://doi.org/10.1175/2008JCLI2405.1, 2008.

Pifferetti, M., Cat Berro, D., Mercalli, L., Ricciardi, G. and Buffa, A.: La neve nella Pianura Padano-veneta: nuova cartografia 1961-2017, Nimbus, 77, 2017.

Pinheiro, J. C. and Bates, D. M.: Mixed-effects models in S and S-PLUS, Springer, New York., 2000.

Prein, A. F. and Gobiet, A.: Impacts of uncertainties in European gridded precipitation observations on regional climate analysis, International Journal of Climatology, 37(1), 305–327, https://doi.org/10.1002/joc.4706, 2017.

Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., Ikonen, J., Takala, M., Cohen, J., Smolander, T. and Norberg, J.: Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018, Nature, 581(7808), 294–298, https://doi.org/10.1038/s41586-020-2258-0, 2020.

RCoreTeam: R: A language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria., 2008.

Resch, G., Chimani, B., Koch, R., Schöner, W. and Marty, C.: Homogenization of long-term snow observations, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-8807, https://doi.org/10.5194/egusphere-egu2020-8807, 2020.

Salzmann, N. and Mearns, L. O.: Assessing the Performance of Multiple Regional Climate Model Simulations for Seasonal Mountain Snow in the Upper Colorado River Basin, J. Hydrometeor., 13(2), 539–556, https://doi.org/10.1175/2011JHM1371.1, 2011.

Scherrer, S. C. and Appenzeller, C.: Swiss Alpine snow pack variability: major patterns and links to local climate and large-scale flow, Climate Research, 32(3), 187–199, https://doi.org/10.3354/cr032187, 2006.

Scherrer, S. C., Wüthrich, C., Croci-Maspoli, M., Weingartner, R. and Appenzeller, C.: Snow variability in the Swiss Alps 1864–2009, International Journal of Climatology, 33(15), 3162–3173, https://doi.org/10.1002/joc.3653, 2013.

Schöner, W., Auer, I. and Böhm, R.: Long term trend of snow depth at Sonnblick (Austrian Alps) and its relation to climate change, Hydrological Processes, 23(7), 1052–1063, https://doi.org/10.1002/hyp.7209, 2009.

Schöner, W., Koch, R., Matulla, C., Marty, C. and Tilg, A.-M.: Spatiotemporal patterns of snow depth within the Swiss-Austrian Alps for the past half century (1961 to 2012) and linkages to climate change, International Journal of Climatology, 39(3), 1589–1603, https://doi.org/10.1002/joc.5902, 2019.

Schwaizer, G., Keuris, L., Nagler, T., Derksen, C., Luojus, K., Marin, C., Metsämäki, S., Mudryk, L., Naegeli, K., Notarnicola, C., Salberg, A.-B., Solberg, R., Wiesmann, A., Wunderle, S., Essery, R., Gustafsson, D., Krinner, G. and Trofaier, A.-M.: Towards a long term global snow climate data record from satellite data generated within the Snow Climate Change Initiative, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-19228, https://doi.org/10.5194/egusphere-egu2020-19228, 2020.

Steger, C., Kotlarski, S., Jonas, T. and Schär, C.: Alpine snow cover in a changing climate: a regional climate model perspective, Clim Dyn, 41(3–4), 735–754, https://doi.org/10.1007/s00382-012-1545-3, 2013.

Steiger, R. and Stötter, J.: Climate Change Impact Assessment of Ski Tourism in Tyrol, Tourism Geographies, 15(4), 577–600, https://doi.org/10.1080/14616688.2012.762539, 2013.

Storch, H. von and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, Cambridge., 1999.

Taylor, M. H., Losch, M., Wenzel, M. and Schröter, J.: On the Sensitivity of Field Reconstruction and Prediction Using Empirical Orthogonal Functions Derived from Gappy Data, J. Climate, 26(22), 9194–9205, https://doi.org/10.1175/JCLI-D-13-00089.1, 2013.

Terzago, S., Cassardo, C., Cremonini, R. and Fratianni, S.: Snow Precipitation and Snow Cover Climatic Variability for the Period 1971–2009 in the Southwestern Italian Alps: The 2008–2009 Snow Season Case Study, Water, 2(4), 773–787, https://doi.org/10.3390/w2040773, 2010.

Terzago, S., Fratianni, S. and Cremonini, R.: Winter precipitation in Western Italian Alps (1926–2010), Meteorol Atmos Phys, 119(3), 125–136, https://doi.org/10.1007/s00703-012-0231-7, 2013.

Thackeray, C. W., Derksen, C., Fletcher, C. G. and Hall, A.: Snow and Climate: Feedbacks, Drivers, and Indices of Change, Curr Clim Change Rep, 5(4), 322–333, https://doi.org/10.1007/s40641-019-00143-w, 2019.

Valt, M. and Cianfarra, P.: Recent snow cover variability in the Italian Alps, Cold Regions Science and Technology, 64(2), 146–157, https://doi.org/10.1016/j.coldregions.2010.08.008, 2010.

Valt, M., Cagnatti, A., Crepaz, A. and Cat Berro, D.: Variazioni Recenti del Manto Nevoso sul Versante Meridionale delle Alpi, Neve e Valanghe, 63 https://issuu.com/aineva7/docs/nv63, 2008.

Venables, W. N. and Ripley, B. D.: Modern Applied Statistics with S, 4th ed., Springer-Verlag, New York., 2002.

1165