HOW MANY SNOW PROFILES CAN YOU PROCESS? MAKING THE WEALTH OF INFORMATION INCLUDED IN LARGE-SCALE SNOWPACK SIMULATIONS MORE ACCESSIBLE FOR OPERATIONAL AVALANCHE FORECASTING

Florian Herla[1]*, Pascal Haegeli[1], Simon Horton[1,2], Patrick Mair[3]

[1] Avalanche Research Program, Simon Fraser University, Burnaby BC, Canada
[2] Avalanche Canada, Revelstoke BC, Canada
[3] Harvard University, Cambridge MA, USA

ABSTRACT: Snowpack models can provide detailed and continuous insight about the evolution of the snow stratigraphy in ways that are not possible with direct observations. However, the volume of data generated by the simulations can easily become overwhelming, and since simulated snow profiles are characterized by a rather complex, multidimensional data format, it is challenging to analyze the rich information manually. The available information is therefore commonly reduced to bulk properties and summary statistics of the entire snow column or individual grid points. This is only of limited value for operational avalanche forecasting where knowledge about thin, critical avalanche layers is important. In our opinion, the lack of efficient ways to access and mine large numbers of snow profiles is one of the key reasons for the limited operational use of spatially distributed snowpack simulations and ensemble systems so far.

We discuss recently developed tools for numerically processing snow profiles to make large volumes of snowpack model output more accessible for practitioners in relevant ways. This includes algorithms that compare and assess generic snow profiles by matching corresponding layers and aligning them before effectively synthesizing many profiles into a meaningful overall perspective. Our approach enables the computation of informative summary statistics and distributions of snowpack layers, as well as the dynamic clustering of profiles into groups with distinct conditions. Our algorithms are based on customized versions of Dynamic Time Warping (DTW) and DTW Barycenter Averaging (DBA), well established methods in the data sciences.

To be presented at the 2023 International Snow Science Workshop in Bend, Oregon, from October 8-13, 2023.

HOW MANY SNOW PROFILES CAN YOU PROCESS? MAKING THE WEALTH OF INFORMATION INCLUDED IN LARGE-SCALE SNOWPACK SIMULATIONS MORE ACCESSIBLE FOR OPERATIONAL AVALANCHE FORECASTING

Florian Herla[1,*], Pascal Haegeli[1], Simon Horton[1,2], Patrick Mair[3]

[1] *Avalanche Research Program, Simon Fraser University, Burnaby, BC, Canada*
[2] *Avalanche Canada, Revelstoke, BC, Canada*
[3] *Harvard University, Cambridge, MA, USA*

ABSTRACT: Snowpack models can provide detailed and continuous insight about the evolution of the snow stratigraphy in ways that are not possible with direct observations. However, the volume of data generated by the simulations can easily become overwhelming, and since simulated snow profiles are characterized by a rather complex, multidimensional data format, it is challenging to analyze the rich information manually. The available information is therefore commonly reduced to bulk properties and summary statistics of the entire snow column or individual grid points. This is only of limited value for operational avalanche forecasting where knowledge about thin, critical avalanche layers is important. In our opinion, the lack of efficient ways to access and mine large numbers of snow profiles is one of the key reasons for the limited operational use of spatially distributed snowpack simulations and ensemble systems so far.

We discuss recently developed tools for numerically processing snow profiles to make large volumes of snowpack model output more accessible for practitioners in relevant ways. This includes algorithms that compare and assess generic snow profiles by matching corresponding layers and aligning them before effectively synthesizing many profiles into a meaningful overall perspective. Our approach enables the computation of informative summary statistics and distributions of snowpack layers, as well as the dynamic clustering of profiles into groups with distinct conditions. Our algorithms are based on customized versions of Dynamic Time Warping (DTW) and DTW Barycenter Averaging (DBA), well established methods in the data sciences.

Keywords: Avalanche forecasting, snowpack modeling, data mining, comparing, summarizing, clustering.

1. INTRODUCTION

The layered nature of the snowpack is a necessary condition for the formation of snow avalanches (e.g., Schweizer et al., 2016), and snow stratigraphy information is crucial for developing a meaningful understanding of existing avalanche conditions (Statham et al., 2018). Snowpacks are inherently spatially variable due to the complex interactions of the meteorological forcing and terrain (Schweizer et al., 2007), and layer depths, thicknesses and properties can therefore vary substantially between different locations even over short distances. In some circumstances, some layer sequences might even be missing entirely. To understand the conditions at various spatial scales, avalanche forecasters observe snow profiles at targeted point locations, and then synthesize the gathered information into a mental model of the regional scale snowpack conditions, which are often represented in hand-drawn summary profiles. The documented layers in these idealized snow profiles represent key features of the conditions that forecasters expect to exist within their region. Local field observations are then used to validate and localize the regional understanding of the conditions. As the season progresses, forecasters continuously revise their mental model and update their summary profile throughout the winter as new observations become available.

While avalanche forecasters have developed meaningful strategies for synthesizing manual snowpack observations, the potential volume of data generated by snowpack simulations is too vast for human processing (Morin et al., 2020). While effective visualization designs can help guide human perception to data features that prompt human reasoning (Horton et al., 2020b; Nowak, 2023), visualizations of large data sets that include both spatial and temporal dimensions remain challenging. Since computer-based tools excel at applying repetitive tasks to big data sets, numerical data processing algorithms have great potential to allow avalanche forecasters to make better use of large scale snowpack simulations.

---

[*] Corresponding author address:
Florian Herla, SFU Avalanche Research Program,
email: fherla@sfu.ca

The objective of this paper is to highlight and summarize recent efforts by Herla et al. (2021, 2022a) and Horton et al. (in preparation) to develop a suite of modeling tools for numerically processing simulated snow profiles in ways relevant for avalanche forecasting applications.

2. COMPARING OF PROFILES

Designing numerical tools for snow profile analysis and synthesis needs to start with the development of algorithms that enable basic processing functionalities. This includes the tasks of comparing pairs of snow profiles by matching their corresponding layers and aligning them before assessing the similarity of the two profiles.

Inspired by Hagenmuller and Pilloix (2016), who adapted methods for numerically processing one-dimensional snow ram resistance measurements from field observations, Herla et al. (2021) designed algorithms to process generic snow profiles similar to the ones recorded manually in snow pits. The tools center around a snow profile alignment algorithm that is based on Dynamic Time Warping (DTW), a popular data science method originally developed for speech recognition (Sakoe and Chiba, 1978). To identify corresponding layers between two profiles that might be characterized by different properties or located at different depths, the algorithm makes use of one or multiple layer characteristics, such as grain type, hand hardness, layer dates, or more recently also density and optical grain size. After the matching of layers, a similarity measure that is specifically geared to avalanche forecasting purposes assesses the similarity of all corresponding layers (including any unmatched layers) based on grain type categories or more in-depth information about layer instabilities.

While primarily developed for comparing simulated snow profiles in ways that are tangible for forecasters, the alignment algorithm and similarity measure can technically be applied to manual snow profiles as well (Figure 1). This creates opportunities to quantitatively compare modeled and observed snow profiles for model validation purposes.

3. SUMMARIZING OF PROFILES

To allow forecasters to tap into a wealth of simulated snowpack information from a distributed system across large areas (e.g., forecast region) more easily, Herla et al. (2022a) implemented tools to summarize the data in relevant and familiar ways by following the design principle ”overview first, details on demand” (Shneiderman, 1996). First a representative snow profile provides a quick overview of the prevalent snowpack features before detailed distributions of layer characteristics provide more nuanced insight about the conditions.

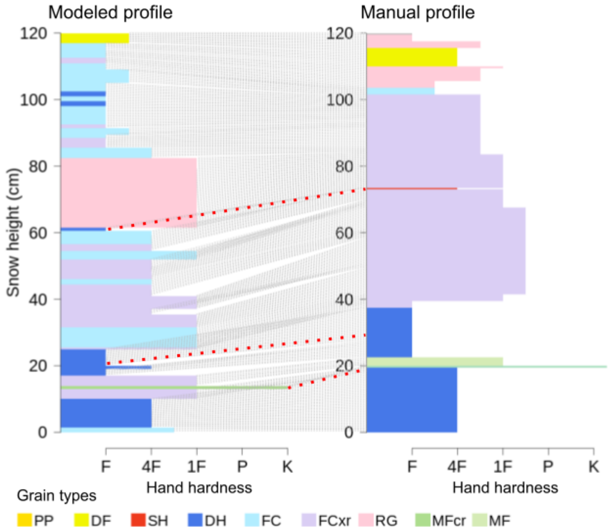

Figure 1: A modeled and manual snow profile from the same drainage. Gray lines visualize how the layer matching algorithm identifies corresponding layers between the two profiles although many layer properties vary substantially between the two profiles. Red dashed lines draw attention to important specific snowpack features that are matched correctly, such as a buried surface hoar layer, or a weak snowpack base with an embedded crust.

3.1. *Computing representative profiles*

Herla et al. (2022a) extended the functionality of the previous algorithms with a snow profile averaging algorithm based on a data science method called DTW Barycenter Averaging (DBA, Petitjean et al., 2011) that can efficiently be applied to large sets of profiles (Figure 2a, c). Computing a representative profile for every day of the season allows for monitoring the evolution of the average conditions within the region (Figure 3a). The algorithm is designed to match the median snow height of the individual profiles (black solid line) and preserve other useful metrics, such as the median thickness of new snow (dashed line) or the median depths of specific weak layers (dotted lines). The time series of the representative profile depicted in Figure 3 clearly identifies all layers that were of operational concern for forecasters during the season.

3.2. *Extracting distributions of layer properties*

By matching all layers from all individual profiles against the representative profile, the DBA algorithm enables the computation of insightful spatial distributions of layer characteristics. This includes but is not limited to layer instability, load on spe-

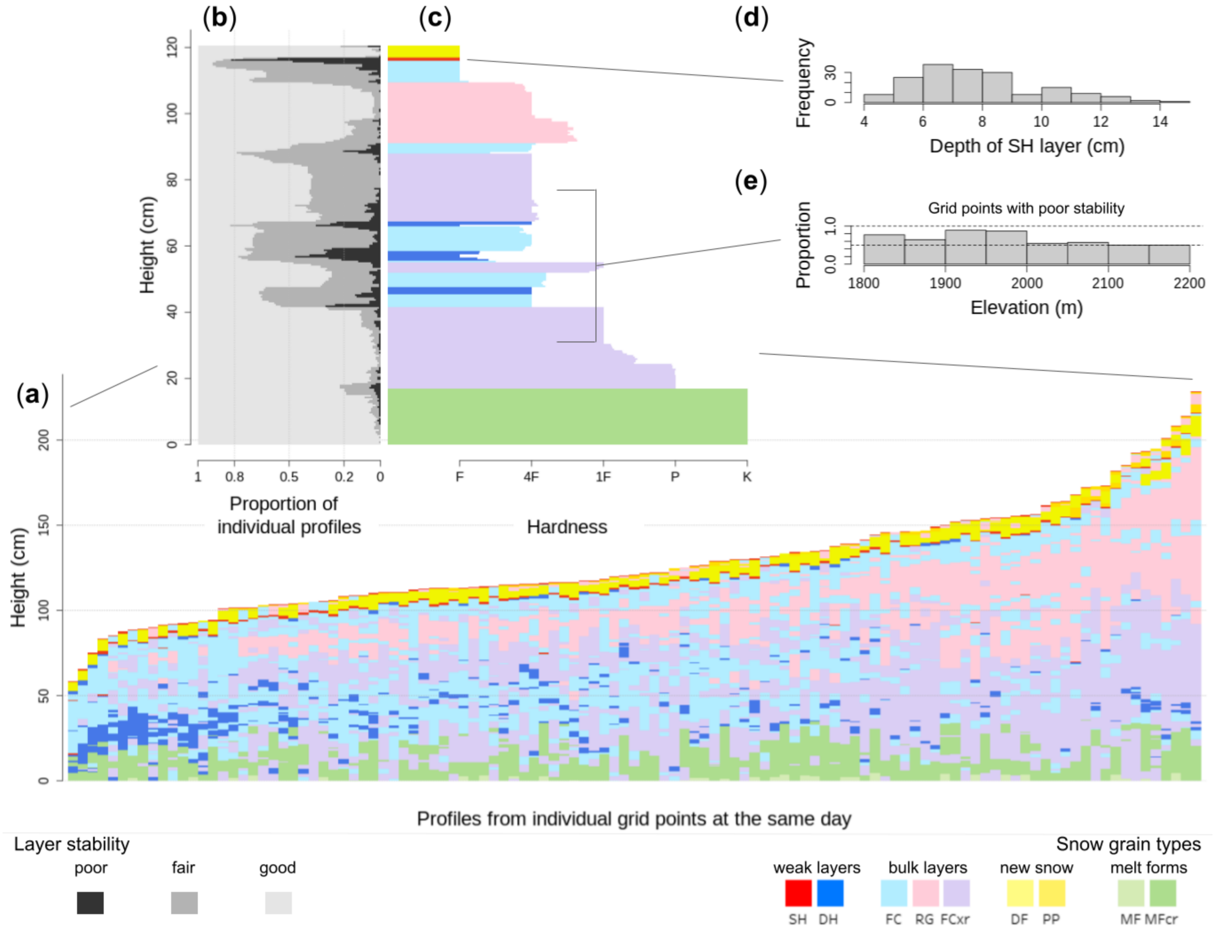

Figure 2: (a) Simulated snow profiles from 112 model grid points in an avalanche forecast region on the same day, (c) that are summarized by the representative profile. The representative profile provides access to distributions of layer and profile properties, such as, for example, (b) the distribution of layer instabilities, (d) the depth distribution of a SH layer that is starting to get buried, and (e) the elevation distribution of those profiles that contain weak layers in mid-snowpack.

cific layers, or elevation bands of specific problems (Figure 2b, d, e). Overplotting the representative time series with information from the underlying instability distribution allows users to quickly absorb which layers caused widespread instabilities and when these instabilities were most likely to occur (Figure 3b).

4. CLUSTERING OF PROFILES

Since avalanche warning services increasingly publish forecasts for dynamic region boundaries that vary depending on the conditions, snowpack clustering is an operationally attractive, but yet infant area of research (Herla et al., 2021; Reuter et al., 2023; Horton et al., in preparation). The snow profile alignment algorithm and similarity measure highlighted in Section 2 directly enable the clustering of snow profiles by providing the link to existing clustering methods. Horton et al. (in preparation) extended this approach into an elegant strategy for spatial fuzzy clustering that partitions forecast sub-regions into groups with similar snowpack conditions.

5. DISCUSSION AND CONCLUSIONS

The highlighted algorithms aim to form a comprehensive suite of tools for effectively processing snowpack simulations in support of avalanche forecasting by facilitating comparing, summarizing, and grouping of snow profiles. The common objective that connects the developed methods is the desire of making the rich output of large-scale distributed snowpack simulations accessible in ways that are relevant and informative for forecasters. Recognizing the importance of the layered structure of the snowpack for avalanche formation, all methods preserve the stratigraphy information and enable users for the first time to efficiently access and explore this information in both space and time instead of being limited to one dimension or summary statistics of snowpack bulk properties (e.g., height of snowpack, amount of new snow). Since all methods were inspired by existing forecaster practices, the developed tools enable forecasters to tap into the wealth of simulated information in familiar ways and at relatively low cognitive cost. Moreover, by

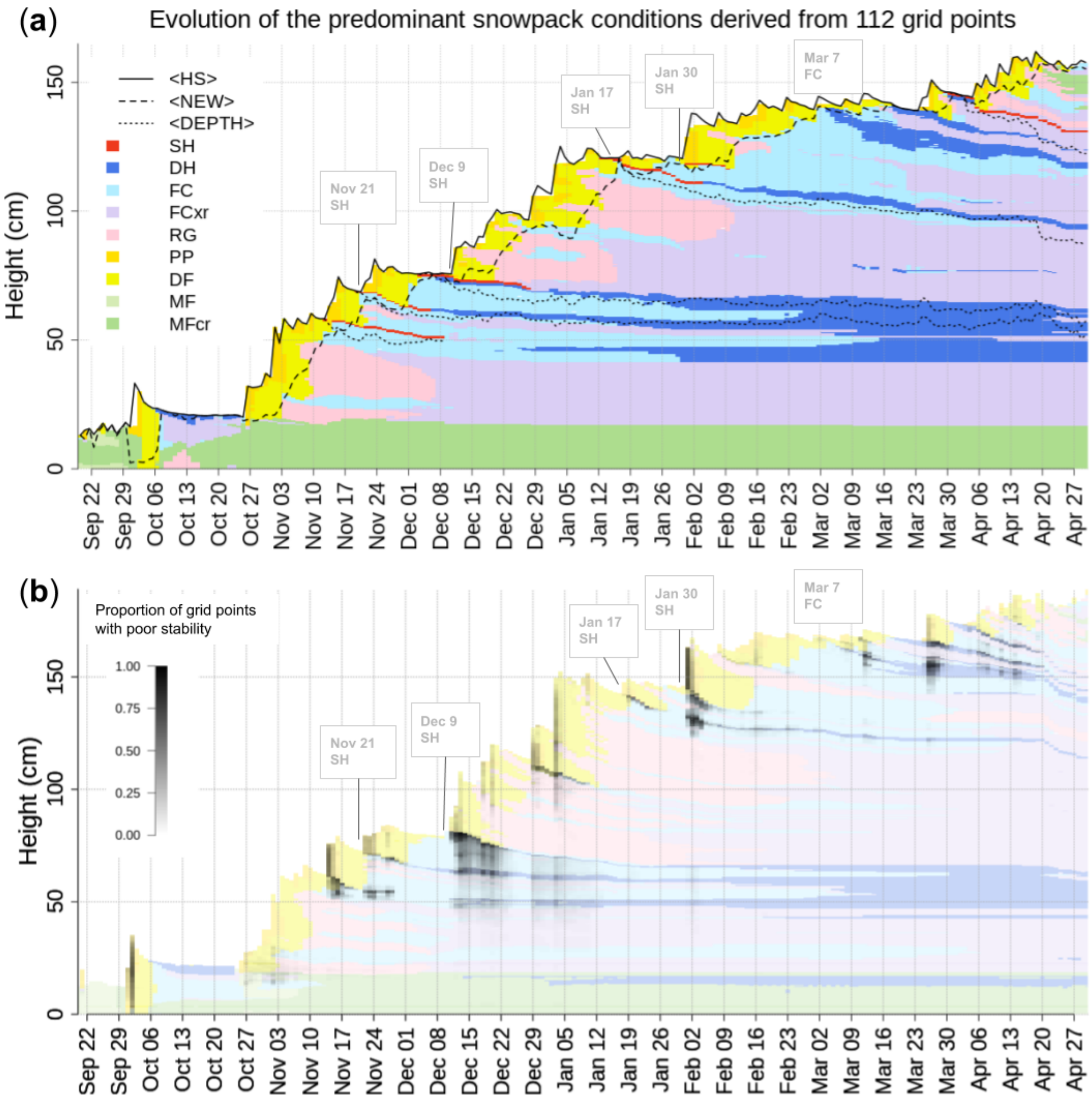

Figure 3: (a) Time series of the representative snow profile that illustrates the evolution of the space-averaged snow stratigraphy (visualized by grain types). The algorithm captures the median total snow height (solid line, <HS>), the median amount of new snow (dashed line, <NEW>), and the median depth of several persistent weak layers (dotted lines, <DEPTH>). Five labels highlight all weak layers that were of operational concern for forecasters during the season. (b) The time series of the average profile overplotted with the distribution of grid points that contain layers with poor stability.

providing detailed access to snowpack simulations at the layer level, the tools invite scrutiny and allow users to identify and trace potential errors in the simulations, which ensures transparency and aims to facilitate adoption.

We imagine the integration of snowpack simulations into operational avalanche forecasting as follows. Traditionally, avalanche risk management operations, such as public warning services or backcountry guiding operations, obtain vital information about the snowpack through manual snow pit observations at select point locations. Simulated snow profiles across a forecast region or operating area can sample the snowpack conditions similarly to field observations, but with a higher spatiotemporal coverage and independent from external circumstances. Our methods can group the simulated profiles according to similar conditions, potentially uncovering different avalanche problems (see also Herla et al., 2023b). The numerical grouping can quantify the prevalence of those conditions, and the conditions can be linked to their specific location (and eventually also elevation and aspect). The different conditions can then be summarized by computing representative profiles, and forecasters can test their hypotheses by querying detailed distributions of layer characteristics. In this vision, the simulated data represents an additional, independent information source that complements existing data sources. Under certain circumstances, the simulations might prompt targeted field observations to verify simulations and assessments alike. Lessons from the use of snowpack simulations in forecast operations and insights from recent research projects (Herla et al., 2023a, 2023b) have shown that being able to continuously compare model output with human observations and assessments is crucial for a meaningful integration into operational avalanche forecasting. This helps us not only interpret individual predictions, but also gain insights into the strengths and weaknesses of the different data sources and how they can best complement each other.

The development of the presented methods was embedded in a larger effort to increase the adoption of snowpack simulations in the avalanche safety community in Canada based on feedback from forecasters (Floyer et al., 2016; Morin et al., 2020) and insights from the history of numerical weather forecasting (Benjamin et al., 2019). Prototypes of the numerical tools have been iteratively integrated into the experimental and eventually operational snowpack simulation infrastructure at Avalanche Canada, the primary public avalanche warning agency in Canada (Horton et al., 2023). Furthermore, the algorithms promise to support their vision of expanding existing interactive dashboards that compare different data sources and facilitate auto-generated snowpack summaries (Horton et al., 2023).

The methods presented can be thought of as numerical infrastructure that facilitates effective and insightful mining of snowpack simulations. While the current focus of snowpack modeling at Avalanche Canada is on the regional scale, the methods are independent from the spatial resolution of the model configuration and can readily be applied to finer resolutions. Furthermore, the methods are equally suited for processing other sets of simulations, such as ensemble simulations that use different initial conditions and/or model formulations. By quantifying uncertainty through probabilistic predictions, ensemble modeling revolutionized weather forecasting (Leutbecher and Palmer, 2008; Benjamin et al., 2019), but has only found limited application in snowpack modeling for avalanche forecasting so far (Vernay et al., 2015; Lafaysse et al., 2017). We attribute this, at least partially, to a lack of numerical tools for meaningfully mining the immense volume of stratigraphic information generated by ensemble simulations. Hence, the presented methods create promising new opportunities that go beyond their current application.

To make our methods accessible to the community, we have published them as part of the open-source software packages `sarp.snowprofile` (Horton et al., 2020a) and `sarp.snowprofile.alignment` (Herla et al., 2022b), which can be downloaded and used for free using the R language and environment for statistical computing. The detailed documentation and code vignettes provided in these packages explain the rich functionality and allow for reproducing all visuals. Once the manuscript mentioned in Section 4 has been published, the clustering method will also be made accessible in the packages.

ACKNOWLEDGEMENT

We are grateful to the Coast Salish peoples on whose traditional and unceded territories Simon Fraser University and our research program resides. This research was supported by the Government of Canada Natural Sciences and Engineering Research Council via the NSERC Industrial Research Chair in Avalanche Risk Management at Simon Fraser University. The industry partners of the Research Chair include Canadian Pacific Railway, HeliCat Canada, Mike Wiegele Helicopter Skiing, and the Canadian Avalanche Association. The re-search program receives additional support from Avalanche Canada and the Avalanche

Canada Foundation.

References

Benjamin, S. G., Brown, J. M., Brunet, G., Lynch, P., Saito, K., and Schlatter, T. W.: 100 Years of Progress in Forecasting and NWP Applications, Meteorological Monographs, 59, 13.1–13.67, doi:10.1175/amsmonographs-d-18-0020.1, 2019.

Floyer, J., Klassen, K., Horton, S., and Haegeli, P.: Looking to the 20s: Computer-Assisted Avalanche Forecasting in Canada, in: Proceedings of the 2016 International Snow Science Workshop, Breckenridge, CO, USA, International Snow Science Workshop, Breckenridge, Colorado, 2016.

Hagenmuller, P. and Pilloix, T.: A New Method for Comparing and Matching Snow Profiles, Application for Profiles Measured by Penetrometers, Frontiers in Earth Science, 4, doi:10.3389/feart.2016.00052, 2016.

Herla, F., Horton, S., Mair, P., and Haegeli, P.: Snow profile alignment and similarity assessment for aggregating, clustering, and evaluating of snowpack model output for avalanche forecasting, Geoscientific Model Development, 14, 239–258, doi:10.5194/gmd-14-239-2021, 2021.

Herla, F., Haegeli, P., and Mair, P.: A data exploration tool for averaging and accessing large data sets of snow stratigraphy profiles useful for avalanche forecasting, The Cryosphere, 16, 3149–3162, doi:10.5194/tc-16-3149-2022, 2022a.

Herla, F., Horton, S., Mair, P., and Haegeli, P.: sarp.snowprofile.alignment, URL `https://cran.r-project.org/web/packages/sarp.snowprofile.alignment/index.html`, 2022b.

Herla, F., Haegeli, P., Horton, S., and Mair, P.: A Large-scale Validation of Snowpack Simulations in Support of Avalanche Forecasting Focusing on Critical Layers, EGUsphere [preprint], doi:10.5194/egusphere-2023-420, 2023a.

Herla, F., Haegeli, P., Horton, S., and Mair, P.: A quantitative module of avalanche hazard—Comparing forecaster assessments of avalanche problems with information from distributed snowpack simulations, in: Proceedings of the International Snow Science Workshop Bend, OR, 2023b.

Horton, S., Herla, F., and Haegeli, P.: An R package for snow profile analysis and visualization, in: Proceedings of the 2020 Virtual Snow Science Workshop VSSW, Fernie, BC, Canada, 2020a.

Horton, S., Nowak, S., and Haegeli, P.: Enhancing the operational value of snowpack models with visualization design principles, Natural Hazards and Earth System Sciences, 20, 1557–1572, doi:10.5194/nhess-20-1557-2020, 2020b.

Horton, S., Klassen, K., Haegeli, P., Floyer, J., and Helgeson, G.: Adopting snowpack models into an operational forecasting program: successes, challenges, and future outlook, in: Proceedings of the International Snow Science Workshop Bend, OR, Bend, OR, 2023.

Horton, S., Herla, F., and Haegeli, P.: Clustering simulated snow profiles to produce avalanche forecast regions, in preparation.

Lafaysse, M., Cluzet, B., Dumont, M., Lejeune, Y., Vionnet, V., and Morin, S.: A multiphysical ensemble system of numerical snow modelling, Cryosphere, 11, 1173–1198, doi:10.5194/tc-11-1173-2017, 2017.

Leutbecher, M. and Palmer, T. N.: Ensemble forecasting, Journal of Computational Physics, 227, 3515–3539, doi:10.1016/j.jcp.2007.02.014, 2008.

Morin, S., Fierz, C., Horton, S., Bavay, M., Dumont, M., Hagenmuller, P., Lafaysse, M., Mitterer, C., Monti, F., Olefs, M., Snook, J. S., Techel, F., Van Herwijnen, A., and Vionnet, V.: Application of physical snowpack models in support of operational avalanche hazard forecasting : A status report on current implementations and prospects for the future, Cold Regions Science and Technology, 170, 1098–1107, doi:10.1016/J.COLDREGIONS.2019.102910, 2020.

Nowak, S.: Designing for ambiguity in sensemaking: Visual analytics in risk analysis and prediction, Ph.D. thesis, Simon Fraser University, 2023.

Petitjean, F., Ketterlin, A., and Gançarski, P.: A global averaging method for dynamic time warping, with applications to clustering, Pattern Recognition, 44, 678–693, doi:10.1016/j.patcog.2010.09.013, 2011.

Reuter, B., Hagenmuller, P., and Eckert, N.: Snow and avalanche climates in the French Alps using avalanche problem frequencies, Journal of Glaciology, 190, 1–13, doi:10.1017/jog.2023.23, URL `https://doi.org/10.1017/jog.2023.23`, 2023.

Sakoe, H. and Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing, 26, 43–49, doi:10.1109/tassp.1978.1163055, 1978.

Schweizer, J., Kronholm, K., Jamieson, J. B., and Birkeland, K. W.: Review of spatial variability of snowpack properties and its importance for avalanche formation, Cold Regions Science and Technology, 51, 253–272, doi:10.1016/j.coldregions.2007.04.009, 2007.

Schweizer, J., Reuter, B., Van Herwijnen, A., and Gaume, J.: Avalanche Release 101, in: Proceedings of the 2016 International Snow Science Workshop, Breckenridge, CO, USA, URL `https://arc.lib.montana.edu/snow-science/item/2235`, 2016.

Shneiderman, B.: Eyes have it: a task by data type taxonomy for information visualizations, in: IEEE Symposium on Visual Languages, Proceedings, pp. 336–343, IEEE, doi:10.1016/b978-155860915-0/50046-9, 1996.

Statham, G., Haegeli, P., Greene, E., Birkeland, K. W., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: A conceptual model of avalanche hazard, Natural Hazards, 90, 663–691, doi:10.1007/s11069-017-3070-5, 2018.

Vernay, M., Lafaysse, M., Mérindol, L., Giraud, G., and Morin, S.: Ensemble forecasting of snowpack conditions and avalanche hazard, Cold Regions Science and Technology, 120, 251–262, doi:10.1016/j.coldregions.2015.04.010, 2015.