Brief communication: A numerical tool for averaging large data sets of snow stratigraphy profiles useful for avalanche forecasting

Florian Herla[1], Pascal Haegeli[1], and Patrick Mair[2]

[1]Simon Fraser University, Burnaby, BC, Canada
[2]Harvard University, Cambridge, MA, USA

Correspondence: Florian Herla (fherla@sfu.ca)

Abstract. Snowpack models can provide detailed insight about the evolution of the snow stratigraphy in ways that is not possible with direct observations. However, the lack of suitable data aggregation methods currently prevents the effective use of the available information, which is commonly reduced to bulk properties and summary statistics of the entire snow column or individual grid cells. This is only of limited value for operational avalanche forecasting. To address this challenge, we present
5 an averaging algorithm for snow profiles that can effectively synthesize large numbers of snow profiles into a meaningful overall perspective of the existing conditions. Notably, the algorithm enables compiling of informative summary statistics and distributions of snowpack layers, which creates new opportunities for presenting and analyzing distributed and ensemble snowpack simulations.

1   Introduction

10   The layered nature of the snowpack is a necessary condition for the formation of snow avalanches (e.g., Schweizer et al., 2003, 2016; Reuter and Schweizer, 2018), and information about the snow stratigraphy is crucial for developing a meaningful understanding of existing avalanche hazard 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
15   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. In the field, local observations are used to validate and localize the regional understanding of the conditions. As the season progresses,
20   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 limited numbers of 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

25   (Horton et al., 2020b), 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 aggregation algorithms have the potential to allow avalanche forecasters to make better use of large scale snowpack simulations. Inspired by Hagenmuller and Pilloix (2016), who introduced *Dynamic Time Warping*, a well establish algorithm for measuring similarity between two sequences, to the snow community to analyze one-dimensional hardness profiles, Herla et al. (2021) developed a set of nu-
30   merical algorithms for comparing multidimensional, mixed data type snow profiles also based on Dynamic Time Warping. However, their medoid approach for the computation of a representative profile has substantial limitations. Since the medoid is simply the profile within a given group that is most similar to all other profiles, it does not actually aggregate the available information and therefore does not accurately represent the snowpack features that exist within the group. Furthermore, it is not suited for tracking average conditions over time as the medoid within a data set might differ between time steps resulting
35   in a disjointed and difficult to interpret time series. Finally, medoid calculations are computationally costly and thus only of limited applicability in operational contexts. All these reasons make the medoid aggregation approach unsuitable for avalanche forecasting.

The objective of this brief communication is to introduce an averaging algorithm for snow profiles that extends the snow profile processing tools of Herla et al. (2021) with a global averaging method that is based on the approach proposed by
40   Petitjean et al. (2011). Our goal is to compute an average snow profile that provides a quick and familiar overview of the predominant snowpack features that are captured within large sets of profiles in a way that is useful for avalanche forecasting. To be useful for operational avalanche forecasting and support existing practices, our averaging approach should highlight information relevant for assessing avalanche hazard concealed in large volumes of snowpack simulations and facilitate the tracking of space-averaged snowpack conditions over time. Analyzing large volumes of snowpack simulations in these novel
45   ways will significantly improve their accessibility to data features and data views that are relevant for avalanche forecasting. The algorithm described in this paper has been implemented in the open source R package `sarp.snowprofile.alignment` (Herla et al., 2022a), and is freely available to researchers and practitioners.

2   Description of the snow profile averaging algorithm

Our approach is based on Petitjean et al. (2011) who developed an averaging method specifically for sequential data called
50   *Dynamic Time Warping Barycenter Averaging* (DBA). Unlike the medoid approach used for snow profile aggregation by Herla et al. (2021), the average sequence derived with DBA consists of an entirely new sequence that represents the notion of an average of all individual sequences. This makes this approach more suitable for snow profile applications because it actually provides an average perspective of the conditions. In addition, DBA also uses considerably less computation time since DBA does not rely on pairwise comparisons across the entire data set, but rather uses an iterative approach that starts with an initial
55   condition that acts as a reference for the calculations. In our implementation of the approach for snow profiles, all layers from all profiles within the data set are matched against that reference profile to retrieve sets of corresponding layers. The corresponding layers are then averaged, and the averaged layer properties are used to update the reference profile. This process
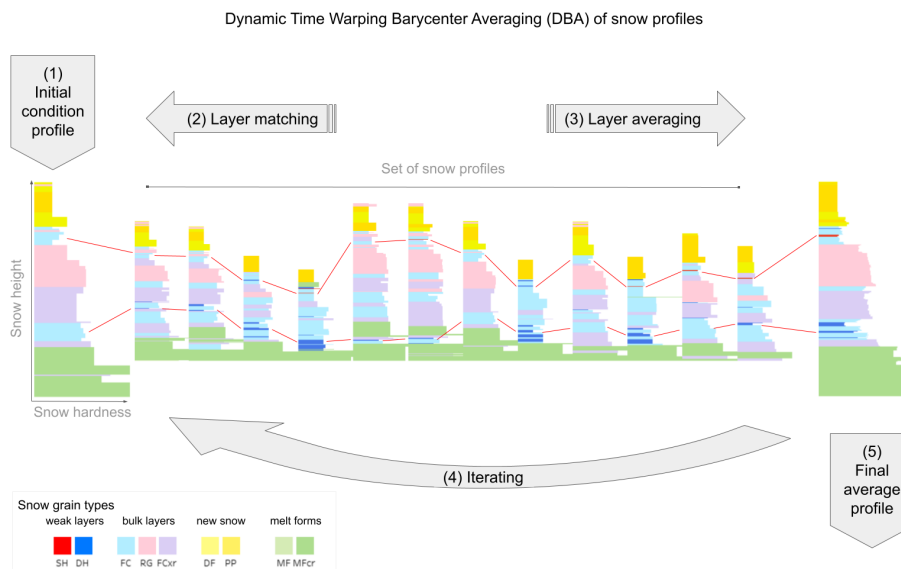
Figure 1. Conceptual flowchart demonstrating how Dynamic Time Warping Barycenter Averaging (DBA) can be applied to snow stratigraphy profiles. While the figure emphasizes steps (2) and (3) for the corresponding layers of two snowpack features, these steps are actually carried out for all layers.

of matching and updating the reference profile is repeated several times until the reference profile does not change substantially anymore. At this point, the reference profile represents the average profile of the data set. Figure 1 illustrates the workflow of

60 the algorithm based on a small set of profiles. The red lines highlight how the corresponding snowpack layers in the profiles are matches to the layers in the reference profile.

The layer matching step is implemented as described in Herla et al. (2021) with one exception. While the original approach for the alignment required all snow profiles to be rescaled to identical snow heights before their alignment, this requirement has been removed for the updated implementation presented here. All profiles can therefore be aligned on their native height

65 grids. Since there are meaningful use cases for both approaches, the updated version of our R package allows users to choose between the two options. We have found alignments on the native height grid to be of slightly higher quality than alignments of rescaled profiles, but more importantly, alignments on native height grids are easier to interpret.

Our approach for averaging multidimensional, mixed data type snow stratigraphy layers starts with summarizing *grain types* because it is a fundamental layer characteristic and plays a critical role in snow profile processing tools described by Herla

70 et al. (2021). Following the approach of Huang (1998) for summarizing categorical variables in k-means clustering, we average corresponding layers by first calculating the predominant grain type (i.e., the grain type mode). The average of other (ordinal or numerical) layer properties are then expressed by the median properties of the layers of the predominant grain type.

Since thin weak layers are particularly important when analyzing snow profiles for avalanche forecasting, our algorithm includes a flexible approach for ensuring critical weak layers are meaningfully highlighted in the averaged profiles even if their

75 grain types are not necessarily consistent and the most prevalent. Instead of averaging weak layers solely based on their grain type mode, we implemented a subroutine that can be used to label *layers of interest* based on users' particular needs and data availability. While layers of interest will typically be weak layers, they can also be other layers such as crusts. For example, the label can be based on grain type classes (e.g., all persistent grain types), or it can include additional relevant measures like stability thresholds (e.g., threshold sums; Schweizer and Jamieson, 2007; Monti and Schweizer, 2013). If the majority of

80 corresponding layers is labeled as layers of interest, the resulting averaged layer properties are the median properties of all labeled layers regardless of the actual grain types. Note that this approach still eliminates weak layers that only occur in a few profiles but might nevertheless be relevant for avalanche considerations. To address this issue, users can either query the profile set for the list of weak layers that are not included in the average profile, or they can change the hyperparameter that specifies the occurrence frequency threshold for labeled layers to be included in the average profile away from the default 50%. This

85 ensures that the final average profile represents the predominant and/or relevant snowpack features and that layer properties are internally consistent.

While the stochastic and iterative nature of the DBA approach is responsible for its computational efficiency[1], it also makes it sensitive to initial conditions. We turned this potential weakness into an opportunity to steer the averaging algorithm in a more informative direction by choosing initial condition profiles strategically. Since it is important that relevant thin weak layers

90 are captured by the average profile, we select initial profiles with an above average number of weak layers in as many depth ranges as possible. While weak layers that do not exist in the rest of the data set are automatically averaged out during the first iteration, more prevalent weak layers are more likely to get matched and included in the final average profile. Additionally, we rescale the initial condition profile to the median snow depth to maximize the number of meaningful layer matches and ensure that the final profile represents the snow depth distribution of the data set in a meaningful way. To avoid exaggerated rescaling,

95 we only select initial condition profiles whose total snow depth is within the interquartile range of the snow depth distribution.

After several DBA iterations the reference profile will only change marginally. To assess the iterative changes between the reference profiles and stop the iteration cycle, we use a similarity measure for snow profiles analogous to Herla et al. (2021). If the similarity between the reference profiles of two subsequent iterations is beyond a certain threshold, the algorithm is stopped. Reaching a similarity threshold of 0.99 usually does not take more than five iterations. However, if computational speed is of

100 the essence, a threshold of 0.90 that is attained twice in consecutive iterations yields comparable results. If the algorithm is started with several initial conditions, the best average profile among the different realizations is chosen by converting the similarity measure between the reference profile and the profile set to a mean squared error. The average profile with the lowest mean squared error is chosen as the final realization.

Our testing has shown that these rules combined with rerunning the algorithm with different initial conditions (as suggested

105 by Petitjean et al., 2011) consistently produce reasonable average snow profiles suitable for avalanche forecasting.

---

[1]While computing the medoid for a given profile set of length N requires $\mathcal{O}(N^2)$ profile alignments, the DBA approach requires $\mathcal{O}(N \cdot I \cdot IC)$, where I is the number of iterations and IC is the number of different initial conditions.
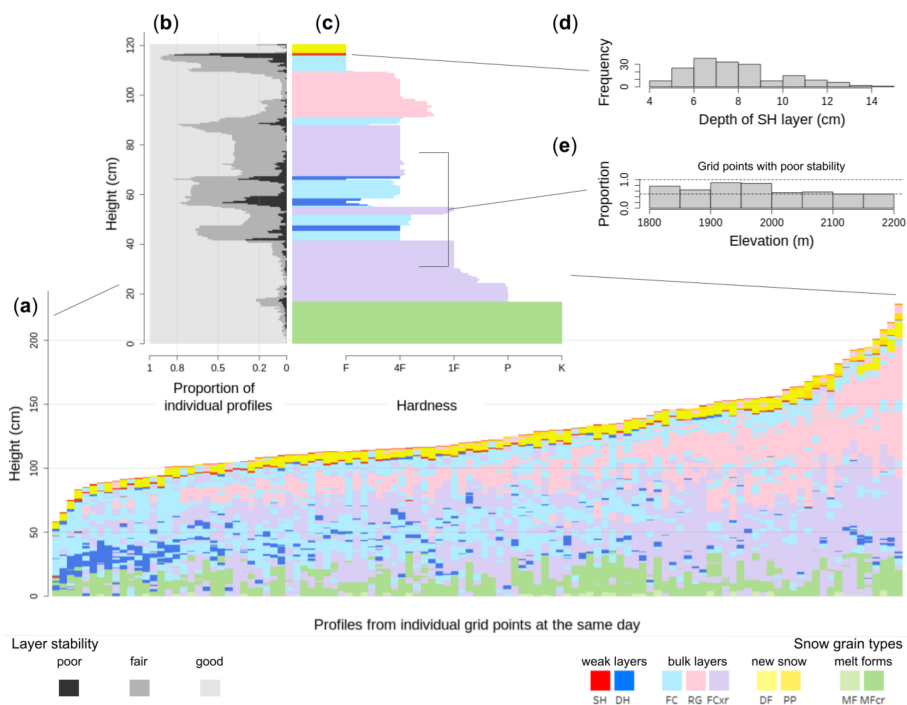
Figure 2. (a) Snow profile set from an avalanche forecast region, (c) which is summarized by the average profile. The average profile provides access to distributions of layer and profile properties, such as e.g. (b) the distribution of layer stabilities derived from threshold sums, (d) the depth distribution of a SH layer that is starting to get buried, (e) the elevation distribution of the proportion of profiles that contain layers with poor stability in mid snowpack.

3 Applications

3.1 An inconspicuous asset: Overview first, details on demand

Being able to efficiently calculate average snow profiles allows synthesizing large volumes of snow profile data into an overall perspective of existing conditions. Figure 2a shows simulated profiles from 112 model grid points within a forecast region
110 ordered by snow depth, and Fig. 2c illustrates how an average profile summarizes the characteristics of the entire profile set in a meaningful way. The average profile highlights three distinct weak layers buried in the mid to lower snowpack with a thick and consolidated slab above. Close to the surface, the snow is loose, and a surface hoar layer is starting to get buried. This overview provides an insightful overview and important context for interpreting the layer sequences of individual profiles if more specific information is required.
115 However, there is more to the averaging algorithm than just providing a graphic representation of the average condition across a forecast area. Since all individual layers are matched against the average profile, each of the averaged layers can be traced throughout the entire data set. The average profile therefore acts as a navigation tool that connects all layers and

enables the tracking of layers across space as well as the computation of distributions of layer and profile properties. Hence, the average profile embodies the important and broadly used "overview first, details on demand" data visualization principle

120 that was first proposed by Shneiderman (1996). To illustrate this capability, Fig. 2b shows the stability distribution of each layer using the *threshold sums* (Schweizer and Jamieson, 2007), which we classified into three categories (poor: ≥ 5; fair: ≥ 4; good: < 4). About 50 % of all profiles in the data set promote structural instability along what is predominantly a DH layer in mid snowpack, and almost all profiles contain a surface hoar layer with poor stability that will be concerning upon deeper burial. To dig deeper, it is easy to retrieve the depth distribution of that shallow surface hoar layer and confirm that it

125 is currently buried mainly by 7 cm and up to 15 cm of new snow (Fig. 2d). Note that similar charts can be computed for other stability indices or any other layer properties available in the profiles, and that calculating distributions on subsets of layers with particular properties is also straightforward. For example, in Fig. 2e we show the elevation distribution of the proportion of profiles that contain layers with poor stability in mid snowpack. While Fig. 2a suggests that these weak layers mainly exist in shallower profiles, Fig. 2e also confirms that these layers are more likely to be found in lower elevations. In summary, the

130 average profile enables efficient and user-friendly access to large volumes of snowpack simulations in support of answering critical questions of avalanche forecasting like *Which weak layers exist?*, and *How distributed and sensitive to triggering are they?*.

The largest source of error in averaging snow profiles originates from the layer matching step. As discussed in Herla et al. (2021), applying the matching algorithm to highly diverse data sets in an unsupervised manner will inevitably include some

135 alignment inconsistencies and errors since a single hyperparameter setting can naturally not be optimal for the full range of observed conditions. While this is negligible for high-level overview results, it is important to remember that the detailed distributions of layer and profile characteristics do not represent the truth but need to be evaluated in light of this source of uncertainty. Poor layer matches between snow profiles are more likely if their snow depths differ considerably or their layer sequences show very few common patterns. It is therefore users' responsibility to judge whether it is meaningful to compute an

140 average profile for a specific set of profiles, and we advise to allow for more margin of error the more diverse the snow profiles are.

3.2 A representation of the predominant conditions over the course of the season

Since snowpack and avalanche conditions evolve continuously throughout a season, being able to effectively present the evolution of the predominant conditions across forecast regions is critical for supporting forecasters' assessment process and mental

145 models of the existing conditions. Our averaging algorithm can be used to represent a temporal perspective of the average conditions in a consistent way by looping it over the course of the season and using the average profile from the previous day as initial condition.

The height of the snowpack grows by matching the current day individual profiles against the previous day average profile in an open-ended bottom-up alignment approach (for more details, see Herla et al., 2021). This allows new snow layers that are

150 not present in the previous average profile to get stacked on top of the old snow column if more than 50 % of the grid points
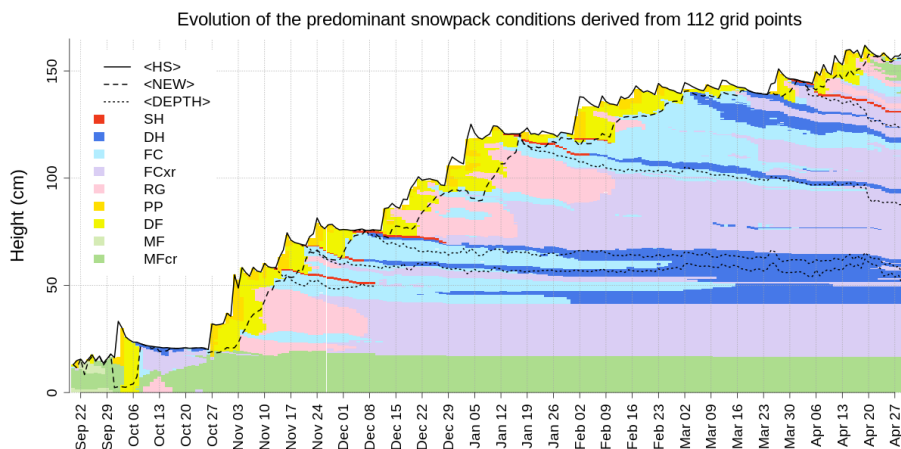
Figure 3. Time series of the average snow profile that illustrates the space-averaged evolution of the snow stratigraphy (visualized by snow grain types). The algorithm captures the median total snow height (solid line, <HS>), the median amount of new snow (dashed line, <NEW>), and the median depth of several persistent weak layers (dotted lines, <DEPTH>).

contain new layers. The new snow amounts in the updated average profile therefore represent the median new snow amounts reported in the profile set. The same effect allows for the growing of thin weak layers at the snow surface.

To capture settlement and melting, the average profile is rescaled after each day if it exceeds the median snow height. While layers that were added at the same day remain unchanged, the upper part of the old snow column is rescaled with a uniform scaling factor determined by the snow height offset at that day. This models the settlement of freshly buried layers while avoiding unrealistic settlements in the deeper layers at the same time. Since the average profile encapsulates information about the median depths of all layers, the extent of the rescaled column is adjusted each day, so that the resulting layers indeed align with their median depths.

Applying our averaging algorithm in this temporal fashion to the same simulated profiles from a forecast region yields a continuous time series of averaged profiles that has a very similar look and feel as the time series of the snowpack evolution at individual grid points but contains information from the entire data set (Fig. 3). To appreciate the capabilities of the algorithm to capture important summary statistics, study the black lines in Fig. 3: the solid line represents the median snow height and follows precisely the height of the average profile; the area between the solid line and the dashed line represents the median thickness of new snow and is well captured by the corresponding layers of PP and DF shown in yellow; and finally, the dotted lines represent the median depths of several weak layers, which align closely with the red and blue colors highlighting the presence of SH and DH layers in the average profile.

To produce a realistically looking time series of average profiles (Fig. 3), several hyperparameters and algorithmic adjustments are necessary. While these adjustments come with the benefits of temporal consistency and computational efficiency[2],

---

[2]Averaging about 100 grid points with daily profile sampling along the season takes about one hour on a local computer, mostly dependent on the snow depths of the profiles.

they increase the possibility of minor unrealistic features. One inconvenience, for example, can occur when new snow falls
from one day to the next, and surface hoar forms on top of the new snow. In this situation, the surface hoar layer is often-
times only captured by the average time series upon burial instead of at the snow surface. Furthermore, since the time series
is designed to represent several different summary statistics of the profiles, cases exist when these summary statistics are not
completely internally consistent. For example, the surface hoar layer that got buried on April 6 in Fig. 3 happens to primarily
exist in profiles that subsequently received above average snowfall amounts (not shown). Consequently, its median layer depth
is deeper than what can be represented in the average time series given an accurate representation of the median snowfall
amounts and the median snow height. Additionally to that effect, the lower elevation profiles in the data set transitioned into
an isothermal state around April 20, which replaced their original layer sequences with melt forms (not shown). As a result,
the median layer depths of the January 17 and April 6 surface hoar layers increased abruptly, which explains the abrupt drop
in the dotted lines in late April.

4   Conclusions

The two snow profile averaging algorithms described in this brief communication continue a line of development that aims to
make snowpack simulations more accessible and relevant to avalanche warning agencies and practitioners. Building on the tools
introduced by Herla et al. (2021), the presented methods support the analysis of large volumes of snowpack simulations along
both space and time by (i) providing quick summary visualizations that help assess the evolution of snow depths, new snow
amounts, weak layer and slab combinations, and by (ii) facilitating retrieval of various summary statistics and distributions of
layer and profile characteristics.

Without appropriate tools, the operational processing and analysis of simulated stratigraphic information has mainly been
restricted to individual grid points, or along either one dimension of space/time. This led to configurations of snowpack simula-
tions in support of avalanche forecasting to primarily be station-based or semi-distributed (Morin et al., 2020), and approaches
for evaluating distributed and/or ensemble simulations have been limited to bulk properties and summary statistics of the snow-
pack (Morin et al., 2020; Vernay et al., 2015). By providing summary statistics of layers instead of the entire snow column, our
algorithms provide new opportunities for how distributed or ensemble snowpack simulations can be validated and exploited.
These new ways of mining available and relevant information aim to inspire new approaches for the operational use of dis-
tributed snowpack simulations that are more useful for avalanche forecasting. Furthermore, synthesizing snow profile sets into
a representative perspective provides an important and necessary step towards clustering snow stratigraphy information.

While our algorithms open the door for powerful analysis of large data sets of snowpack simulations, there are a number of
limitations that should be considered when applying them. It is important to remember that our algorithms are not designed to
extract true summaries (i.e., precise average grain size of a particular layer) but rather to facilitate *meaningful explorations* of
data sets that are too big for human forecasters to analyze manually.

Even though the impetus of the development of our algorithms was avalanche forecasting, they might also be of use for other cryospheric researchers interested in the examination of large data sets of snow profiles. Our general approach might also have application for the processing of profiles and time series in other geophysical contexts.

Code availability. The presented algorithms are implemented in the R language and environment for statistical computing (R Core Team, 2020) as part of the package `sarp.snowprofile.alignment` (version 1.1.3). The open source package is available from the Comprehensive R Archive Network at https://cran.r-project.org/package=sarp.snowprofile.alignment (Herla et al., 2022a). A static version of the package as well as an annotated demo script to reproduce the figures in this paper are accessible from a permanent repository (Herla et al., 2022b). Our package builds upon the open source packages `dtw` (https://dynamictimewarping.github.io/, last access: 12 January 2022, by Giorgino, 2009), which contains the Dynamic Time Warping implementations, and `sarp.snowprofile` (https://cran.r-project.org/package=sarp.snowprofile, last access: 12 January 2022, by Horton et al., 2020a), which contains basic functionality for reading and manipulating snow profile data.

Author contributions. All authors conceptualized the research that FH conducted; FH designed, implemented, and tested the software and wrote the initial manuscript; all authors contributed to reviewing and editing the manuscript; PH supervised the research and acquired the funding.

Competing interests. The authors declare that they have no conflict of interest.

References

Giorgino, T.: Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package, J. Stat. Softw., 31, https://doi.org/10.18637/jss.v031.i07, 2009.

Hagenmuller, P. and Pilloix, T.: A New Method for Comparing and Matching Snow Profiles, Application for Profiles Measured by Penetrometers, Front. Earth Sci., 4, https://doi.org/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, Geosci. Model Dev., 14, 239–258, https://doi.org/10.5194/gmd-14-239-2021, 2021.

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

Herla, F., Mair, P., and Haegeli, P.: An averaging algorithm for snow profiles—Data and Code, https://doi.org/10.17605/OSF.IO/7MA6G, 2022b.

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, Nat. Hazards Earth Syst. Sci., 20, 1557–1572, https://doi.org/10.5194/nhess-2019-344, 2020b.

Huang, Z.: Extensions to the k-means algorithm for clustering large data sets with categorical values, Data Min Knowl Disc, 2, 283–304, 1998.

Monti, F. and Schweizer, J.: A relative difference approach to detect potential weak layers within a snow profile, in: Proceedings of the 2013 International Snow Science Workshop, Grenoble, France, pp. 339–343, https://arc.lib.montana.edu/snow-science/item.php?id=1861, 2013.

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 Reg. Sci. Technol., 170, 1098–1107, https://doi.org/10.1016/J.COLDREGIONS.2019.102910, 2020.

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

R Core Team: R: A Language and Environment for Statistical Computing, https://www.r-project.org/, 2020.

Reuter, B. and Schweizer, J.: Describing Snow Instability by Failure Initiation, Crack Propagation, and Slab Tensile Support, Geophys. Res. Lett., 45, 7019–7027, https://doi.org/10.1029/2018GL078069, 2018.

Schweizer, J. and Jamieson, J. B.: A threshold sum approach to stability evaluation of manual snow profiles, Cold Reg. Sci. Technol., 47, 50–59, https://doi.org/10.1016/j.coldregions.2006.08.011, 2007.

Schweizer, J., Jamieson, J. B., and Schneebeli, M.: Snow avalanche formation, Rev. Geophys., 41, https://doi.org/10.1029/2002rg000123, 2003.

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 Reg. Sci. Technol., 51, 253–272, https://doi.org/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, 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, https://doi.org/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
255   conceptual model of avalanche hazard, Nat. Hazards, 90, 663–691, https://doi.org/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 Reg. Sci. Technol., 120, 251–262, https://doi.org/10.1016/j.coldregions.2015.04.010, 2015.