Classification of Snow Depth Measurements for tracking plant phenological shifts in Alpine regions

Jan Svoboda[1] **Michael Zehnder**[1,3,4] **Marc Ruesch**[1] **David Liechti**[1]
Corinne Jones[2] **Michele Volpi**[2] **Christian Rixen**[1,4] **Jürg Schweizer**[1]
[1]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[2]Swiss Data Science Center, ETH Zürich and EPFL, Zürich, Switzerland
[3]Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland
[4]CERC, Davos Dorf, Switzerland
{jan.svoboda,michael.zehnder,ruesch,david.liechti,rixen,schweizer}@slf.ch
corinne.jones@epfl.ch
michele.volpi@sdsc.ethz.ch

Abstract

Ground-based snow depth measurements are often realized using ultrasonic or laser technologies, which by their nature measure the height of any underlying object, whether it is snow or vegetation in snow-free periods. We propose a machine learning approach to the automated classification of snow depth measurements into a snow cover class and a class corresponding to everything else, which takes into account both the temporal context and the dependencies between snow depth and other sensor measurements. Through a series of experiments we demonstrate that our approach simplifies the detection of seasonal snowmelt and corresponding onset of plant growth, which we used to assess climate-change related phenological shifts in otherwise rather poorly monitored high alpine regions.

1 Introduction

An excellent way to assess the effects of global warming on organisms is through phenology, which studies the timing and recurring patterns of natural events. Long-term studies have shown that the spring phenology of plants has advanced over the last decades [4, 7, 19, 23, 25, 20]. These findings have led to the establishment of phenological monitoring networks, which are, however, mostly situated in areas below the treeline. Besides evidence from satellite observations [25], field evidence of phenological shifts in high alpine regions is scarce due to the high costs and the labor involved.

In seasonally snow-covered regions, snowmelt and spring temperatures are the most significant phenological cues that initiate plant growth [6, 8, 9, 22, 26, 12]. The length of the snow-covered season in the Eureopean Alps has receded in the 20^{th} century with a significant shift in the timing of snowmelt, particularly below the treeline [10, 18, 21, 24]. A good proxy for high alpine regions, demonstrating corresponding snowmelt trends, is the climatological station located at Weissfluhjoch (Alpine range in Switzerland) (2540 m a.s.l.) [17], which suggests that the advancement of snowmelt timing is most likely transferable to higher elevations as well.

We propose to mitigate the lack of evidence above the treeline with the use of automated weather station (AWS) data from the Swiss Intercantonal Measurement and Information System (IMIS) [14]. IMIS employs ultrasonic sensors to measure snow height, which only measure the distance from the sensor to the underlying object, whether it is snow or vegetation. The previous attempts to disentangle snow measurements from the rest were based on thresholding [2], which is generally cumbersome to transfer to new stations, and random forests [3], which however cannot explicitly model the temporal

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2024.

structure of the data, deemed important to judge whether the signal coming from the sensor shows snow or vegetation.

In this work, we instead approach this binary classification problem using a Temporal Convolutional Network (TCN) [13] that explicitly accounts for the temporal relationships between different points in time series data. This accurate classification of snow height measurements allows us to track snowmelt and the start of plant growth over a climate relevant period (1998-2023), showing evidence of phenological shifts in high alpine regions – an excellent indicator of climate change [7].

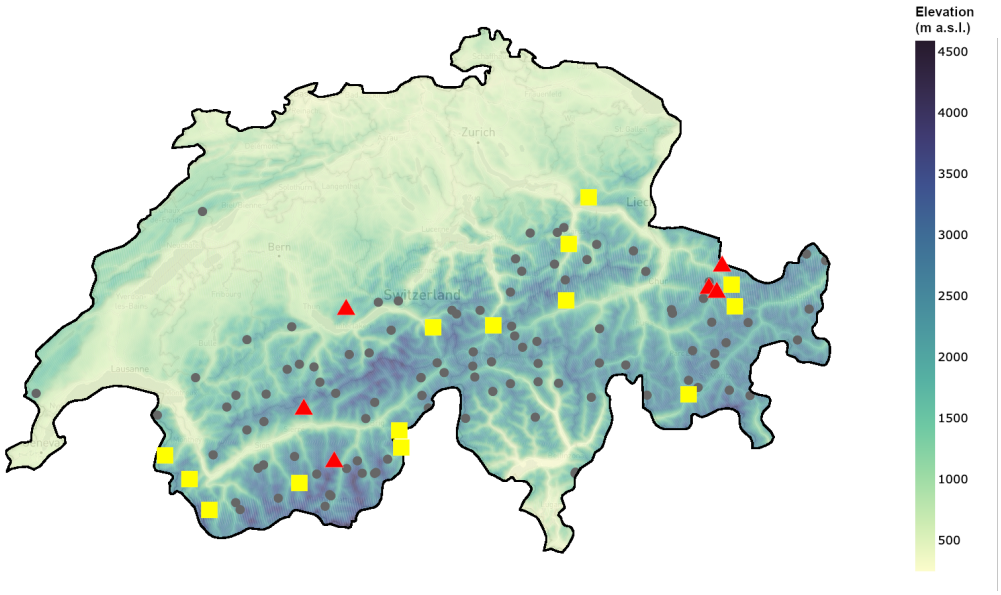


Figure 1: Map of IMIS stations in Switzerland. Stations marked as full gray circles were not used in model development. Yellow squares are the stations that have been used for training (14 stations) and red triangles indicate stations used for testing (6 stations). Colours indicate elevation in m a.s.l.

2 Data

We used snow height data from the IMIS, a network of 132 AWS (as of May 2024) focused on snow measurements that are distributed throughout the Swiss Alps and the Jura region (see Figure 1), mostly located above 2000 m a.s.l. The stations acquire snow height and meteorological data regularly in 30 minute intervals. Manual exploration showed that the following sensor measurements are key in disentangling snow from soil and vegetation measurements: snow height (HS), air temperature (TA), snow surface temperature (TSS), ground temperature (TG) and reflected short-wave solar radiation (RSWR).

Snow/no-snow dataset For model development and validation, we prepared a dataset with reliable ground truth information. A subset of 20 stations (see Figure 1 and Appendix A) which span different regions and elevations were selected and manually annotated with two-class ground truth information regarding snow height data into *snow* and *no snow* (e.g. vegetation, soil, rocks, etc.).

3 Methodology

The backbone of our model is a 4-layer TCN [13] architecture as shown in Figure 2, which has 4-dimensional time series with 48 time steps as the input. The number of layers and filter sizes were selected so that the output representation of the last point in the input time series is an aggregation of all previous time steps. This representation is fed to an MLP classifier, which first produces a series representation and then uses this representation to produce output class probabilities. Due to the properties of our dataset, we opt for the so-called focal loss [15], which allows the model to focus

and train preferentially on hard examples while down-weighting the simple cases throughout the training process.

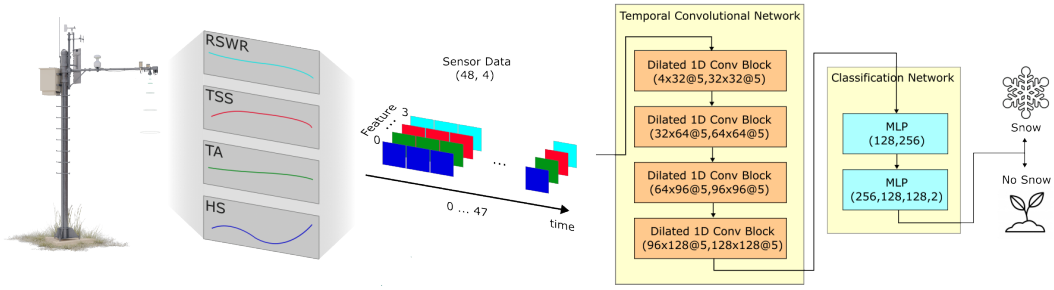


Figure 2: Structure of the input data and architecture of the modified TCN employed. Each dilated 1D conv block has filters described in the format (*input_features* $\times$ *output_features* @ *kernel_size*). The composition of each MLP is described as (*input_features*, *hidden_features_1*, . . . , *output_features*).

Experimental setup. We ran training for a maximum of 300 epochs, feeding the model with a batch of 64 samples in each iteration. We allowed for the possibility of early stopping if the validation loss has not improved for more than 50 epochs. The optimization process was governed by the AdamW [16] optimizer with an initial learning rate of 10^{-3}. The learning rate was subject to step decay with factor 0.1, three times, after 50, 100 and 150 epochs.

We took 14 stations from our dataset for training and use a 90/10 split with a fixed random seed to obtain the training and validation set. The remaining 6 stations (SLF2, SHE2, KLO2, TRU2, WFJ2, STN2) were used solely for performance evaluation purposes.

Phenological analyses. For plant phenological analyses, the 30min interval measurements of TA, TG and HS measurements classified using our TCN were aggregated to daily medians. Snowmelt was defined as the first day following the longest continuous winter snow cover, while snow-in marks the first day of the continuous snow cover in fall. Summer or vegetation season ground temperatures are determined as the median temperatures between meltout and snow-in. Spring temperature is defined as the median temperature from snow-out until the vegetation growth onset. Of the 132 AWS stations, 124 stations between 1600 and 2950m yielded sufficient data for snowmelt and temperature trends, with a mean coverage of 22 years per station. To extract the plant growth signal, days corresponding to HS signal classified as snow during the vegetation period were removed. We fitted a logistic curve [11] to the growth signal. Growing years with a poor fit (R<0.75) were excluded from the analysis, resulting in 40 stations between 1600 and 2700m with a mean coverage of 19 years for plant-phenological analysis. We defined an ecologically meaningful Start of Growth based on the tangent at the point of the maximum growth and the intersection with the x-axis [5] (details in Appendix C). To model temporal shifts, we applied linear-mixed effect models [1] with a phenological or climatic response variable, with "year" as a fixed effect and "station ID" as random effect.

4 Conclusions and Impact

We propose a new approach to classification of snow height measurements coming from AWS stations. Results in Appendix D demonstrate the efficiency of the model in disentangling snow height measurements during snowmelt and vegetation growth periods, predesposing it to be used in vegetation phenology and snow climatology. Early analysis of the classified time series from 1998-2023 suggests the advance of the timing of snowmelt and strong warming across all AWS stations. We additionally observe a shortening of the timelag between snowmelt and initiation of plant growth, indicating a warming-driven phenological advancement at a large subset of AWS stations suited for tracking plant growth. Our preliminary vegetation phenology study highlights the importance of long-term monitoring and automated machine learning approaches in understanding climate-induced phenological shifts, with implications for ecosystem dynamics in remote alpine regions (see also Appendix D.

References

[1] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4, 2014.

[2] M. Bavay and T. Egger. Meteoio 2.4.2: a preprocessing library for meteorological data. *Geoscientific Model Development*, 7(6):3135–3151, 2014.

[3] G. Blandini, F. Avanzi, S. Gabellani, D. Ponziani, H. Stevenin, S. Ratto, L. Ferraris, and A. Viglione. A random forest approach to quality-checking automatic snow-depth sensor measurements. *The Cryosphere*, 17(12):5317–5333, 2023.

[4] C. Collins, S. Elmendorf, R. Hollister, G. Henry, K. Clark, A. Bjorkman, I. Myers-Smith, J. Prevey, I. Ashton, J. Assmann, J. Alatalo, M. Carbognani, C. Chisholm, E. Cooper, C. Forrester, I. Jónsdóttir, K. Klanderud, C. Kopp, C. Livensperger, and K. Suding. Experimental warming differentially affects vegetative and reproductive phenology of tundra plants. *Nature Communications*, 12:3442, 06 2021.

[5] L. Gu, W. Post, D. Baldocchi, A. Black, A. Suyker, S. Verma, T. Vesala, and S. Wofsy. *Characterizing the Seasonal Dynamics of Plant Community Photosynthesis Across a Range of Vegetation Types*. 05 2009.

[6] D. Inouye. Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology*, 89:353–62, 03 2008.

[7] D. W. Inouye. Climate change and phenology. *WIREs Climate Change*, 13:e764, 2022.

[8] D. Jerome, W. Petry, K. Mooney, and A. Iler. Snowmelt timing acts independently and in conjunction with temperature accumulation to drive subalpine plant phenology. *Global Change Biology*, 27:5054–5069, 07 2021.

[9] T. Jonas, C. Rixen, M. Sturm, and V. Stoeckli. How alpine plant growth is linked to snow cover and climate variability. *J. Geophys. Res.-Biogeosci.*, 113:G03013, 07 2008.

[10] G. Klein, Y. Vitasse, C. Rixen, C. Marty, and M. Rebetez. Shorter snow cover duration since 1970 in the swiss alps due to earlier snowmelt more than to later snow onset. *Climatic Change*, 139, 12 2016.

[11] D. Kong, T. McVicar, X. Mingzhong, Y. Zhang, J. Peña-Arancibia, G. Filippa, Y. Xie, and G. Xihui. phenofit: An r package for extracting vegetation phenology from time series remote sensing. *Methods in Ecology and Evolution*, 04 2022.

[12] C. Körner. *Alpine Plant Life*. 2021.

[13] C. Lea, R. Vidal, A. Reiter, and G. D. Hager. Temporal convolutional networks: A unified approach to action segmentation. *CoRR*, 2016.

[14] M. Lehning, P. Bartelt, B. Brown, T. Russi, U. Stöckli, and M. Zimmerli. Snowpack model calculations for avalanche warning based upon a network of weather and snow stations. *Cold Regions Science and Technology*, 30:145–157, 12 1999.

[15] T. Lin, P. Goyal, R. B. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2999–3007. IEEE Computer Society, 2017.

[16] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *Proceedings of the International Conference on Learning Representations*, 2019.

[17] C. Marty and R. Meister. Long-term snow and weather observations at weissfluhjoch and its relation to other high-altitude observatories in the alps. *Theoretical and Applied Climatology*, 110, 12 2012.

[18] C. Marty, M. Rohrer, M. Huss, and M. Stähli. Multi-decadal observations in the alps reveal less and wetter snow, with increasing variability. *Frontiers in Earth Science*, 11, 06 2023.

[19] A. Menzel, T. Sparks, N. Estrella, E. Koch, A. Aasa, R. Ahas, K. Alm-Kübler, P. Bissolli, O. Braslavska, A. Briede, F. Chmielewski, Z. Crepinsek, Y. Curnel, A. Dahl, C. Defila, A. Donnelly, Y. Filella, K. Jatczak, F. Mage, A. Mestre, O. Nordli, J. Penuelas, P. Pirinen, V. Remisova, H. Scheifinger, M. Striz, A. Susnik, A. van Vliet, F. Wielgolaski, S. Zach, and A. Zust. European phenological response to climate change matches the warming pattern. *Global Change Biology*, 12(10):1969–1976, 2006.

[20] C. Rixen, T. T. Høye, P. Macek, R. Aerts, J. M. Alatalo, J. T. Anderson, P. A. Arnold, I. C. Barrio, J. W. Bjerke, M. P. Björkman, D. Blok, G. Blume-Werry, J. Boike, S. Bokhorst, M. Carbognani, C. T. Christiansen, P. Convey, E. J. Cooper, J. H. C. Cornelissen, S. J. Coulson, E. Dorrepaal, B. Elberling, S. C. Elmendorf, C. Elphinstone, T. G. Forte, E. R. Frei, S. R. Geange, F. Gehrmann, C. Gibson, P. Grogan, A. H. Halbritter, J. Harte, G. H. Henry, D. W. Inouye, R. E. Irwin, G. Jespersen, I. S. Jónsdóttir, J. Y. Jung, D. H. Klinges, G. Kudo, J. Lämsä, H. Lee, J. J. Lembrechts, S. Lett, J. S. Lynn, H. M. Mann, M. Mastepanov, J. Morse, I. H. Myers-Smith, J. Olofsson, R. Paavola, A. Petraglia, G. K. Phoenix, P. Semenchuk, M. B. Siewert, R. Slatyer, M. J. Spasojevic, K. Suding, P. Sullivan, K. L. Thompson, M. Väisänen, V. Vandvik, S. Venn, J. Walz, R. Way, J. M. Welker, S. Wipf, and S. Zong. Winters are changing: snow effects on arctic and alpine tundra ecosystems. *Arctic Science*, 8(3):572–608, 2022.

[21] S. C. Scherrer. Temperature monitoring in mountain regions using reanalyses: lessons from the alps. *Environmental Research Letters*, 15(4):044005, mar 2020.

[22] Y. Vitasse, M. Rebetez, G. Filippa, E. Cremonese, G. Klein, and C. Rixen. 'hearing' alpine plants growing after snowmelt: ultrasonic snow sensors provide long-term series of alpine plant phenology. *International Journal of Biometeorology*, 61(2):349–361, 2017.

[23] Y. Vitasse, S. Ursenbacher, G. Klein, T. Bohnenstengel, Y. Chittaro, A. Delestrade, C. Monnerat, M. Rebetez, C. Rixen, N. Strebel, B. Schmidt, S. Wipf, T. Wohlgemuth, N. Yoccoz, and J. Lenoir. Phenological and elevational shifts of plants, animals and fungi under climate change in the european alps. *Biological Reviews*, 04 2021.

[24] M. Vorkauf, C. Marty, A. Kahmen, and E. Hiltbrunner. Past and future snowmelt trends in the swiss alps: the role of temperature and snowpack. *Climatic Change*, 165, 04 2021.

[25] H. Wang, H. Liu, G. Cao, Z. Ma, Y. Li, Z. Fawei, X. Zhao, X. Zhao, L. Jiang, N. Sanders, A. Classen, and J.-S. He. Alpine grassland plants grow earlier and faster but biomass remains unchanged over 35 years of climate change. *Ecology Letters*, 23, 02 2020.

[26] S. Wipf and C. Rixen. A review of snow manipulation experiments in arctic and alpine tundra ecosystems. *Polar Research*, 29(1):95–109, Apr. 2010.

A Snow/No-Snow Dataset

For model development and validation, we prepared a dataset with reliable ground truth information, which resulted in a new publicly available dataset[1]. Manually annotating snow height data is a tedious process, and doing so for the whole IMIS network is intractable. Therefore, we identified a subset of IMIS stations that we then manually annotated.

It should be mentioned that annotating historical data is problematic, as there is no way of checking whether there really was snow at the station or not. This means that assessing the presence of snow with the help of information from other sensors should be considered a best effort approach.

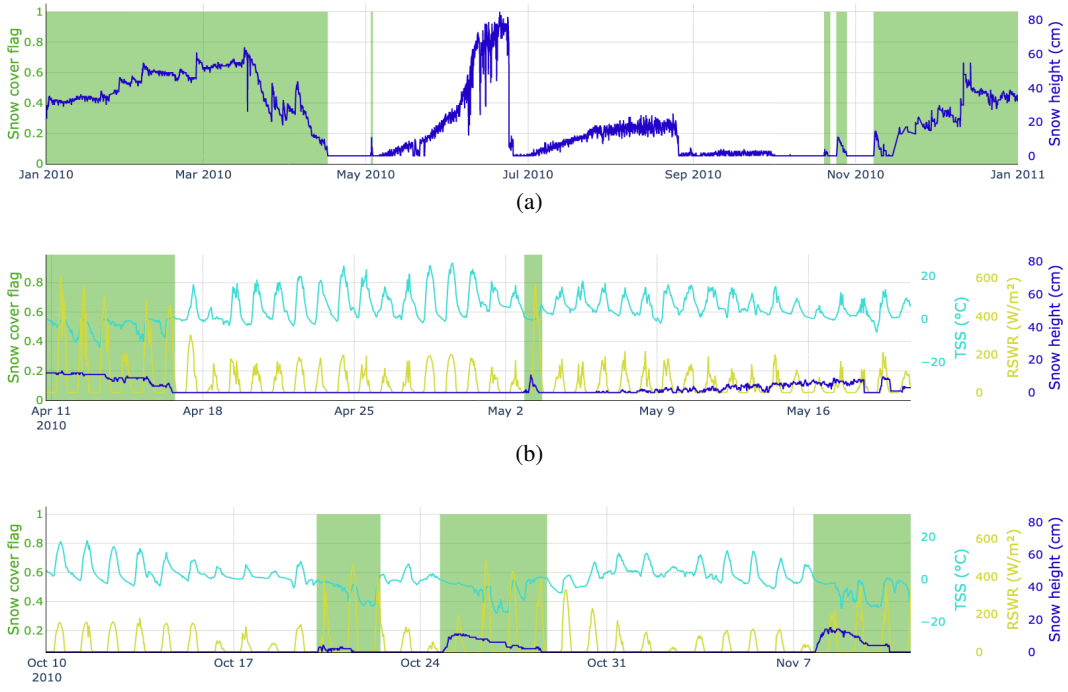


Figure 3: Examples of manually annotated data for the calendar year 2010 at the station SLF2. (a) shows the snow cover flag and snow height; green rectangles mark periods with snow cover. (b) focuses on the end of winter season 2009/2010 illustrating the behavior of TSS and RSWR dependent on whether there is snow or not. (c) is the same as in (b) for the beginning of the winter season 2010/2011.

A subset of 20 stations (see Appendix A.1) which span different locations and elevations and vary in underlying surface (e.g., vegetation, bare ground, glacier, etc.) were manually annotated with binary two-class ground truth information regarding snow height data:

- Class 0 - *snow* - The surface is covered by snow.
- Class 1 - *no snow* - the surface is snow-free (e.g., vegetation, soil, rocks, etc.).

The stations annotated with ground-truth information are depicted in yellow and red in Figure 1. An example of data annotation is shown in Figure 3, with two detailed views that emphasize the differences in behavior of TSS and RSWR in the presence and absence of a snow cover. The selected stations mostly contain data between 2000 and 2023 with a few exceptions for stations that have been built later (BOR2, FLU2, LAG3, RNZ2 and SHE2; see Appendix A.1). The data is recorded regularly at a 30-minute frequency.

[1]Dataset is available at: https://anonymized.link

A.1 List of stations in the snow/no-snow dataset

This section provides the list of IMIS stations used in our snow/no-snow dataset together with their metadata. Table 1 shows the stations ordered by increasing elevation. The column *Subset* indicates whether a station was used for training or testing.

Station ID	Latitude [°N]	Longitude [°E]	Elevation [m]	Available since	Subset
SLF2	46.8127	9.8482	1563	November 1997	**test**
AMD2	47.1708	9.1468	1610	October 1997	train
GLA2	46.9966	9.0375	1632	November 2000	train
SHE2	46.7488	7.8124	1852	October 2001	**test**
ILI2	46.1913	6.8277	2022	March 2000	train
GUT2	46.6793	8.2896	2115	November 1999	train
KLO2	46.9091	9.8738	2147	November 1996	**test**
TUM2	46.7810	9.0214	2191	October 2002	train
FNH2	46.1007	6.9641	2252	September 1997	train
KLO3	46.8412	9.9316	2299	November 1996	train
LAG3	46.4245	9.6977	2300	November 2009	train
FLU2	46.7527	9.9464	2394	October 2003	train
RNZ2	46.6855	8.6267	2400	December 2008	train
TRU2	46.3709	7.5855	2459	November 1996	**test**
BOR2	46.2905	8.1093	2517	September 2001	train
WFJ2	46.8296	9.8092	2536	January 1996	**test**
ARO3	46.0874	7.5620	2602	September 1996	train
SPN2	46.2294	8.1176	2620	November 1996	train
FOU2	45.9717	7.0672	2800	October 1999	train
STN2	46.1678	7.7505	2914	October 1998	**test**

Table 1: List of stations that are part of the snow/no-snow dataset, together with their auxiliary information, ordered by elevation.

A.2 Subsampling of the training data

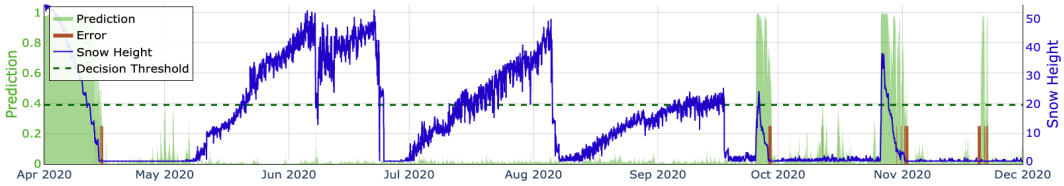
To run experiments in a reasonable time and make sure they were computationally tractable, we sub-sampled the training dataset to reduce the amount of training samples. In Table 2 we list which years were selected for each station for the training set.

Station ID	Selected years
AMD2	1998, 2001, 2004, 2007, 2010, 2013, 2016, 2019, 2022
GLA2	2001, 2004, 2007, 2010, 2013, 2016, 2019, 2022
ILI2	2002, 2005, 2008, 2011, 2014, 2017, 2020, 2023
GUT2	2000, 2003, 2006, 2009, 2012, 2015, 2018, 2021
TUM2	2004, 2007, 2010, 2013, 2016, 2019, 2022
FNH2	2000, 2003, 2006, 2009, 2012, 2015, 2018, 2021
KLO3	1999, 2002, 2005, 2008, 2011, 2014, 2017, 2020, 2023
LAG3	2011, 2014, 2017, 2020, 2023
FLU2	2005, 2008, 2011, 2014, 2017, 2020, 2023
RNZ2	2010, 2013, 2016, 2019, 2022
BOR2	2002, 2005, 2008, 2011, 2014, 2017, 2020, 2023
ARO3	1998, 2000, 2003, 2006, 2009, 2012, 2015, 2018, 2021
SPN2	1999, 2002, 2005, 2008, 2011, 2014, 2017, 2020, 2023
FOU2	2001, 2004, 2007, 2010, 2013, 2016, 2019, 2022

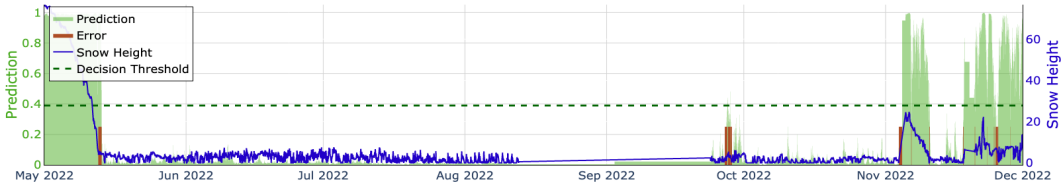
Table 2: List of years for each station that were selected as part of the sub-sampled training dataset.

B Visual examples of snow height classification results

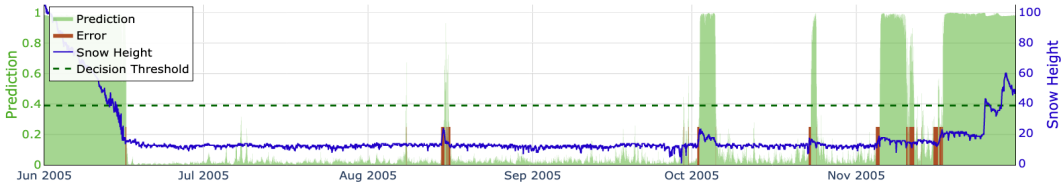
This section provides a few concrete examples of snow height classification results. Figure 4 depicts both examples where the model succeeds as well as some typical cases in which the model fails to correctly classify the snow height signal.



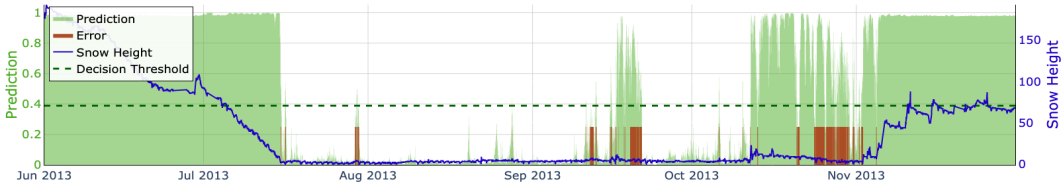
(a) SLF2 (1563 m), Year 2020



(b) SHE2 (1852 m), Year 2022



(c) TRU2 (2459 m), Year 2005



(d) STN2 (2914 m), Year 2013

Figure 4: Examples of classification results. The snow height signal is depicted in blue. The model predictions in terms of probability (0 - 1) are shown in green. The dashed horizontal line denotes the decision threshold for binary classification. The red-shaded areas show classification errors. (a) shows a correct classification of summer vegetation growth. (b) is an example of early October snowfall that has been classified partially correctly. (c) demonstrates the model's capability to detect summer snowfalls as well as scattered snowfalls at the beginning of winter. (d) is evidence that the model does not always perform well, here making mistakes at the beginning of the next winter season.

C Determination of start of growth

An example of fitted a logistic growth curve [11] to the clean plant growth measurements is shown in Figure 5). T he start of plant growth is determined based on the tangent at the steepest point of the logistic growth curve and its intersection with the x-axis [5].

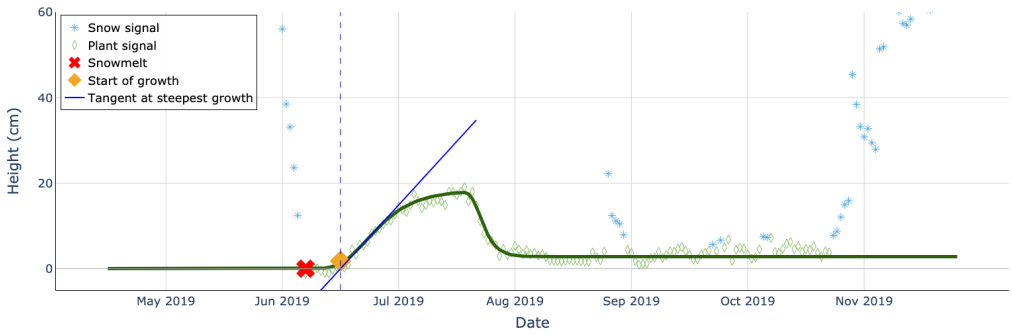


Figure 5: An example of a logistic growth curve (in dark green) fitted to height measurements data from TUJ2, in the vegetation season of year 2019. Snow height data corresponding to snow are shown with blue stars, while plant signal is shown with green diamonds. The red cross marks the snowmelt date, while the orange diamond marks the start of plant growth. The start of plant growth is determined using the blue tangent line as described in [5].

D Results

Classification performance. Figure 6 shows the classification performance of our model depending on the time of the year. The model performs very well in April, May, June and July, which are typically the months when snowmelt occurs (see also Appendix B), achieving an F1-score of 98% for snow and 99% for snow-free ground. In addition, the model performs very well in classifying the *no snow* class during the period of vegetation growth with an F1-score of 99.2%. These results suggest that our model can be applied in phenological studies focusing on snowmelt and vegetation growth shift analysis.

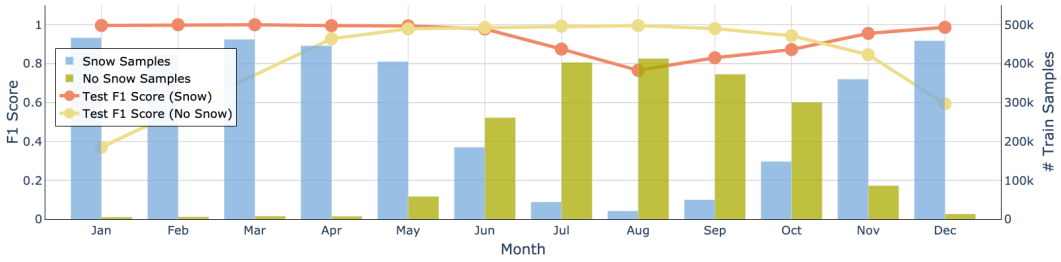
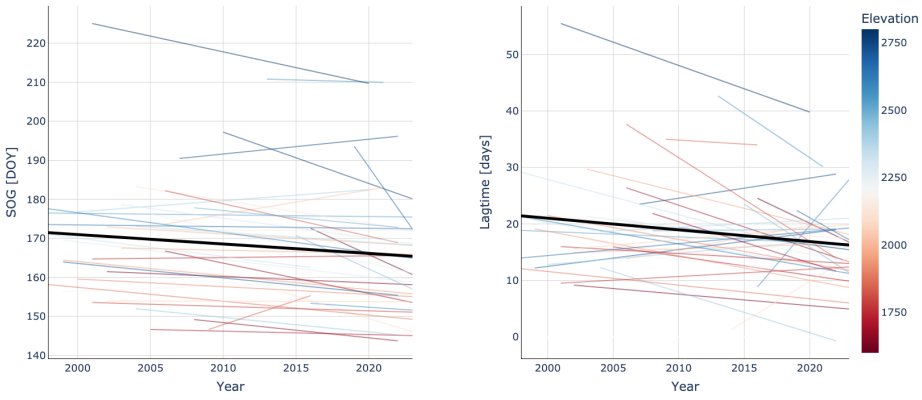


Figure 6: Performance of the model for each month of the year separately. The F1-score is shown separately for the classification of snow (red line) and no snow (yellow line). The blue columns indicate the distribution of snow samples, while the green columns indicate the distribution of the no-snow samples.

Snowmelt timing and warming trends. Spring temperatures have risen by 1.95°C ± 0.35° over the study duration of 25 years with a rate of 0.78°C/decade (95% CI: 0.42 – 1.13, $p < 0.001$). For AWS stations with phenological data (n=40), the timing of snowmelt has not advanced significantly with -0.28 days/decade (95% CI: -2.04 – 1.48, $p = 0.755$). Notably, when we look at all AWS stations (n= 126), we observe an advance of the snowmelt by -1.35 days/decade (95%: -2.15 – -0.55, $p = 0.001$) amounting to -3.38 days ± 3.29 (see Figure 7(a)).

Phenological shifts. The Start of Growth was directly linked to the timing of snowmelt, consistent with other studies [8, 9], while late snowfall events shifted the start of growth towards later calendar days. The Start of Growth, has changed by -2.41 days/decade (95% CI: -3.96 – -0.91, $p = 0.002$) amounting to an advance of 6.03 days ± 5.35. The unchanged timing of snowmelt at phenological stations and the earlier Start of Growth results in a shortening of the time interval between snowmelt and plant growth initiation; this lagtime has shortened by -2.06 days/decade (95% CI: -3.17 – -0.96, $p = 0.001$) resulting in an overall shorting of 5.15 days ± 2.85 (see Figure 7(b)). In 1998, the mean lagtime between snow cover disappearance and growth onset was 21.1 days, while in 2023 this period only lasts 16.3 days. Thus, alpine grasslands green up sooner after snowmelt.



(a) (b)

Figure 7: The changing trend in the timing of (a) Start of Growth and (b) Lagtime. Color lines show the trends for individual stations while the black line shows the overall trend for all stations.