AUTHOR RESPONSE

We thank Karsten Müller, Jan-Thomas Fischer, and Michael Warscher for their constructive reviews. We have revised our manuscript to address their concerns. We first respond to some of their common comments below, then address their remaining comments individually.

A common theme in the comments was requests for additional features in the visualizations. While we appreciate the suggestions, the intent of our paper was not to design the ultimate visualization tool for avalanche forecasting, but rather to argue that snowpack models can be more useful by applying avalanche forecasting principles and design principles to develop user-focused tools. To address this potential miscommunication, we tried to strengthen the main messages in our revised manuscript and emphasize the reasons for the included visualization examples more clearly.

Another common theme was concerns over our limited user survey. We have removed the survey and instead simply discuss the validity of our designs. The design principles included in our paper are grounded in a large body of literature in visualization research and do not need to be explicitly validated. However, our revised manuscript clearly emphasizes the need for extensive user testing for the successful implementation of operational visualizations.

We also agree with the reviewers that incorporating validation data would add great value to these visualizations. We added more discussion about this being another critical next step, but choose not to include this in our visualization examples, as this is would be rather complex and could dilute the main message of our paper.

Finally, there were several comments about applying the visualizations to different forecasting contexts and snowpack conditions. We've added more discussion about different forecasting contexts and added an Appendix figure with the dashboard images for different snowpack conditions over the course of a season in our study area.

RESPONSE TO RC1 (Karsten Müller)

The manuscript applies the visualization design framework proposed by Munzner (2009) to an established workflow for avalanche hazard assessment (CMAH). The goal is to enhance the interpretability and increase the relevance of numerical snowpack models for the avalanche forecaster. Snowpack models in avalanche forecasting are the equivalent to numerical weather prediction models (NWP) in weather forecasting. While weather forecasting nowadays heavily relies on NWPs, snowpack models are only sparsely used in operational avalanche forecasting. Accessibility, interpretability, relevance and integrity of snowpack models are not yet good enough for operational purposes a recent study by Morin et al claims. While issues of accessibility and integrity are not addressed in this manuscript, the main reason for poor interpretability is accorded to poor visualization of snowpack model output, which also reduces their relevance to the avalanche forecaster. Existing visualization tools are designed by the model developers with evaluation of model performance in mind, but not the operational forecaster (end-user). This manuscript presents a top-down approach for visualizing snowpack models with the forecaster/user in mind. The authors suggest that the visualization of snowpack model output should help the forecaster to answer the key questions from the conceptual model of avalanche *hazard (CMAH). They demonstrate their design by applying it to a regional avalanche forecasting scenario. The study concludes that the presented top-down design approach is superior in an operational setting based on a small user survey.*

GENERAL COMMENTS

This work is a relevant technical contribution. It describes a thorough process on how to improve snowpack visualization in operational avalanche forecasting. The avalanche community in North America, but also internationally, will benefit from these results. I found it well written and structured. Figures are of high quality. I recommend to publish the manuscript. I only have minor suggestions on how to improve the paper.

It seems like the target audience for the suggested visualizations are regional avalanche forecasters that operate with forecasting areas of several hundreds to thousands sq.km. The language used and the given example address this audience. I miss a discussion on other operational settings and more extreme cases, i.e. very large forecasting areas and high resolution model (large amount of data) and small forecasting regions and poor model resolution (too little data). Could you discuss what needs to be done to transfer the presented plots to a smaller scale? Can these visualizations be beneficial for managing avalanche safety in a ski resort? Is there a minimum number of simulated snow profiles to apply your designs? On the other hand, an avalanche forecaster can normally not wait for several minutes for a plot to be displayed. Could you discuss performance on a standard workstation/PC? How long does it take to load and update the dashboard in your example on common hardware?

Our focus on regional scale forecasting was in response to regional scale forecasters having the greatest interest in snowpack models so far. We have added more discussion about how the design principles could apply to other contexts by: at the start of Sect. 3 explaining the purpose of our examples are to show applications of design principles, acknowledging how different contexts have different spatial scales and descriptions of terrain that could benefit from maps or other geospatial views (other than elevation-aspect classes), recognizing the model needs to be configured to capture the variability of snowpack conditions for the type of forecast area (which requires model expertise). We also try to generalize more of our discussion about how the design principles could be applied to any forecasting application by emphasizing the need to focus on user needs and their specific tasks and questions. As far as technical specifications, we added a summary of the practical speed and data set size limitations of visualizations in Sect. 4.2.

In my opinion an important part of showing model output is to provide a measure/ display of uncertainty or an indication of when the model is off. In my experience, forecasters rejected model output because it is often hard and time consuming to evaluate if the model is providing reasonable results. Thus, be able to plot model output against field data or other sources will help to improve integrity. You mention this at the end of your conclusion. Could you, in addition, discuss briefly why this is not part of your study? I suggest to add a short section that summarizes the main issues with assimilating snowpack observations in snowpack models.

We agree including field data to validate the models is a critical step towards forecasters trusting models. We have expanded our discussion of this step in a new section

(4.2 Steps towards operational implementation), and although field data could be added to some of our example visualizations, this is non-trivial and we think beyond the scope of the main message of our paper. We also discuss the how visualization could play a helpful roll in assimilating field data into models by allowing forecasters and researchers to explore relationships between the two data in rich ways.

A drawback of this study is the small and rather arbitrary user survey on the effectiveness of the proposed visualization design.

We agree the user survey was small. We are currently in the early stages of the feedback process with forecasters and do not have a full user analysis to report, so we removed reference to the survey efforts and instead described the importance of an iterative design process where qualitative user feedback plays a very important role in informing and improving designs. Instead of discussing the validity of our specific designs, Sect. 4.1 now focuses on recommendations for designs in various forecasting contexts at each level of the nested model of visualization design.

SPECIFIC COMMENTS

- *p2 l8: should be "Morin et al. (in press)"as in prev sentence*
  Updated to fully published reference (Morin et al., 2020).
- *p2 l10: What is meant by "workstations"? To me this is hardware - a PC! They are normally not specifically designed for showing field data. I assume you address the lack of proper software that can make model output accessible to the forecasters. A hardware issue might be lack of CPU/GPU power to effectively handle large amounts of data/images.*
  Workstation was the term used in the Benjamin et al (2019) meteorology paper, but to be more specific we replaced that term with either "visualization tool" or "workflow" throughout the manuscript as appropriate.
- *p*2 l13: Snowpack models "relevance"comes from their ability to produce information over a large spatial scale, something field observations can not. So I think their relevance is less of a problem than their integrity, i.e. difficult to compare to field observations due to scale issues in the forcing data.
  We agree that snowpack models inherently should have 'relevant' information, but despite that, when Morin et al. (20209) discuss the information quality of snowpack models, under the 'relevance' section they suggest the added value of model information is unknown . The integrity is a serious issue, but the focus of our paper was relevance and interpretability. We have expanded our discussion on the importance of improving model integrity through visualization techniques.
- *p2 l17: Could you provide an example of a "conventional method- individual snow profiles?*
  As suggested, we added "manual snow profiles" as an example.
- *p3 l19: You could provide "assess the spatial distribution of weak layers"as an example for "major needs".*
  Done.
- *Figure 1: The figure caption could be more detailed and explain the main features and abbreviations of the shown chart. This would ease the understanding for readers outside the avalanche community.*

We have added a more detailed explanation of the information in this photo (and replaced with a high quality image).

- *p7 l18-19: The given example is typical for a regional avalanche forecast and the presented visualizations work well in this given case. Could you discuss (later in the text) the extremes, very large forecasting areas and high resolution model (large amount of data) and small forecasting regions and poor model resolution (too little data).*
  These cases are now discussed in Sect. 4.1 under design considerations at the algorithm level.
- *Table 2: Layers of large facets can be an avalanche problem, too. However, facets (FC) are treated as bulk layers here. What is the criteria (in SNOWPACK) that separates DH from FC - size only - if yes, what is the threshold?*
  This grouping is now explained and justified in more detail. In short, SNOWPACK has a size threshold where FC larger than 1.5 mm are called DH. Thus most FC layers in the model appear in bulk layers while most problematic FC layers appear as DH. We also cite Schweizer and Jamieson (2007) for common rules about identifying weak layers (i.e. the combined importance of grain type and size).
- *p9 l5: remove "and"*
  Done.
- *p9 l11: "...each day OF the season."*
  Done.
- *Figure 4: Can you explain why the percentage exceeds 100% on some days? E.g. Oct 21 or Noc 22 and 25. Does the plot only evaluate if a layer is present (Boolean - regardless of layer thickness) in a simulated profile or is the percentage each layer takes up of the total snow depth within each simulation regarded and used as a form of weight?*
  We corrected the algorithm to avoid percentages over 100%. These were caused by cases when a single profile had multiple layers on the same date with different grain types, but to simplify the data we now choose a single layer per date (with priority for weak layers over new snow over bulk layers over melt forms). And yes, the plot only counts whether a layer is present and does not account for layer thickness. We have expanded our description of this method in the text to make it clearer.
- *Figure 5: Have you considered a "polar/radar plot"for each elevation band? I can imagine that it will show the presence of layers with regard to aspect more clearly. However, the information on snow depth will be difficult to integrate.*
  Yes, we've experimented with circular plots as these are a familiar idiom for weather and avalanche forecasting. Based on visualization literature we found circular plots are effective and intuitive for overview tasks for simple types of data, but precise comparisons are difficult due to the skewed and unaligned scales. So we added a sentence explaining this rationale and suggest circular plots for simple data (category, ordinal), but argue complex data is better suited to rectilinear plots, especially when the intent of the visualization is precise comparison tasks.
- *p16 l 2: "...the principles outlineD in..."*
  Sentence removed.

- *p16 l 16: Could you provide examples for "other operational tasks"? And discuss briefly benefits and challenges (see General Comments).*
  This section has been rewritten and now focuses more broadly on how snowpack model data could be visualized for a broad selection of contexts and tasks.
- *p16 l26-27: What do you mean by "...prototypes were accessed externally from existing workstation..."? Were the plots provided by a server? Why is it critical to integrate visualizations into forecasting workstations? Do you mean integrate into software used by the forecasters?*
  We meant there was an accessibility barrier because the models were accessed on a separate webpages for the forecasters familiar workflows (software, bookmarks, etc.). However, we no longer discuss the specific user feedback.
- *p17 l3: Consider to add "...two of the four major perceived issues (BESIDES ACCESSIBILITY AND INTEGRITY) with ..."*
  Done.
- *p17 l11-12: The sentence "As highlighted by..."should be removed or rephrased since it is not clear how it fits into your conclusions. Please be more specific on your findings.*
  Sentence removed.

RESPONSE TO RC2 (Jan-Thomas Fischer)

In this manuscript the authors present a new methodological approach to enhance the operational information value of snowpack models. One of the main developments (that could get even more attention) is the (easy to use) CMAH dashboard tool, providing an online, interactive approach to reproduce the main outcomes/figures of the paper. The authors successfully demonstrate how their approach enhances the information value of snowpack modelling, with particular emphasis of layer prevalence and their spatial distribution. The paper also shows that future research and work is necessary to implement an appropriate measure and visualization for stability. The authors succeed in highlighting the spatially distributed character of the results (which could deserve an additional spatial description, see comment below). Besides this spatial character it would be worth to elaborate (or at least mention) how this method can be applied with respect to temporal variability. This could e.g. be achieved by (1) evaluating/providing an additional CMAH dashboard for a different date (in mid December?) as supplementary material (if the corresponding workload allows to?) or by (2) briefly discussing how the content and information value changes/develops throughout a season (cf. seasonal variation in Fig 2).

We agree that temporal trends are important for forecasting, and although now a specific question in the CMAH, the fact the forecasting process is iterative (and usually repeated daily) tracking temporal changes becomes important. The temporal continuity of snowpack models are well suited to understanding temporal trends, and we have added some discussion in Sect. 4.1 about how this is a specific task requires its own designs for this purpose (such as stratigraphy timelines). We added examples of the CMAH dashboard for several dates over the course of the season for Glacier National Park in the Appendix. This both shows potential to track temporal variability, and shows the effectiveness of the visualizations for different snowpack conditions.

All in all the paper is well written (although the authors could review/explain which and why so many terms are italicized). The paper has a good mixture of technical terms and corresponding descriptions. It is of high quality, enjoyable to read and fits to the scope of NHESS.

We removed most of the italics.

- *p1 l 8, ...in terrain.: Please specify on which scale(s).*
  We generalized the sentence and added "at relevant spatial scales."
- *p2 l 1, ...can range from individual slopes in backcountry guiding.: I do not see how individual slopes are relevant in the context of this work. Is it possible to distinguish?*
  This list is meant to show the diversity of avalanche forecasting contexts, we address how our visualizations and design principles transfer to other spatial scales in later sections.
- *p3 l1, ...so far prevented...: ... so far limited..*
  Done.
- *p5 l5, ...is analogous to manual snow stratigraphy...: This is very important and could be further highlighted throughout the paper by providing (a) manual profile(s) representative for the date / location(s) of the main analysis (e.g. 8 January 2018). Further the value of Figure 1 could be enhanced by connecting the snowpack information of the generalized profiles to the study area/time.*
  Yes, many of the ideas behind our designs are inspired from existing methods with manual profiles. We have added some more discussion of this, for example explaining how some of methods for grouping weak layers, characterizing sensitivity to triggers, etc. draw from established methods for interpreting manual profiles.
  We appreciate the value of having manual or whiteboard profiles for the same study area and period, however this could direct the reader towards a model validation exercise, which is not the focus of the paper. We think whiteboard profile summaries are an excellent example of how forecasters assimilate and summarize information into an abstract representation, which is more along the theme and inspiration of our study. Unfortunately, we do not have a photo of visual snowpack summary for our study area or study periods, so instead use one from somewhere else.
- *p7 l17-19: Could you comment on the specs of the profile locations and how they are chosen/defined to be representative for the study region? Is the number of study plots important? What number is expected to provide a representative analysis for the region (or spatial density)? Since spatial distribution is an important point of your analysis it could be worth to provide a corresponding map overview of the study region and profile locations (in particular for readers that are not familiar with the region).*
  We have added a map of the study area (Fig. 2) and a discussion about the importance and challenges of configuring models to obtain a representative sample. We further describe our method and acknowledge that obtaining a representative sample of profiles is a difficult question that will vary between context. We also discuss some of the visualization challenges of small or large datasets in the Discussion. In certain forecasting contexts it could be helpful and easy to implement a map in the dashboard where groups of profiles

are selected by location, but this is not necessary in the context of regional forecasting where the area is pre-defined.

- *p7 l20-25: Why is the emphasis undesired? Could you elaborate a bit with respect to what some features have high or low importance?*

  We've added some examples including the fact that thin surface hoar layers are important but often have minimal contrast with surrounding layers in the standard palette.
- *p8 l7-9, Table 2: Could you please add some references and comments on the simplified groups of grain types, with particular emphasis on why it would or why it would not be appropriate to summarize RG and FC as bulk layers (with respect to different types of metamorphism / underlying physical process). In your example (see e.g. Fig 2) it could appear also appropriate to add MF as bulk layer?*

  This grouping is now explained and justified in more detail. In short, SNOWPACK has a size threshold where FC larger than 1.5 mm are called DH. Thus most FC layers in the model appear in bulk layers while most problematic FC layers appear as DH. We also cite Schweizer and Jamieson (2007) for common rules about identifying weak layers (i.e. the combined importance of grain type and size).

  Yes our example has some thick layers of melt forms, however the colour palette places bulk layers and melt layers as equals on the visual hierarchy of importance. It could be argued to use a single hue for bulk and melt layers, however melt layers can be unique in terms of avalanche release since they are often much harder and can either bridge a weak layer or for a bed surface. Using distinct groups can help identify some of these stratigraphy patterns.
- *p9 l11-12: Please Specify (see also comment on Figure 4). Are you displaying the percentage for a specific date (8 January 2018)?*

  We corrected our aggregation algorithm to avoid percentages greater than 100% and extended our figure caption and text description to explain it is the percentage profiles in the study area with a given layer date and grain type combination.
- *p11 l14: Could you provide a (technical) reference for the "jitter" plot?*

  We now cite Ellis and Dix (2007) who discuss techniques for clutter reduction, including jitter.
- *p12 l10-14: I think it would be worth to (1) mention the availability of the median values in the interactive dashboard (which are way more instructive than the figure) and (2) to comment on the spatial variability of throughout your profiles, e.g. by mentioning the standard deviation and median values for the expected depth main avalanche problems.*

  We now mention the relevance of browsing the distribution of layer depths and the fact interactive tools can help by showing summary statistics for selected features.
- *p13 l6: Is prevalence really connected to spatial distribution (including all sample pits) or is a total measure of occurrence?*

  The CMAH defines spatial distribution of a problem as "the spatial density and distribution of an avalanche problem and the ease of finding evidence to support or refute its presence". We argue the total measure of occurrence is

related to the spatial density (assuming the data is a representative sample of snowpack variability).

- *p13 l9-10: I think it would be worth to shortly discuss why the storm slab (that previously appeared as one of the main problems) is not highlighted in the sensitivity to triggers?*
  We make it clearer that the SSI is best suited for problems related to deep weak layers (since it neglects layers within ski penetration). This also leads to our expanded discussion about using different attributes to characterize the likelihood of different avalanche problem types.
- *p14 sec. 3.5: I find this section highly instructive to understand figures and conclusions in this paper and would like to see interactive CMAH dashboard mentioned earlier in the paper, since it e.g. provides more instructive/clear visualization than the printed terrain class visualization (Fig 5).*
  We agree with this suggestion because the interactive CMAH dashboard tool is our best example of applying visualization principles, so we now highlight the tool earlier in the paper and frame each individual visualization as a component of the tool. The dashboard now receives acknowledgment throughout Sect. 3.
  The figures generally appear clear but would benefit from a (more) self-sufficient description (by e.g. referring to the interactive dashboard where applicable and indicating the specific date in all plots/captions (where applicable, e.g. of Fig 2 (right) and Fig 3,4, 5, 6, 7, 8)):
  We provide a more complete description of the data included in each figure caption.
- *Figure 1: Could you enhance this Figure / increase readability and describe which information is given in the generalized snow profile or alternatively provide a manual profile of the study region/time as reference (additionally a map view of modelled study plots could be beneficiary here, see comment below)?*
  We have replaced the image with a clearer image from a different date and provide a more detailed description of how to interpret the information. We have also added a map of the study area (Fig. 2).
- *Figure 2: Please indicate/describe somewhere (text and caption) what the difference between the left (timeline of stratigraphy) and right (hardness vs depth for a specific date (which?)) plot are?*
  The timeline and stratigraphy profiles are now explained in greater detail in the caption.
- *Figure 3: It would be helpful somehow give an overview of the study area and where the profiles are simulated.*
  The new study area figure (Fig. 2) will help readers unfamiliar with our study area. Maps can easily be added to interactive tools to select profile locations, which could be useful in some contexts, but is not critical to for this example to answer the questions of the CMAH.
- *Figure 4, Figure 7: How can 100% be exceeded (e.g. Nov 24/23 and Oct 22)? Please double check your scale or explain (see comment above concerning prevalence). In the caption - persistent grain types or grain types associated to potential/persistent weak layers (e.g. these are mentioned as weak layer in the dashboard, please double check for the sake of consistency)?*

We have corrected the algorithm and checked for consistent terminology between the text and caption.

- *Figure 7: ...shows spatial distribution... Is it really the spatial distribution in this case (like e.g. Fig 5), or rather a total measure of occurrence (see comment above). Would it be beneficial/feasible to also use the size scaling of Fig 6 allowing for a visual comparison/connection btw. the Figures?*
  We argue the total measure of occurrence is related to the spatial density, and serves as an approximate indication if a problem is widespread/specific/isolated in the terrain. We have added the size scaling to make a connection between the figures.
- *Figure 8: Why are IF not specified/displayed in the dashboard?*
  The legend hid IF because they were not present in any of the profiles, so we manually added IF to the legend for this figure.

RESPONSE TO RC3 (Michael Warscher)

In their manuscript "Enhancing the operational value of snowpack models with visualization design principles", the authors present the application of different visualization design principles in the domain of avalanche forecasting using data from the widely used model SNOWPACK.

GENERAL COMMENTS

The manuscript is technically very well written, as well as easily readable and understandable. I list some general comments and specific remarks in the following.

- *I fully understand and appreciate the usefulness of the presented visualizations and their simplification and aggregation character, however, I still would like to additionally see some conventional map plots at the top-level of the dashboard. This would be very helpful to get an overview of the domain and the spatial distribution of specific snow characteristics and avalanche problems. Examples of such visualizations are presented in e.g. Morin et al. 2020. Your aggregated plots would be a perfect summarizing and aggregating approach in a second visualization step.*
  We agree maps are a very valuable visualization tool. Following the CMAH workflow, the first step is establishing operational objectives and spatiotemporal context. Our context is a regional forecast, hence the study area is pre-defined at this stage. We added a map of the study area to help readers establish this context (Fig. 2). However, once the context is established there is no need for maps to answer the CMAH questions. We agree that integrating maps into the dashboard could be useful for other contexts, but don't think it is essential to illustrate our main message of applying visualization design principles.
- *In my opinion, the most important missing approach in the presented framework is the implementation of validation data. You state in different parts of the manuscript that practitioners lack trust in the integrity of model data. They won't gain any if they do not see the model performance at some validation points at a glance in the operational setup or at least in some hindcast simulations. I think some of the presented visualizations are perfectly suited to include observed validation*

data. You could simply include an interface to integrate measured snow profiles and plot them right into your visualizations as single highlighted data points or in the best case, somehow link them to their respective model grid point (this way, they could be included in all your visualizations, even the "sorted-by-depth" ones). I understand that it could be complicated to do this in a visually attractive way, but I think it would be well worth the effort.

We agree this would be a very valuable addition, but one that deserves special attention that goes beyond the scope of our focus on emphasizing the potential value of user-focused designs. However, we added a dedicated section called "Sect. 4.2 steps towards operational implementation" where we emphasize the importance and potential approaches for designing visualizations that support validation and understanding uncertainty.

- *While I very much like the presentation of your new color profiles, I am kind of torn as they are very much tailored to previous existing expert knowledge (potential weak layer = surface/depth hoar = highly visible) and is not very generic. Of course, this is very useful to detect the targeted wind slab avalanche problems, but what about other common avalanche problems (e.g. wet-snow avalanches). Are they also clearly visible in your visualizations? Regarding this remark, - if feasible - it would be very beneficial for the manuscript to include an additional example for a very different avalanche situation in the same domain.*

  Additional figures in the Appendix now show the visualizations for different snowpack conditions. The colour palettes are heavily tuned to identifying persistent slab problems largely due to operational feedback from forecasters who see this as the potential greatest added value from snowpack models. We've added some discussion acknowledging similar design principles could be applied to identify attributes associated with other problem types (such as visualizing weather data to identify wind slab problems or snow temperature to identify wet problems). The colour palettes we propose are simply examples of how perceptual considerations should be leveraged to direct our attention towards features of interest. We added an Appendix figure with examples of the visualizations for different snowpack conditions throughout the season.

- *I don't see the point of having so many words printed in italic letters even if they refer to specific technical terms. I think this is not necessary here and they could all just be changed to normal fonts.*

  We removed the italicized terms.

- *As the manuscript provides a technical report of the application of a visualization concept, it would be very beneficial to add information about the minimum requirements for a snowpack model in terms of resolution, simulation variables and output that is needed to feed the visualization software and dashboard. It is obvious that the software was developed for the use with SNOWPACK as a well-known and established snow (layer) model, but it would be interesting to read some more technical details about input requirements and portability.*

  We address this in two ways. One, we add more description to emphasize the importance of configuring a snowpack model to capture variability within a region, which depends on context and requires model expertise. Second, our discussion of the algorithm design level in Sect. 4.1 now includes issues that

arise with large or small datasets.

- *It would also be useful to include some more variables displayed in your visualizations, e.g. depth profiles of snow temperatures or snow density which might also be useful for avalanche practitioners and should be provided by the SNOWPACK model.*
  As explained above, we did not intend to propose an optimized forecasting tool. However, we have added more acknowledgement that additional attributes are likely needed to identify and characterize different avalanche problem types.
- *The user survey presented in section 4 is very little explained and far from being representative, so you should consider removing the section and just move the last sentence of the section to your conclusions.*
  We removed the survey and clearly emphasized the need for more user testing in Sect. 4.2.
- *I have two other comments, which might well be beyond the scope of this paper, but could be a useful addition for the future development of the presented approach: In addition to the above-mentioned validation data, it would be very useful to provide a framework for ensemble simulations including uncertainty measures. The implementation of visualizations for multi-model results and corresponding model spreads and uncertainties (ensemble model outputs from e.g. different initial conditions, different meteorological forcing data, and different snow pack models) would be a logical and highly valuable (or even necessary) next step for the application of snowpack models in real-world operational avalanche forecasting settings (similar to NWP). You should add this somewhere in your conclusions. Another helpful addition for avalanche forecasters and practitioners would be the visualization of the meteorological input in your visualization framework, e.g. wind speed and gusts, (min./max./mean.) air temperature, liquid/solid precipitation, SW/LW radiation, all separated for elevation and aspect bands and sectors (of course depending on resolution and origin of the gridded meteorological forcing, domain size, etc.).*
  Thank you for sharing these ideas. We agree there is great potential for applying visualization principles to combine various types of weather, snowpack, and avalanche data. We added a specific section (Sect. 4.2) to address some of these potential next steps and hope this paper serves as a foundation for how visualization approaches can help advance both the research and operational use of snowpack models.

SPECIFIC COMMENTS

- *P. 1, L. 8/9: Rephrase the sentence "Examples of visualizations that support these tasks are presented and follow established perceptual and cognitive principles from the field of information visualization.", to e.g. "Examples of visualizations that support these tasks and follow established perceptual and cognitive principles from the field of information visualization are presented."*
  Done.
- *P.1, L. 18 and others: Regarding the term "workstations". Maybe Benjamin et al. 2019 labelled the development of software, more powerful computers and more available model and observation data as kind of mythical "workstations", I*

would prefer just to call it what it is, namely more powerful computers, more data, and better visualization tools that gradually developed in NWP and of course in all other fields.
Workstation was the term used in Benjamin et al (2019) meteorology paper, but to be more specific we replaced that term with either "visualization tool" or "workflow" throughout the manuscript as appropriate.

- *P. 2, L. 7 and others: update citation Morin et al., is published now.*
  Done.
- *P.5, L. 10: "as hardness profiles" instead of "as a hardness profiles"*
  Done.
- *P. 6. Fig. 1: Do you have a version with better image quality available? The figure is very hard to read. However, I would suggest to remove Fig. 1 anyways as it does not contain important information in the context of the manuscript. If you decide to keep it, you should add some more information to the manuscript explaining what the reader is supposed to see in the figure.*
  We replaced with a higher quality figure. We think the whiteboard profile summary is an excellent example of how forecasters assimilate and summarize information into an abstract representation, and was a major inspiration for some of our designs. We add more description of these profiles and explain more in the text why it was included in the paper.
- *P. 9, Fig. 2, x-axes right panel-plots: Please add explanation for the hardness abbreviations and a "hardness" x-axis label. It becomes clear from the text, but should be included in the figure or at least in the figure caption. That also holds for the hardness test abbreviations (F, 4F, 1F, P, K) which are clear for an avalanche practitioner (fist, 4 fingers, 1 finger, pencil, knife), but the article might be interesting for a broader (snow) scientific audience. Please add explanations.*
  Done.
- *P. 9, L. 15: "Herla et al., in preparation" should be removed if not already published by now.*
  Removed citation as this work is not published yet.
- *P. 10, Fig. 3: Even if it is clear when reading the manuscript and figure caption, I would prefer to have an arrow-type label on the x-axis (e.g. "Thinnest snowpack ¡-¿ Thickest snowpack")*
  Done.
- *P. 11, L. 7: "way to visualize" instead of "way visualize"*
  Done.
- *P. 11, L. 7: "way to visualize" instead of "way visualize"*
  Done.
- *P. 11, L. 10: I suggest to rephrase the sentence: "Instead, using eyes to: : :", e.g. "Instead, simultaneously comparing 1D/2D visualizations: : :"*
  Done.
- *P. 13, Fig. 6, caption: "slab" instead of "slabs"*
  Done.
- *P. 14, Fig. 7: Labels "Sep 30" and "Sep 23" overlap, please solve this issue.*
  Removed overlapping labels.
- *P.14, L. 2: Please use italic here ("Tableau") as this seems to be the name of a commercial software developing company. Just a comment: it would be very*

beneficial if you would develop the dashboard in R or a similar open source programming language, as you have already done with the visualizations. This would foster the use of your very useful software by different target groups.
Tableau is a fast and easy visualization prototyping software, while R is very limited in terms of interactive visualization. While similar designs can be done with open source visualization libraries (e.g. D3), they tend to be more rigid and are better suited for later stages in the design process. We intend to make more of our tools openly available when they get to that stage.

- *P. 15, Fig. 8: Could you provide a screenshot with better quality? The very useful dashboard is kind of hard to acknowledge here.*
  We increased the resolution of the figure.

Enhancing the operational value of snowpack models with visualization design principles

Simon Horton[1], Stan Nowak[1], and Pascal Haegeli[1]

[1]Simon Fraser University, Burnaby, BC, Canada

Correspondence: Simon Horton (shorton@avalanche.ca)

Abstract. Forecasting snow avalanches requires a reliable stream of field observations, which are often difficult and expensive to collect. Despite the increasing capability of simulating snowpack conditions with physical models, models have seen limited adoption by avalanche forecasters. Feedback from forecasters suggest model data is presented in ways that are difficult to interpret and irrelevant to operational needs. We apply a visualization design framework to enhance the value of snowpack models to avalanche forecasters. An established risk-based ~~workflow for avalanche forecasting~~ avalanche forecasting workflow is used to define the ways forecasters solve problems with snowpack data. We ~~address~~ suggest model data should be visualized in ways that directly support common forecasting tasks such as identifying snowpack features related to avalanche problems ~~, summarizing snowpack features within a forecast area, and locating~~ and locating avalanche problems in terrain at relevant spatial scales. Examples of visualizations that support these tasks ~~are presented~~ and follow established perceptual and cognitive principles from the field of information visualization are presented. Interactive designs play a critical role in understanding these complex datasets and are well suited for forecasting workflows. ~~Preliminary feedback suggests these design principles produce visualizations that are~~ Although extensive user testing is still needed to evaluate the effectiveness of these designs, visualization design principles open the door to more relevant and interpretable applications of snowpack model for avalanche forecasters~~, but additional operational testing is needed to evaluate their effectiveness. By addressing issues with interpretability and relevance, this~~. This work sets the stage for implementing snowpack models into ~~workstations~~ visualization tools where forecasters can test their operational value and learn their capabilities and deficiencies.

1 Introduction

Numerical environmental and weather prediction models have dramatically transformed the accuracy of weather forecasts and the role of weather forecasters since the 1980s (Benjamin et al., 2019). As model performance improved, forecasting tasks shifted from predicting weather conditions to interpreting and communicating model guidance. A centerpiece in the adoption of prediction models by weather forecasters was the development of ~~workstations~~ visualization tools that allowed them to work directly with gridded modelled data in combination with in-situ weather observations and remote sensing data (Benjamin et al., 2019). This setup allowed forecasters to visualize model output along with observations and gradually learn the operational value of the models.

The work of avalanche forecasters is similar in nature and complexity to the work of weather forecasters. The objective of avalanche forecasting is to develop an accurate mental model of the current and future nature of avalanche hazard by integrating avalanche, snowpack, and weather information from a variety of sources (Canadian Avalanche Association, 2016b). This assessment is then combined with terrain information to make risk management decisions regarding specific elements at risk. The spatial scale of avalanche forecasting can range from individual slopes in backcountry guiding, to groups of avalanche paths when protecting infrastructure, ~~and~~ to entire mountain ranges in public avalanche warnings. To assist avalanche forecasters at the higher end of the spatial scale spectrum, physical snowpack models such as Crocus (Brun et al., 1992) and SNOWPACK (Lehning et al., 1999) were developed in the 1990s to provide supplementary data about snowpack conditions. Despite the fact ~~snowpack model developers~~ model developers have created numerous operational tools to visualize model output~~for avalanche forecasters~~, snowpack models have so far only seen limited adoption into operational workflows when compared to weather prediction models ~~(?)~~(Morin et al., 2020).

~~Morin et al. (accepted)~~ Morin et al. (2020) employed the information quality framework of Bovee et al. (2003) to describe issues with operational snowpack model tools in terms of ~~*accessibility*, *interpretability*, *relevance*, and *integrity*~~ the accessibility, interpretability, relevance, and integrity of the information. ~~*Accessibility*~~ Accessibility to snowpack models is limited by the time constraints in forecasting environments and ~~workstations that are optimally~~ workflows that are designed for field data rather than model data. Existing tools are also difficult to ~~*interpret*~~ interpret as model output is complex and in their current form require expertise or substantial training to comprehend and ~~utilize. The *relevance*~~ apply. The relevance of the information they provide is also questioned, as similar information may be available from other sources. The ~~*integrity*~~ integrity of model output is also difficult to evaluate in an operational setting where there is limited validation data. ~~For example, snowpack~~ Snowpack models can produce snow stratigraphy profiles for multiple parameters (e.g. grain size, hardness, temperature) at different time intervals at potentially hundreds or thousands of locations. ~~Furthermore, the output from snowpack models are~~ This data can be so complex and voluminous that it becomes ~~extremely~~ challenging for operational forecasters to make sense of in its raw form using conventional methods such as viewing manual snow stratigraphy profiles. This has been described as “information overload” and characterizes “big data” environments (De Mauro et al., 2016). As avalanche forecasting requires substantial cognitive effort to continuously maintain a mental model of conditions (Maguire and Percival, 2018), introducing additional complex data can disrupt this process and have adverse effects on performance. Based on their analysis, ~~?~~ Morin et al. (2020) aptly conclude that while it was important for researchers to focus on improving the accuracy of snowpack models, we are now at a point where addressing issues with the design of operational tools is critical for making snowpack models truly valuable for avalanche forecasting.

To address the challenges of big data and make it tractable for human analysis, the field of visual analytics blends automatic analysis with human analysis via visual interfaces (Keim et al., 2008). Specifically, visualization in combination with interaction techniques support a process of iterative inquiry into data to support sense-making. This reduces the cognitive work needed to perform analytic tasks by leveraging the pattern detecting abilities of the human visual system for processing complex information that would normally exceed cognitive limits (Ware, 2012). Visual analytics has made complex problems and model output tractable for non-scientists and non-model experts in a variety of domains including physics, business, intelligence analysis, and disaster management (Keim et al., 2008). Effective visualization techniques are particularly valuable for environmental data, which is often complex due its spatiotemporal dimensions and uncertainties (Grainger et al., 2016). For example, studying visualization design principles has improved the interpretability of complex data sets in the fields of meteorology (Rautenhaus et al., 2018; Stauffer et al., 2015) and oceanography (Thyng et al., 2016).

Judging from the success of visual analytics applications in other disciplines of environmental science, we believe ~~that~~ applying a visualization design perspective to snowpack models has the potential to substantially address some of the shortcomings that have so far ~~prevented~~ limited their operational use. In this paper, we present design principles ~~and visualization examples that aim to~~ for visualization tools that increase the interpretability and relevance of snowpack models for operational avalanche forecasters. These design principles are informed by information visualization, ~~operational~~ avalanche forecasting practices, and the unique features of snowpack model data. First, we apply a visualization design framework to the domain of avalanche forecasting to outline principles of how data should be visualized to solve operational problems (Sect. 2). Then we provide examples of visualizations where these principles are applied with snowpack model data (Sect. 3)~~. The validity of the design principles is discussed~~, followed by suggestions for next steps towards operational applications (Sect. 4) ~~followed by conclusions and recommendations for adoption by avalanche forecasters~~ and conclusions (Sect. 5).

2 Visualization design principles for avalanche forecasting

2.1 Nested levels of visualization design

The nested model for visualization design described by Munzner (2009) has established itself as a valuable framework for ~~evaluating and designing~~ designing and evaluating visualization tools. This framework considers four nested levels where distinct design issues arise, and where issues at one level can cascade to other levels. The issues with operational snowpack model tools identified by ~~?~~ Morin et al. (2020) relate to design issues at each level of the nested model. ~~The considerations included in these~~ These four levels provide designers with a tangible framework for understanding the users' problems, showing the appropriate information, and presenting it both effectively and efficiently:

1. *Domain situation level.* The domain situation describes the target users, their field of interest, their questions, and their data. A domain has unique vocabulary for describing its data and problems, and usually has an existing workflow for how data is used to solve problems. Issues arise when designers misunderstand the users' needs. For example, existing tools that present snowpack model data may not address the major needs and questions of avalanche forecasters~~(*relevance*~~, such as assessing the spatial distribution of an avalanche problem (relevance).

2. *Task and data abstraction level.* Task and data abstraction maps domain-specific problems into generic vocabulary that clearly describe what type of data is being visualized and why. Tasks are described with generic verbs (e.g. locate, compare) and data is described with generic nouns and adjectives (e.g. table, network, ordered, categorical). Issues arise when the functions and data types in a design do not solve the intended problem. For example, detailed snow

stratigraphy profiles provided by snowpack models may not be the type of information needed for specific forecasting tasks (~~*relevance*~~relevance).

3. *Visual encoding and interaction idiom level.* This level creates visual representations of the data. A distinct visual representation is called an idiom. Data is encoded by arranging it along spatial dimensions and mapping attributes to non-spatial visual features such as colour, size, and shape, while interaction idioms allow the user to change the view. Issues arise when idioms are ineffective at visualizing information. Existing idioms for visualizing snowpack model data are often complex, busy, and difficult for non-model experts to understand (~~*interpretability*~~interpretability).

4. *Algorithm level.* This is the level where idioms are produced from raw data with a computer. Issues arise when algorithms are too slow. At the algorithm level, most snowpack model visualizations are too time consuming for forecasters because they are poorly integrated into their ~~workstations (*accessibility*~~workflows (accessibility).

Munzner (2014) also describes that visualization problems can be attacked from two possible directions within the nested model: top-down approaches that first understand the domain and tasks and then design visual idioms accordingly, and bottom-up approaches that start with developing new algorithms and idioms. Most existing snowpack model visualizations were developed with bottom-up approaches that began with model development followed by the creation of visualizations of the model output. Bottom-up approaches allow novel visualizations that reveal nuances and anomalies in new types of data, but also have the potential to not solve the intended problem (Munzner, 2014). While it is worth considering bottom-up designs that take advantage of the unique capabilities of snowpack models, it is also important to carefully examine the domain and tasks of avalanche forecasting to establish top-down design principles that support forecasting needs.

2.2 Domain of avalanche forecasting

Avalanche forecasting is a common task for all operations that manage short-term avalanche risk (e.g. ski areas, transportation corridors, backcountry warnings, resource extraction). The forecasting process consists of iterative data analysis and is dominated by human judgement and inductive logic (LaChapelle, 1980; McClung, 2002). Statham et al. (2018) surveyed existing operational practices within North American avalanche forecasting operations to develop a standard framework for this process. The resulting ~~*conceptual model of avalanche hazard*~~ conceptual model of avalanche hazard (CMAH) identifies the key components of avalanche hazard and provides standard workflow and terminology to guide the forecasting process. The CMAH is a risk-based framework that is consistent with other natural hazard disciplines and can be applied to any scale in space or time. A central part of the CMAH is the concept of avalanche problems that represent individual, identifiable operational concerns that can be described in terms of their potential avalanche type, location, likelihood, and size (Statham et al., 2018). Under the CMAH, avalanche forecasting is viewed as sequentially answering four questions:

1. What type of avalanche problems exist?

2. Where are these problems located in the terrain?

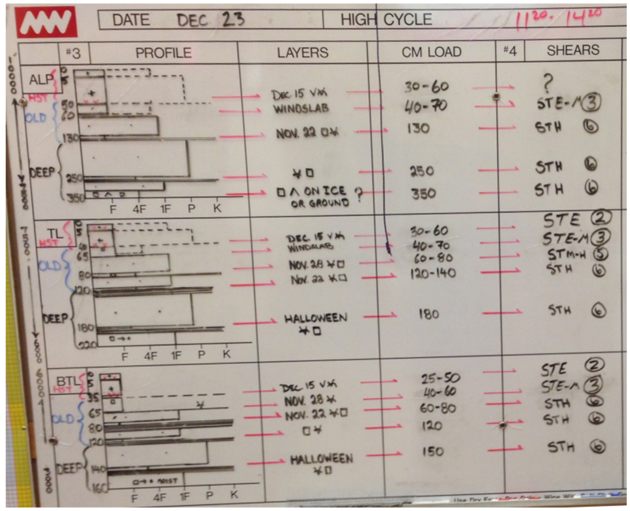

Figure 1. Example of a snowpack summary for a large forecast area drawn on a whiteboard to summarize relevant snowpack conditions. This example was drawn on 23 December 2017 by an avalanche forecaster at Mike Wiegele Helicopter Skiing in Canada. Each row shows conditions for a different elevation band (alpine, treeline, and below treeline). The profile column has stratigraphy profiles showing typical layer depths and hardness (widespread layers are shown with solid lines and localized layers with dashed lines). The remaining columns identify important layers and provide details on their burial date, grain type, depth (in cm), and representative shovel shear test results (Photo: Mike Wiegele Helicopter Skiing).

3. How likely is it that an avalanche will occur?

4. How destructive will the avalanche be?

Over the past decade, the CMAH has been widely adopted by all industry sectors in North America (Statham et al., 2018), which clearly indicates that it is a useful model to describe the domain situation of avalanche forecasting.

2.3 Task and data abstractions for snowpack analysis

Given the importance of avalanche problems in avalanche forecasting~~practices, any operational visualization of data should consider this abstraction~~, operational visualization tools should be designed to help forecasters identify and characterize avalanche problems. Assessing avalanche problems consists of integrating a complex array of data that includes observations of avalanches, snowpack, weather, and terrain (Statham et al., 2018). There is no structured or standardized way this data is used to answer the CMAH questions, as the analysis relies on subjective judgement and heuristics (LaChapelle, 1980), however there are common practices for interpreting field observations.

Snowpack models produce data that is analogous to manual snow stratigraphy profiles, which is a key type of field data used by forecasters to understand snowpack conditions. Building off familiar visual representations is an effective way for people to understand new types of information (Blackwell, 2001), and thus examining existing practices for visualizing and analyzing manual snow profiles provides insight into ways snowpack models could be visualized to support forecasting tasks.

Forecasters perform several analysis tasks with ~~snow stratigraphy profiles to help them assess avalanche problems and~~ manual snow profiles to develop a comprehensive mental model of hazard conditions. ~~Individual~~ Manual snow profiles are either recorded in tables of unstructured text or illustrated as ~~a hardness~~ hardness stratigraphy profiles (Canadian Avalanche Association, 2016a). Forecasters learn to ~~*identify* relevant snowpack *features*~~ identify relevant snowpack features in these profiles, and then ~~*compare*~~ compare multiple snow profiles along with other observations to ~~*summarize*~~ summarize the snowpack conditions within a forecast area. Forecasters summarize snowpack data ~~by writing a~~ in a written snowpack summary that gives a concise overview of ~~snowpack~~ conditions in their forecast area. The goal of a written snowpack summary is to organize and reduce data, focusing on average conditions along with potential anomalies and outliers (Canadian Avalanche Association, 2016a). Some ~~operations also illustrate~~ forecasters visualize their snowpack summary with ~~generalized stratigraphy profiles~~ a representative profile for their forecast area, which helps them organize and communicate relevant information (Fig. 1). These visual snowpack summaries are an example of where forecasters already use visualization to help summarize and understand complex information.

Tracking ~~*trends*~~ trends in snowpack conditions over time is another common forecasting task, which is most often done with tables of text. Temporal trends in the likelihood and size of avalanches are particularly relevant. For example, the InfoEx forecasting workflow allows forecasters to track weak layers in their forecast area with qualitative summaries of their status and depth each day of the season (Haegeli et al., 2014). ~~Simple observed snowpack data is~~ Basic snowpack observations are plotted as time series (e.g. daily snowfall at fixed observation sites), but complex data like snowpack structure is rarely visualized temporally.

To help forecasters ~~answering~~ answer the four key questions about avalanche problems posed by the CMAH, visualizations of snowpack ~~model~~ data should help forecasters ~~*identify*, *compare*, and *summarize* snowpack *features* and highlight *trends*~~ identify, compare, and summarize snowpack features in their forecast area and highlight trends over time. These specific tasks should be considered when designing tools to visualize either field data or snowpack model data.

~~Snowpack conditions within a large forecast area are summarized with generalized snow profiles highlighting important snowpack features at different elevation bands (Photo: Mike Wiegele Helicopter Skiing).~~

2.4 Information visualization principles

The field of information visualization studies how to leverage the human visual system to off-load cognitive work and visualize information effectively. Information visualization principles should be considered when designing the visual appearance and interactive components of tools for snowpack model data (i.e. the visual encoding and interaction idiom level of the nested model). These principles consider effective ways of representing data visually and are explained in greater detail in textbooks by Ware (2012) and Munzner (2014). The following list summarizes information visualization principles that are relevant when visualizing snowpack model data:

- When representing information visually, designers encode data to visual features such as: spatial position, size, color, or shape among others. Color can be further divided into hue (the actual color), luminance (the brightness or darkness of a color), and saturation (the intensity of the color). Through years of perception studies, standard guidelines for mapping these visual features to data types have been established (Cleveland and McGill, 1984).

- Visual encodings should present data in ways that match the capabilities of our visual system. Hence, categorical and ordered data should be encoded with visual features that match human visual aptitudes. For example, when using colours, hues should be used for categorical attributes such as avalanche problem types and luminance (lightness or brightness) should be used for ordered attributes such as avalanche likelihood.

- Designs should prioritize the importance of information and encode data to visual features that are perceived more quickly, accurately, and draw our attention to make this information more salient (i.e. noticeable) and discriminable (Cleveland and McGill, 1984). Spatial position is perceived the fastest and most accurately, and thus the most important attributes should be encoded by their position in a visualization. After spatial position, designs should consider the hierarchy of salience for non-spatial visual features. For example, size features such as length and area are more salient than colour features such as hue, luminance, and saturation. For a comprehensive breakdown of this hierarchy see Munzner (2014).

- Choose designs that are accessible and effective for common types of colour blindness. For example, red-green colour blindness (deuteranopes) affects roughly 8 % of males of European descent (Birch, 2012).

- Interaction reduces cognitive load and helps users understand data by asking questions and performing queries. A practical guideline for designing interaction idioms is the visual information seeking mantra of Shneiderman (1996): “overview first, zoom and filter, then details on demand”. The initial visualization should provide an overview of the entire dataset, and then the interactions should allow the users to change the view to see subsets of the data, and then visualize details about features of interest. This design approach offers users a flexible way to explore data, while being able to maintain a sense of context and orientation.

- Comparison tasks are most effective on aligned scales. Furthermore, comparisons of large amounts of data are often more effective when seeing multiple frames in a single side-by-side view rather than changing views over time. The human perceptual system is effective at reading spatial information in parallel, whereas changing views with animations or multiple tabs relies on human memory and results in substantial cognitive load (Ware, 2012).

- Visualization idioms should present data with the smallest number of spatial dimensions, avoiding three-dimensional visualizations and using one-dimensional lists where possible. Displaying three-dimensional data on planar surfaces has numerous issues with depth-perception and over plotting (Ware, 2012).

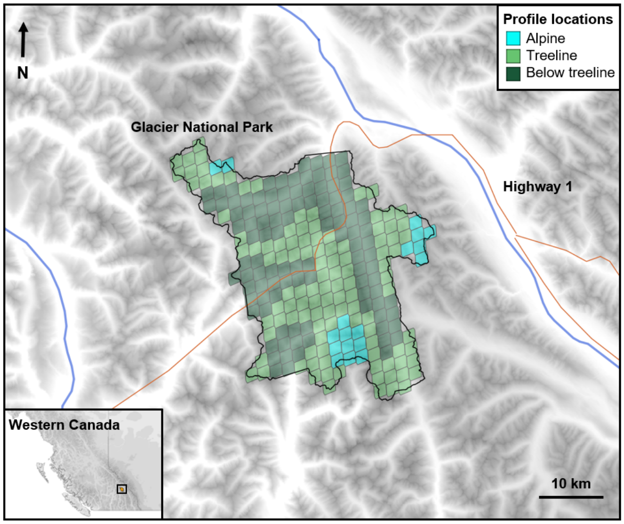

Figure 2. Location of Glacier National Park in western Canada, and locations of grid points where meteorological data was extracted to produce simulated snow profiles (coloured according to elevation band).

3 Applications of visualization design principles

This section presents ~~applications of the~~ examples of how visualization design principles ~~using simulated snowpack data for 8 January 2018 in~~ can be applied to enhance the relevance and interpretability of snowpack model data. Rather than presenting the optimal avalanche forecasting tool, these examples show how to apply a top-down approach to design. Each visualization addresses a specific question posed by the CMAH, which can be combined into a single interactive tool that allows sequential question asking. This section starts by introducing each individual visualization then finishes with an example of an interactive forecasting tool that combines them.

The following examples are for Glacier National Park, Canada~~. On this day the avalanche danger rating was considerable~~, a forecast region covering $1354\,\mathrm{km}^2$ of mountainous terrain (Fig. 2). The examples focus on the needs of regional scale avalanche forecasters (considerations for other forecasting contexts are discussed in Sect. 4). The examples use simulated snowpack data for 8 January 2018, as this day had interesting snowpack conditions with considerable avalanche danger at all elevation bands ~~with two~~ and two common avalanche problems (Parks Canada, 2018): a storm slab problem at all elevations (size 1 to 2 avalanches were possible to likely) and a persistent slab problem at treeline and below treeline elevations (size 1 to 3 avalanche were possible to likely). ~~Simulated~~ The appendix provides additional examples of the visualizations for several other days throughout the 2017-18 season.

Simulated snow profiles were produced by forcing the physical snowpack model SNOWPACK (Lehning et al., 1999) with gridded meteorological data from the Canadian HRDPS numerical weather prediction model (Milbrandt et al., 2016). ~~Meteorological data was extracted at~~ Numerous configurations of weather inputs and geometries are possible with snowpack models (Morin et al., 2020). The ideal configuration for avalanche forecasting should produce a representative sample of snow profiles that capture the spatial variability across the forecast region. Choosing an optimal configuration remains an open research question that requires model expertise and field validation. To produce a sample of profiles that cover the type of locations considered by regional forecasters, a gridded approach was used to extract meteorological data from all 236 grid points in the ~~park and at each grid point a~~ forecast region. A single flat field profile and four virtual slope profiles were simulated at each grid point (38 ° slopes in four cardinal directions) ~~. A~~ resulting in a total of 1180 profiles covering ~~an area of provide a sample data set to present visualizations of regional snowpack conditions~~a range of aspect and elevation bands.

3.1 Identify snowpack structure patterns with colour

Snowpack features related to avalanche problems should be easy to identify in visualizations of snowpack structure. ~~The~~ For example, thin weak layers are important for slab avalanche problems, and so these layers should have high contrast from surrounding layers. From a perceptual perspective, the standard colour palette for snow grains (i.e. Fierz et al., 2009) ~~creates~~ may cause undesired emphasis on certain types of snow ~~. Important features such as thin weak layers have relatively low perceptual salience while less important features such as~~ due to the relative contrast between colours. For example, the fuchsia colour used for surface hoar has little contrast with surrounding layers while melt forms and ice formations ~~have relatively high salience~~are highly emphasized (despite being less important for identifying most avalanche problems). The colours also make it difficult for individuals with colour blindness to distinguish important features ~~.~~ (e.g. precipitation particles and melt forms are difficult to discern for individuals with red-green colour blindness).

Table 1. A perception-informed colour palette for snow grain types that emphasizes features related to avalanche problems and is effective in grayscale and for common types of colour blindness.

Grain type	Standard	Perception -informed	Grayscale	Deutera-nomaly	Deutera-nopia	Hex code
Surface hoar (SH)						#ff0000
Depth hoar (DH)						#0078ff
Precipitation particles (PP)						#ffde00
Decomposing and fragmented particles (DF)						#f1f501
Rounded grains (RG)						#ffccd9
Rounding faceted particles (FCxr)						#dacef4
Faceted crystals (FC)						#b2edff
Melt forms (MF)						#d5ebb5
Melt-freeze crust (MFcr)						#addd8e
Ice formations (IF)						#a3ddbb

Table 2. Simplified colour palette for groups of grain types related to avalanche problems.

Group	Grain types	Colour	Hex code
Persistent weak layers	SH, DH		#95258f
New snow layers	PP, DF		#ffde00
Bulk layers	RG, FCxr, FC		#dacef4
Melt and ice layers	MF, MFcr, IF		#d5ebb5

We propose a perception-informed colour palette for snow grain types that ~~emphasize~~ emphasizes features related to avalanche problems (Table 1). Similar perception-informed colour palettes have been proposed to improve the interpretation of visualizations in meteorology and oceanography (Stauffer et al., 2015; Thyng et al., 2016). The proposed colour palette

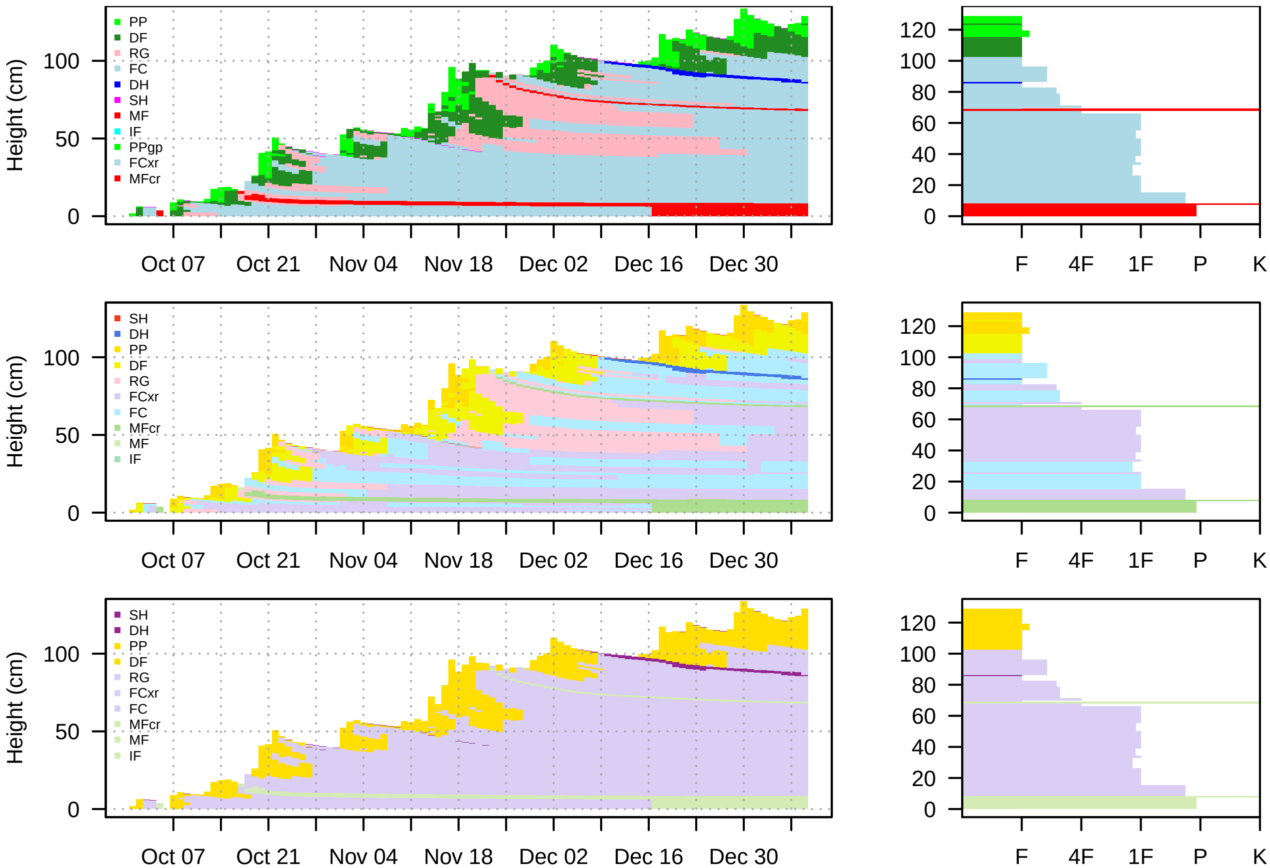

Figure 3. Applying different colour palettes to common snowpack model visualizations. The same timeline and stratigraphy profiles are shown with the standard colours for grain types (top row), perception-informed colours for grain types from Table 1 (middle row), and perception-informed colours for grain type groups from Table 2 (bottom row). Timeline profiles (left) show the evolution of layer heights and grain type from 1 October 2017 to 8 January 2018. Stratigraphy profiles (right) show layer height, grain type, and hand hardness (F = fist, 4F = four finger, 1F = one finger, P = pencil, K = knife) on 8 January 2018.

groups grain types into four categories based on their role in avalanche problems: persistent weak layers (surface hoar and depth hoar), new snow layers (precipitation particles and decomposing and fragmented particles), bulk layers (rounded grains and faceted crystals), and melt and ice form layers. While faceted crystals are typically considered persistent weak layers, the SNOWPACK model classifies any faceted crystal with a grain sizes greater than 1.5 mm as depth hoar. This rule causes most modelled layers composed of large faceted crystals (i.e. those associated with persistent weak layers) to be classified as depth hoar, while layers with smaller faceted crystals tend to be thicker and associated with slabs. While these grain type groups are defined by model behaviour, they are consistent with common snow profile analysis techniques that consider a combination of grain type and grain size (amongst other properties) to identify weak layers (Schweizer and Jamieson, 2007).

These groups were visually related using analogous color schemes (e.g. the hues are perceptually close to each other) that remained visually discriminable. The visual salience of these groups was adjusted using properties of color such as how dark they appear (i.e. luminance) and how vivid the colors are (i.e. saturation). In this way a visual hierarchy of importance was created. Weak layers that tend to take up the smallest area were made the most salient by using strong contrast against other grain types, next new snow was made salient. Finally, the other layers formed the lowest level of perceptual salience and serve as a neutral background. All colors were made to be perceptually distinct ~~. Accessibility~~ and accessible for common types of colour blindness ~~was also considered~~ (see Table 1). Unique colours were also assigned to melt-freeze crust and rounding faceted particles, as distinguishing these sub-classes was deemed meaningful for avalanche forecasters. A simplified colour palette was also designed using only the four main categories of grain types for non-model experts (Table 2). The simplified palette uses ~~analogous colours~~ colors that are analogous to the full palette and ~~maintains a similar~~ maintain the established visual hierarchy.

~~**Table 1.** A perception-informed colour palette for snow grain types that emphasizes features related to avalanche problems and is effective in grayscale and for common types of colour blindness.~~

~~**Table 2.** Simplified colour palette for groups of grain types related to avalanche problems.~~

~~Comparison of timeline and stratigraphy profiles with standard colours for grain types, perception-informed colours for grain types (Table 1), and perception-informed colours for grain type groups (Table 2).~~

The colour palettes were tested with common visualization idioms such as hardness and timeseries profiles (Fig. ~~2~~3). Comparing the standard and redesigned colour palettes at a single treeline location in Glacier National Park shows how the new palettes simplify the interpretation of the profiles by drawing attention to the most important snowpack features~~on 8 January 2018.~~ . The increased salience of the thin depth hoar ~~layers~~ layer highlights a potential persistent slab avalanche problem and the new snow highlights a potential storm slab avalanche problem.

3.2 Identify avalanche problem types from multiple profiles

Visualizing information from an ensemble of snow profiles is an effective way to identify snowpack patterns in a forecast area. Identification and summarization tasks can be done fast and effectively by deriving visual summary statistics from distributed visual information. For example, humans can visually calculate correlation coefficients, clusters, and averages with their visual perception systems (Szafir et al., 2016). The volume and continuity of data produced by snowpack models offers

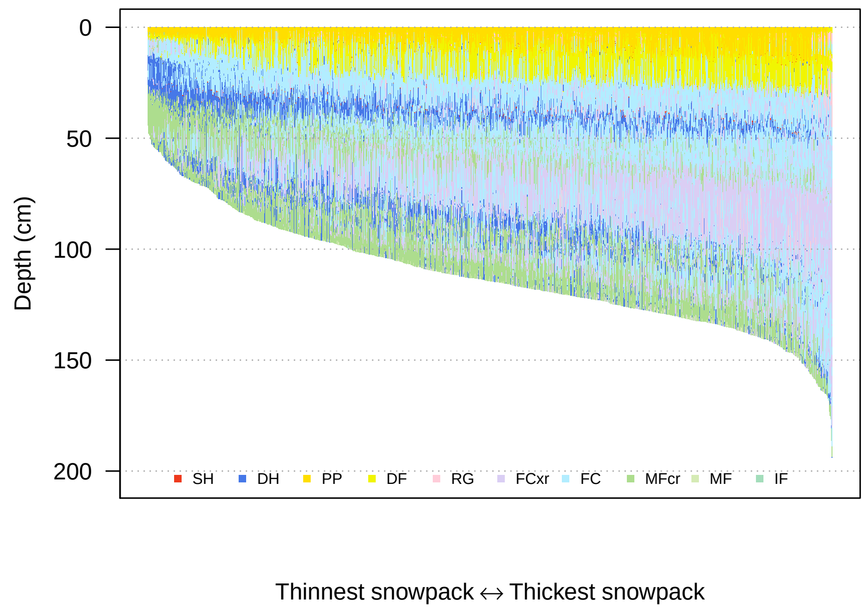

Figure 4. A visualization to identify avalanche problem types from 1180 simulated profiles on 8 January 2018. The profiles are summarized by plotting grain type stratigraphies side-by-side and sorting the profiles from thinnest to thickest. Grain types are coloured using the perception-informed palette from Table 1. The storm slab avalanche problem is emphasized with yellow surface layers and the persistent slab avalanche problem is emphasized by the band of blue depth hoar layers 30 to 50 cm below the surface.

new opportunities for summarizing snowpack structure that are not possible with ~~human observed~~ manual snow profiles. When used in combination with a colour palette that emphasizes snowpack features related to avalanche problems, profile ensemble visualizations can help forecasters identify prominent avalanche problem types.

A simple and powerful summary is obtained by plotting multiple grain type profiles side-by-side (Fig. ~~3~~4). In this example, 1180 profiles are sorted from thinnest to thickest and over 46,000 individual snow layers are shown in a single view. Despite the large volume of data, ~~and~~ a few prominent snowpack features pop-out and attention is drawn to the main snowpack patterns in the forecast area. Since this visualization is specifically designed for the task of identifying potential avalanche problem types, other idioms are required for visualizing geospatial patterns in a meaningful way (see Sect. 3.3).

~~Snowpack layers from 1180 simulated profiles are summarized by plotting grain type stratigraphies side-by-side and sorting the profiles from thinnest to thickest. Grain types are coloured using the perception-informed palette from Table 1.~~

~~Prevalence of snowpack layers from 1180 simulated profiles by aggregating layers by their age. A diverging scale distinguishes the percentage of profiles with layers containing persistent grain types (i.e. surface hoar and depth hoar) on the right from the percentage of profiles with layers containing other grain types on the left. Grain types are coloured using the perception-informed palette from Table 1.~~

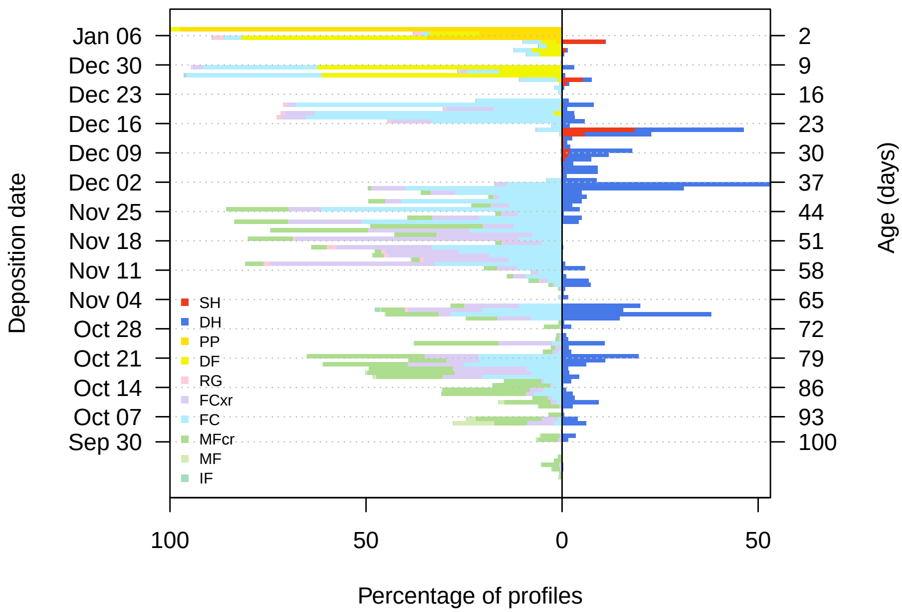

Figure 5. A visualization to identify avalanche problem types from 1180 simulated profiles on 8 January 2018. Snowpack layers are aggregated by their age to show their prevalence throughout the region (with widespread layers appearing in a greater percentage of profiles). A diverging scale distinguishes the layers with grain types associated with persistent weak layers (i.e. surface hoar and depth hoar) on the right from the layers containing other grain types on the left. Grain types are coloured using the perception-informed palette from Table 1.

Another summary visualization that draws attention to potential avalanche problem types is produced by aggregating layers by their age ~~or deposition date~~ (Fig. ~~4). Simulated profiles can be aligned and aggregated by the deposition date of each layer to summarize the main features amongst a~~ 5). The simulated profiles for 8 January 2018 have layers dating back to the start of the winter. Common features amongst the set of profiles ~~. The prevalence of different grain types is determined by counting the percentage of profiles containing grain types for each day the season. Grain~~ such as new snow near the surface and widespread weak layers share similar deposition dates, thus counting the number of profiles with different age and grain type combinations results in a summary of the snowpack structure. In Fig. 5, grain types associated with persistent weak layers are emphasized with a diverging horizontal scale to distinguish them from other grain types. The persistent weak layers are ~~also~~ easier to notice in this visualization than in Fig. 3 and 4, because they occupy a greater ~~area than in Fig. 2 and 3 where their size~~ spatial area in the visualization than visualizations where the spatial area occupied by a layer is proportional to ~~layer thickness. While it~~ its thickness. It is also possible to produce an aggregated stratigraphy profile ~~from aligned layers (e.g. Hagenmuller and Pilloix, 2016; ?),~~ by aligning layers based on other properties such as hardness (e.g. Hagenmuller and Pilloix, 2016), however this requires complex data transformations and assumptions about averaging layer properties. The layer prevalence visualization in Fig. ~~4 is simple to implement and~~ 5 supports the task of identifying potential avalanche problem types in a way that is is fast and simple to implement.

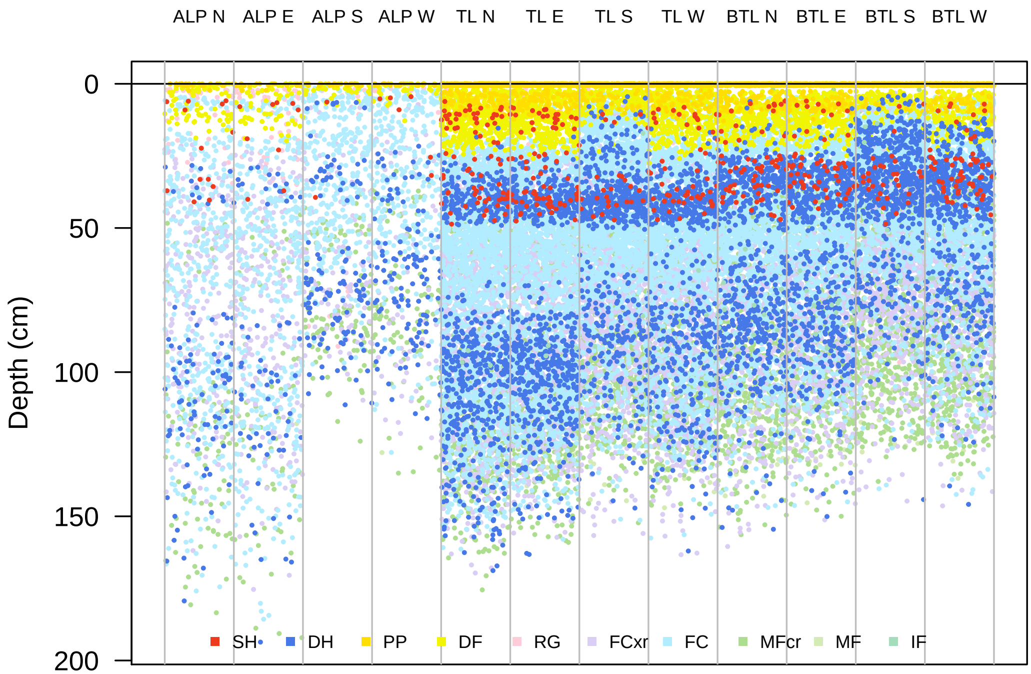

Figure 6. A visualization to locate avalanche problems in terrain. Snowpack layers from 1180 simulated profiles on 8 January 2018 are partitioned into terrain class bins for elevation band and aspect. Elevation bins include alpine (ALP), treeline (TL), and below treeline (BTL) and aspect bins include four cardinal directions (north, east, south, west). Each layer is given a random horizontal position within the bin to allow visual summary statistics. Grain types are coloured using the perception-informed palette from Table 1.

The visualizations in Fig. ~~3 and~~ 4 and 5 use colour and position to draw attention to snowpack features that relate to the storm slab and persistent slab avalanche problems on 8 January 2018. The storm slab problem is apparent from the yellow new snow grains on the surface and a potential persistent slab avalanche problem is apparent from the salient surface hoar and depth hoar layers that are buried 30 to 50 cm below the surface (Fig. ~~3~~4) and formed in early December 2017 (Fig. ~~4~~5).

3.3 Locate avalanche problems in terrain

When locating avalanche problems in terrain, the description of the terrain depends on the context and scale of the forecast (Statham et al., 2018). For example, regional forecasters describe terrain by elevation bands and aspects while highway forecasters reference named avalanche paths. Partitioning snowpack data into distinct terrain classes and comparing side-by-side views of the data for each terrain class is an effective way to visualize complex geospatial patterns. High-dimension (3D) visualizations are tempting to characterize mountainous terrain, particularly with high density model datasets, but there is large potential for misinterpretation on two-dimensional displays due to depth perception issues and over-plotting (Ware, 2012). Instead, ~~using eyes to simultaneously compare~~ simultaneously comparing one- or two-dimensional visualizations for different types of terrain has low cognitive load and less potential for misinterpretation.

~~Snowpack layers from 1180 simulated profiles partitioned into terrain class bins for elevation band and aspect. Elevation bins include alpine (ALP), treeline (TL), and below treeline (BTL) and aspect bins include four cardinal directions (north, east, south, west). Each layer is given a random horizontal position within the bin to allow visual summary statistics. Grain types are coloured using the perception-informed palette from Table 1.~~

To provide insight into the ~~spatial distribution of avalanche problem characteristics~~location of avalanche problems, the simulated profiles from Glacier National Park were partitioned into bins for elevation band and aspect classes to support regional-scale forecasting (Fig. ~~5).~~ 6). Avalanche forecasters often use radial plots to visualize simple aspect and elevation patterns such as danger ratings or the presence of an avalanche problem. While radial plots are familiar and widely used because of their metaphor for cardinal directions, the skewed and unaligned coordinate plane makes precise comparisons much more difficult. Given the complexity of snowpack model data, Fig. 6 uses rectilinear plots with an aligned scale for more accurate comparisons (Cleveland and McGill, 1984). A randomized horizontal position (i.e. jitter) was applied to each layer to reduce over-plotting and randomize the order within a bin (Ellis and Dix, 2007). The jitter plot allows the user to derive visual summary statistics about the snowpack structure in each terrain class and make comparisons between different terrain bins such as:

- snow depth generally increases with elevation, except for south and west facing slopes in the alpine,
- there is more new snow on north and east aspects,
- buried surface hoar layers are more prevalent on north and east aspects, and
- the early December 2017 weak layer is more prevalent at treeline and below treeline elevations.

These types of visual patterns could help forecasters localize avalanche problems in their terrain. Different types of terrain bins could be applied for other forecasting contexts to highlight differences between relevant types of terrain ~~. Examples include~~ such as sub-regions, avalanche paths, or classes of ski terrain (e.g. Sterchi et al., 2019).

3.4 Compare distributions of avalanche size and likelihood

Avalanche size is easily visualized by aligning layers by depth rather than height. Layer depth is more relevant to forecasting avalanches than layer height, as weak layer depths correlate to the destructive potential of slab avalanches (McClung, 2009). From an information visualization perspective, comparisons are more effective on aligned scales, and thus aligning layers by depth allows users to browse the distribution of depths for specific weak layers. From the distribution of layer depths in Fig. ~~3~~ 4 and Fig. ~~5~~6, forecasters could estimate the potential sizes of storm slab and persistent slab avalanches. The distribution of layer depths in these visualizations relates to spatial variability amongst the profile locations. Overlaying summary statistics on the visualizations, such as the median depth of a specific layer, could further help estimating the size of avalanches in different types of terrain (as done in the interactive dashboard in Sect. 3.5).

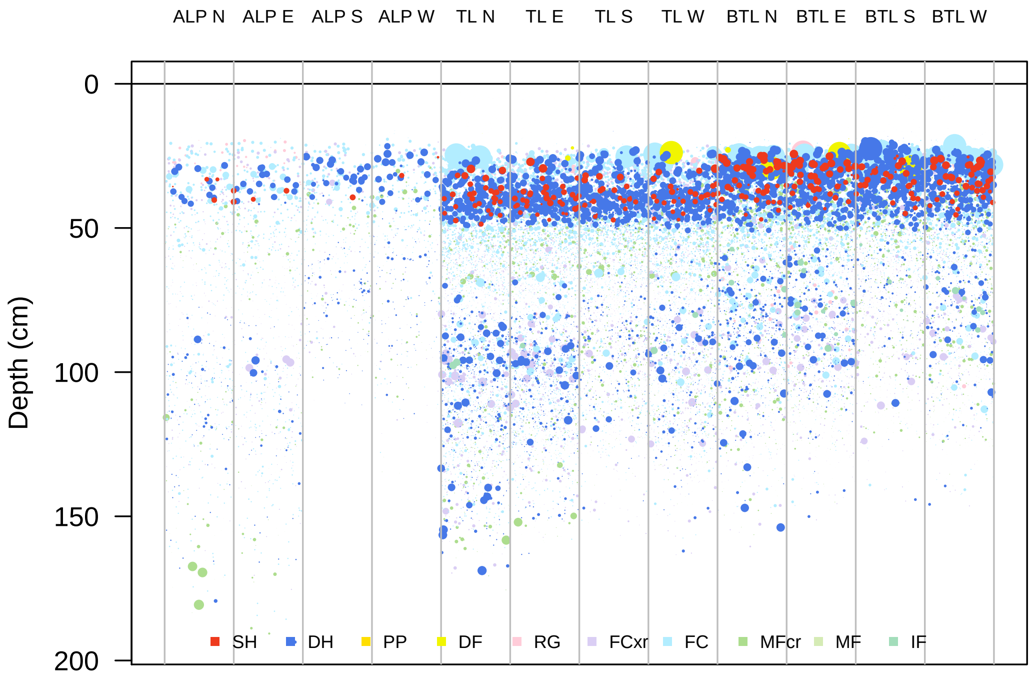

Figure 7. Visualization designed to show the likelihood of persistent slab avalanches by combining spatial distribution and the sensitivity to triggers of snowpack layers. Snowpack layers from 1180 simulated profiles on 8 January 2018 are partitioned into terrain class bins for elevation band and aspect. Elevation bins include alpine (ALP), treeline (TL), and below treeline (BTL) and aspect bins include four cardinal directions (north, east, south, west). The number of dots with persistent grain types in a terrain bin relates to the spatial density of the problem and the size of each layer's dot relates to its sensitivity to triggers (derived from the structural stability index). Each layer is given a random horizontal position within the bin to allow visual summary statistics. Grain types are coloured using the perception-informed palette from Table 1.

The CMAH defines the likelihood of avalanches as a combination of ~~*sensitivity to triggers* and *spatial distribution*~~ sensitivity to triggers and spatial distribution (Statham et al., 2018), making it a relatively difficult attribute to visualize. Options for visualizing avalanche likelihood could include encoding related attributes with visual features such as shape, size, or motion in any of the previous idioms or by designing new idioms that focus specifically on likelihood. ~~Information about the *spatial distribution*~~ We present examples of both approaches using some simple attributes related to sensitivity to triggers and spatial distribution.

When working with snow profiles, one potential method for assessing the spatial distribution of a problem ~~can be derived by counting relevant features~~ is counting relevant layers amongst a set of profiles ~~(e.g. Fig. 4). *Sensitivity to triggers* is an assessment of snowpack instability, which snowpack models estimate using~~ as an indication of spatial density. Meanwhile, sensitivity to triggers can potentially be assessed with snowpack tests, stability indices, or structural criteria such as grain

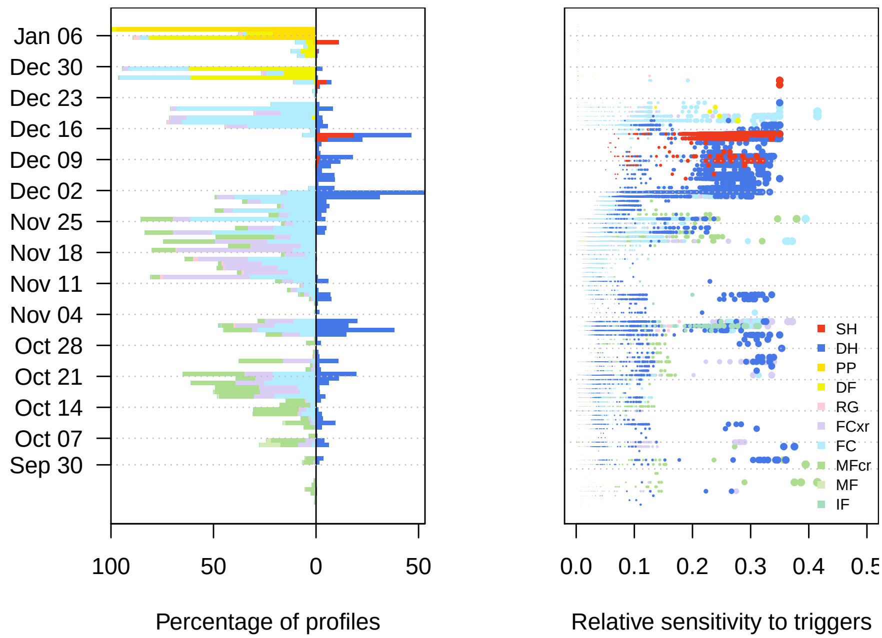

Figure 8. Combining visualizations of spatial distribution and sensitivity to triggers to provide information about the likelihood of avalanches from 1180 simulated profiles on 8 January 2018. Both visualizations aggregate the layers by age and colour them by grain type. The left panel shows spatial distribution by counting the number of profiles with different grain types (see Fig. 5) and the right panel shows the distribution of sensitivity to triggers for these same layers as derived from the structural stability index (with dot size proportional to sensitivity to triggers). Grain types are coloured using the perception-informed palette from Table 1.

size and hardness (Schweizer and Jamieson, 2007). Snowpack models offer several stability indexes based on the mechanical and structural properties of the layers (Schweizer et al., 2006). We derive a relative measure of sensitivity to triggers (S) from SNOWPACK's structural stability index (SSI). The SSI combines a stress-strength ratio with differences in hardness and grain size to calculate a value between 0 and 6, where lower values correspond to less stable layers. The SSI is most effective for avalanche problems associated with deep weak layers (e.g. persistent slab problems), because it ignores surface layers within skier penetration depth. To visually emphasize unstable layers, SSI was transformed into a relative measure of sensitivity to triggers:

$$S \propto exp^{-SSI} \tag{1}$$

where the SSI for each layer is scaled inverse exponentially to produce an ordered variable that correlates with the sensitivity categories from the CMAH (i.e. unreactive, stubborn, reactive, touchy). This transformation produces values between 0 and 1 and exaggerates differences for unstable weak layers with low SSI. The numeric value of the sensitivity measure does not have an interpretable meaning but ~~illustrates~~ can illustrate relative patterns when applied in visualizations.

~~Providing information about the likelihood of persistent slabs avalanches by scaling the size of each layer's dot with its sensitivity to triggers (derived from the structural stability index). Snowpack layers from 1180 simulated profiles are partitioned into terrain class bins for elevation band and aspect. Elevation bins include alpine (ALP), treeline (TL), and below treeline (BTL) and aspect bins include four cardinal directions (north, east, south, west). Each layer is given a random horizontal position within the bin to allow visual summary statistics. Grain types are coloured using the perception-informed palette from Table 1.~~

~~Combining visualizations of spatial distribution and sensitivity to triggers to provide information about the likelihood of avalanches. Both visualizations aggregate the layers by age and colour them by grain type. The left panel shows spatial distribution by counting the number of profiles with different grain types (see Fig. 4) and the right panel shows the distribution of sensitivity to triggers for these same layers as derived from the structural stability index. Grain types are coloured using the perception-informed palette from Table 1.~~

We present two examples of visualizing likelihood information with this relative measure for sensitivity to triggers. The terrain class visualization in Fig. ~~5~~ 6 was modified to scale the dot size of each layer to its sensitivity to triggers (Fig. ~~6~~7). This creates greater emphasis on sensitive weak layers, ~~where the~~ so the combination of the number and size of weak layer dots in a terrain bin relate to the likelihood of persistent slab avalanches in that type of terrain. Another visualization specifically designed for likelihood is given in Fig. ~~7~~8, where the left panel provides information about the spatial distribution of each layer and the right panel provides information about their sensitivity to triggers. ~~Spatial~~ Information about the spatial distribution is shown ~~by the~~ with the same visualization as Fig. 5, where the prevalence of each layer by age ~~(i. e. Fig. 4), while sensitivity~~ is related to the spatial density of the problem. Sensitivity to triggers is shown with the distribution of the relative sensitivity of each layer by age. The side-by-side comparison of spatial distribution and sensitivity to triggers provides information about the potential likelihood of persistent slab avalanche problems. For example, the weak layers that formed in early December 2017 are more widely distributed and sensitive to triggers than the weak layers that formed in late October (i.e. avalanches are more likely).

It is important to note that we are presenting these likelihood visualizations ~~more~~ to illustrate the concept ~~than as a practical decision aid. It is known that the modelled stability index does not provide meaningful information about layers near the surface where storm slab avalanches occur (Schweizer et al., 2006), and Monti et al. (2014) has highlighted issues between the modelled stability indices and field observations of snowpack instability~~of visually encoding stability information rather than suggest these derivations for an operational tool. These derivations are most effective for persistent and deep persistent slab avalanche problems, while the likelihood of other avalanche problem types may be better represented by other attributes such as weather variables or snow temperatures (Haegeli et al., 2010). Deriving stability information from simulated snow profiles is an active research topic (Monti et al., 2014), and new stability indices will likely provide more accurate information about the likelihood of avalanches.

3.5 Interactive dashboard

~~Visualizations~~

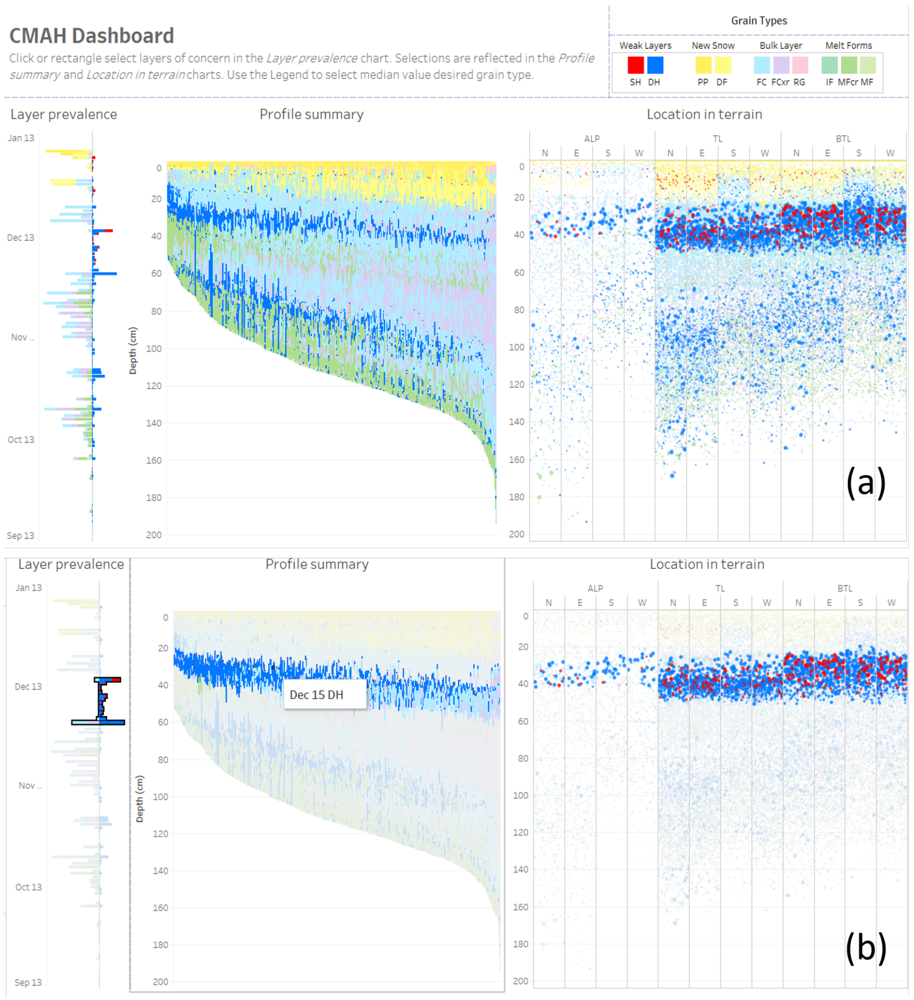

Figure 9. Screenshots of an interactive dashboard that provides visualizations of layer prevalence, profile summary, and location in terrain from 1180 simulated profiles on 8 January 2018. The initial view (a) provides and overview of the entire dataset for the user to assess potential avalanche problems and then (b) the updated view after the user has selected layers that formed between 2 and 15 December 2017 to explore details about the distribution and depth of the persistent slab avalanche problem.

The visualizations presented in this section were combined into an interactive dashboard using ~~Tableau~~ *Tableau* data visualization software (Fig. ~~8 and available online in Code and data availability~~9). The dashboard facilitates the sequential questions of the CMAH by following the “overview first, zoom and filter, details on demand” mantra (Shneiderman, 1996)~~via interactions~~. Interactions that allow the user to change the view by selecting visual features and filters from the legend. The initial view (Fig. ~~8a~~9a) consists of the layer prevalence visualization ~~(Fig. 4)~~from Fig. 5, the profile summary visualization ~~(Fig. 3)~~from Fig. 4, and the location in terrain visualization ~~(Fig. 5).~~ from Fig. 6. The combination of these visualizations provides a visual overview of the snowpack structure to support the first question in the CMAH – identifying potential avalanche problem types~~(Fig. 8a).~~. After identifying potential avalanche problem types from the overview visualizations, users select layers of concern from the layer prevalence panel to update the visualizations. Once a layer of concern is selected, the layer is highlighted in the other panels to provide details about the location in terrain and the distribution of avalanche sizes (Fig. ~~8b~~9b). Horizontal bars show the median depth of the selected layer in each terrain class for comparison of potential avalanche sizes. A tooltip allows the user to hover over any visual feature and see details such as the grain type, deposition date, and depth in a pop-up window. In Fig. ~~8b~~9b, the user has selected all the layers that formed between 2 and 15 December 2017 to investigate the persistent slab avalanche problem. The profile summary shows the position of this layer in the snowpack and the location in terrain visualization shows the layer is more prevalent at treeline and below treeline, with median depths of $40\,\mathrm{cm}$ at treeline and $35\,\mathrm{cm}$ below treeline. The appendix provides examples of this dashboard for several days throughout the 2017-18 winter.

~~Screenshots of an interactive dashboard that provides visualizations of layer prevalence, profile summary, and location in terrain. The initial view (a) provides and overview of the entire dataset for the user to assess potential avalanche problems and then (b) the updated view after the user has selected layers that formed between 2 and 15 December 2017 to explore details about the distribution and depth of the persistent slab avalanche problem.~~

4 Implementation

5 ~~Validating visualization designs~~

4.1 Design considerations

~~While the principles outline~~The visualizations presented in Sect. ~~2 are a good foundation for designing meaningful visualizations,~~ 3 are a starting point of how information from snowpack models can be designed to address specific forecasting needs, but additional user testing is ~~critical to ensure the visualizations have the desired effects. We presented the visualizations from the previous section to avalanche forecasters at ten workshop-style presentations, tested real-time prototypes with three helicopter skiing operations over two winters, and ran an exercise where five users performed basic tasks with the interactive dashboard. Following the~~ necessary for them to evolve into a valuable forecasting tool. The nested model for visualization design ~~framework of Munzner (2009) again, we present the feedback that we received at each design level separately.~~of Munzner (2009) provides a structured approach to evaluating the design of such tools, where issues can be addressed at each specific design level (i.e. domain situation, task and data abstraction, visualization and interaction idiom, algorithm).

At the domain situation level, creating links between snowpack models and ~~the CMAH addresses~~ avalanche forecasting workflows like the CMAH will address operational challenges faced by ~~avalanche forecasters(Statham et al., 2018)~~forecasters. Reflecting the broad adoption of the CMAH (Statham et al., 2018), the proposition of using snowpack models to characterize avalanche problems across forecast regions has gained more interest from the Canadian forecasting community than previous snowpack model tools ~~produced over the past decade. The~~ that focused on individual stratigraphy profiles. However, the CMAH may not characterize ~~the domain situation for all possible snowpack model users, as problems~~ all possible domain situations for snowpack models, as tasks such as terrain selection or civil protection ~~likely require distinct design~~ possibly require distinct domain level considerations.

At the task and data abstraction level, ~~the visualization of snowpack summaries has received consistent positive feedback from forecasters~~designs should focus on specific questions and forecasting tasks. This requires a shift from bottom-up scientific visualizations towards top-down information visualizations. The visualizations in Sect. 3 are specifically designed to answer the four questions posed by the CMAH by focusing on the type of task (e.g. identify, locate, compare). Forecasters have existing methods for performing these tasks with field data, but aggregating and summarizing that data can be challenging and uncertain. The continuous spatial coverage of snowpack models offers unique opportunities to support these tasks. The side-by-side ~~profile summary~~ profiles (Fig. ~~3) visualizes~~ 4) and terrain class plots (Fig. 6) visualize snowpack patterns in ~~a way that is~~ ways that are not possible with traditional snow profile data and can help forecasters build a more complete mental model of the snowpack structure in their forecast area. ~~Other operational tasks could benefit from bottom-up designs that leverage the spatial and temporal coverage of snowpack models, such as using stratigraphy timelines to visualize temporal trends .~~

While the examples in Sect. 3 are particularly suited to storm and persistent slab avalanche problem types, the same principles could be applied to emphasize attributes important to other problem types (such as weather data to identify wind slab avalanche problems and snow temperature data to identify wet avalanche problem types). The task of locating avalanche problems in terrain differs for different forecasting contexts. In many cases maps or other geospatial visualizations could be valuable for this task. While not directly a question in the CMAH, the task of tracking temporal trends is also critical, as the forecasting process is iterative throughout the winter. The continuous temporal data provided by snowpack models offers unique capabilities for tracking snowpack evolution. Stratigraphy timeline visualizations (e.g. Fig. 3) are well suited for tracking snowpack evolution at individual locations, however adding a temporal dimension to spatial information creates additional complexity and requires specific design considerations. A visualization showing the temporal evolution of a snowpack summary would be particularly interesting. Examples of temporal change are showing in the Appendix with examples of the interactive dashboard for six different periods over the 2017-18 winter.

At the visualization and interaction idiom level, ~~some forecasters suggested reducing the number of colours in snow profile visualizations to make them easier to interpret~~following established perceptual and cognitive principles ensures designs are effective at their intended tasks. The perception-informed colour palettes (Table 1 and 2) ~~achieve this while following established perceptual and cognitive~~ are examples of applying these principles to draw attention to the ~~most important features. The user testing exercise evaluated the users' ability to interpret the visualizations by performing simple tasks with the interactive dashboard. Four out of five participants correctly performed task such as comparing snow height over different~~

~~elevations, identifying the depth of prominent weak layers , and summarizing new snow amounts. The remaining participant made mistakes with filtering and selection, highlighting the importance of designing interactions that are simple and intuitive.~~ features in snow profiles that are deemed most important. The standard grain colour palette may emphasize features without intending to do so and is likely ineffective for individuals with colour blindness. The simplified colour palette in Table 2 could potentially be more relevant for forecasters as it shifts the purpose of colour from showing snowpack structure towards identifying avalanche problems. Considering information visualization principles listed in Sect. 2.4 could prevent data from being misinterpreted.

~~At the algorithm level, the operational prototypes provided daily updated visualizations in a timely manner with fast response time for interactions. The main concern at the algorithm level was the prototypes were accessed externally from existing workstations, which created a major barrier to access. Integrating snowpack model visualizations into forecasting workstations is a critical next step. Testing in an operational setting would allow further validation at the domain and abstraction levels by measuring user adoption and observing how designs are used to perform operational tasks. Although the designs presented in this paper follow established visualizationprinciples, testing in real forecasting scenarios is needed to validate their actual operational value.~~ At the algorithm level, interactive tools need to be efficient in terms of time and memory performance. We tested several versions of interactive dashboards with operational forecasters. While these dashboards were not optimized for web performance, they worked at reasonable speeds with maximum wait times of 2-3 seconds for filtering layers in large regions with over 5000 profiles. Over-plotting becomes an issue for large data sets where the total number of layers becomes larger than the number of pixels on the screen, but can be addressed by stratified sampling or implementing subpixel rendering techniques to increase the apparent resolution of the screen. The visualizations may also be less effective with small data sets where there are not enough layers for patterns to emerge. This could be addressed by downscaling the model to increase the number of modelled profiles. An important consideration in data set size is ensuring the model is configured to capture an appropriate amount of spatial variability across the forecast region.

4.2 Steps towards operational implementation

While these designs in Sect. 3 are informed with well established visualization principles, user testing is critical to validate their actual operational value. Various versions of the interactive dashboard presented in Sect. 3.5 have been tested with operational forecasters in Canada, resulting in an agile development process where qualitative feedback has provided new perspectives and identified issues with the designs. An iterative process of feedback and redesign is critical for successful implementation of new visualization tools into operational workflow and is much less risk-prone than developing visualization tools in entirety. For example, the US National Weather Service used an agile development process to deploy their modern forecasting tool over several years in the early 2000s (LeFebvre et al., 2003).

The visualization design principles presented for snowpack model data are equally relevant for visualizing traditional field data. An ideal implementation of snowpack models into forecasting workflows would be combining field data and model data into a single interactive tool. A major motivation for adding model data into forecasting workflows is to reduce uncertainty about snowpack conditions. A visualization tool with mixed data sources would allow forecasters to assess the integrity of the model output as well as place the field observations into a broader context with the continuous spatial coverage of models. Many of the visualizations presented in Sect. 3 could be modified for such comparison tasks. Similarly, visualizing an ensemble of model data sets (e.g. with different meteorological inputs or geometric configurations) would provide insights about the confidence in modelled data.

In addition to improved visualization, model development and validation remains critical to improving the integrity of model output. This should continue in parallel to user testing so forecasters can offer operational feedback on model accuracy. Assimilating field data into snowpack models could greatly improve their integrity (Winstral et al., 2018), however model developers are faced with assimilation challenges such as mismatched spatial scales between gridded models and point field observations. Interactive visualizations of heterogenous field and model data has potential for researchers and forecasters to gain a deeper understanding of how they relate, and the knowledge gained through such a process can translate to improved computational assimilation methods.

5 Conclusions

We present visualization design principles that increase the ~~*interpretability* and *relevance*~~ interpretability and relevance of snowpack model outputs. These are two of the four major perceived issues with operational snowpack model tools identified by ~~?~~Morin et al. (2020) (besides accessibility and integrity). The nested model for visualization design (Munzner, 2009) provides a framework for defining the domain of avalanche forecasting and the necessary tasks that are needed to analyze data. Tasks required to assess avalanche hazard are described by applying the widely adopted conceptual model of avalanche hazard (Statham et al., 2018). From these tasks, we ~~apply~~ show how information visualization principles can be applied to design visual representations of snowpack model data in ways that leverage the human visual system to understand the complex nature of the data. ~~Preliminary feedback from avalanche forecasters suggests these designs are easier to interpret and provide more relevant information than previous visualizations of snowpack model data.~~

A key idea in these designs is shifting from bottom-up scientific visualizations towards information visualizations that address user needs. ~~As highlighted by Grainger et al. (2016), other types of environmental models would likely see improved adoption by shifting towards information visualization. When using numeric models as a tool for assessing natural hazards, visualizations will be more effective when the designers make links to established risk frameworks and carefully consider the tasks performed by operational decision makers.~~

A critical next step is implementing ~~these designs into~~ and testing these designs in operational forecasting workflows. By addressing issues with the interpretability and relevance of snowpack model data, these designs will allow forecasters to learn the capabilities and deficiencies of snowpack models in a meaningful way. The same design principles should be considered when visualizing other types of avalanche and snowpack data, as the same domain situation and task abstractions apply when ~~analyzing~~ forecasters analyze field observations. Interaction idioms should play an important role in understanding ~~of~~ complex model data, as they allow users to perform custom queries, test and validate hypotheses, and discover inconsistencies and anomalies. Interactions that compare model data with observation data would be particularly powerful in building trust in

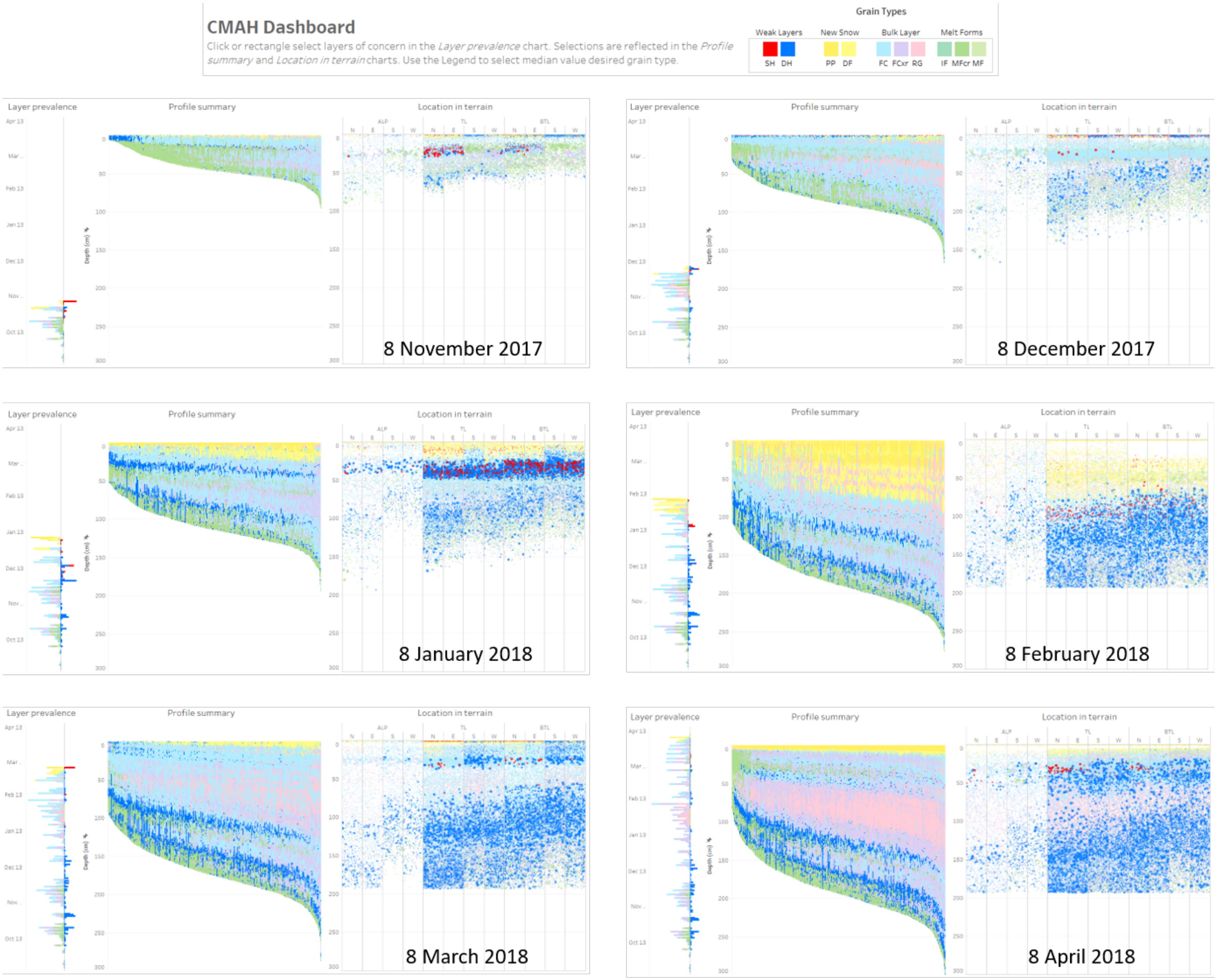

Figure 10. Comparison of the interactive dashboard for different days over the course of the 2017-18 winter. Each dashboard includes visualizations of layer prevalence, a profile summary, and location in terrain for the same 1180 simulated profiles in Glacier National Park.

the models and addressing issues with their integrity. This process was critical in the adoption and trust in numeric weather predictions models by meteorologists (Benjamin et al., 2019), and just like meteorologists, avalanche forecasters could become active participants in model validation and improvement.

Appendix

. The code and data used to produce the visualizations are publicly available at https://osf.io/8wz2v (Horton, 2020). The interactive dashboard is available at https://avalancheresearch.ca/pubs/2019_horton_snowpackvis.

. All authors worked on the conceptualization of this paper. SH prepared the data and software, SN contributed to visualization ideas and designs, and PH provided supervision. SH prepared the manuscript with review and editing from the other authors.

. The authors declare no competing interests.

. Thanks to Lyn Bartram of the Vancouver Institute of Visual Analytics, the Big Data Hub at Simon Fraser University (SFU), and the rest of our colleagues at the SFU Avalanche Research Program. The NSERC Industrial Research Chair in Avalanche Risk Management at SFU is financially supported by Canadian Pacific Railway, Helicat Canada, the Canadian Avalanche Association, and Mike Wiegele Helicopter Skiing. The SFU Avalanche Research Program is further supported by Avalanche Canada and the Avalanche Canada Foundation. We thank Karsten Müller, Jan-Thomas Fischer, and Michael Warscher for their constructive reviews.

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, Meteor. Mono., 59, 13.1–13.67, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0020.1, 2019.

Birch, J.: Worldwide prevalence of red-green color deficiency, J. Opt. Soc. Am, 29, 313–320, https://doi.org/10.1364/JOSAA.29.000313, 2012.

Blackwell, A. F.: Pictorial representation and metaphor in visual language design, J. Visual Lang. Comput., 12, 223–252, https://doi.org/10.1006/jvlc.2001.0207, 2001.

Bovee, M., Srivastava, R. P., and Mak, B.: A conceptual framework and belief – function approach to assessing overall information quality, Int. J. Intell. Syst., 18, 51–74, https://doi.org/10.1002/int.10074, 2003.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, https://doi.org/10.3189/S0022143000009552, 1992.

Canadian Avalanche Association: Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches., Revelstoke, BC, Canada, 2016a.

Canadian Avalanche Association: Technical Aspects of Snow Avalanche Risk Management – Resources and Guidelines for Avalanche Practitioners in Canada, Revelstoke, BC, Canada, 2016b.

Cleveland, W. S. and McGill, R.: Graphical perception: Theory, experimentation, and application to the development of graphical methods, J. Am. Stat. Assoc., 79, 531–554, https://doi.org/10.1080/01621459.1984.10478080, 1984.

De Mauro, A., Greco, M., and Grimaldi, M.: A formal definition of Big Data based on its essential features, Library Rev., 65, 122–135, https://doi.org/10.1108/LR-06-2015-0061, 2016.

Ellis, G. and Dix, A.: A taxonomy of clutter reduction for information visualisation, IEEE T. Vis. Comput. Gr., 13, 1216–1223, https://doi.org/10.1109/TVCG.2007.70535, 2007.

Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., McClung, D., Nishimura, K., Satyawali, P., and Sokratov, S.: The International Classification for Seasonal Snow on the Ground, IHP-VII Tech. Doc. Hyrdol. No. 83, IACS Contrib. No. 1, UNESCO-IHP, Paris, 2009.

Grainger, S., Mao, F., and Buytaert, W.: Environmental data visualisation for non-scientific contexts: Literature review and design framework, Environ. Modell. Softw., 85, 299–318, https://doi.org/10.1016/j.envsoft.2016.09.004, 2016.

Haegeli, P., Atkins, R., and Klassen, K.: Decision making in avalanche terrain - a field book for winter backcountry users, Canadian Avalanche Centre, Revelstoke, BC, Canada, 2010.

Haegeli, P., Obad, J., Harrison, B., Murray, B., Engblom, J., and Neufeld, J.: InfoEx 3.0 – Advancing the data analysis capabilities of Canada's diverse avalanche community, in: Proceedings of the 2014 International Snow Science Workshop, pp. 910–917, 2014.

Hagenmuller, P. and Pilloix, T.: A new method for comparing and matching snow profiles, application for profiles measured by penetrometers, Front. Earth Sci., 4, 52, https://doi.org/10.3389/feart.2016.00052, 2016.

Horton, S.: Snow profile visualization, osf.io/8wz2v, 2020.

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., and Ziegler, H.: Visual analytics: Scope and challenges, pp. 76–90, Springer, https://doi.org/10.1007/978-3-540-71080-6_6, 2008.

LaChapelle, E.: The fundamental processes in conventional avalanche forecasting, J. Glaciol., 26, 75–84, https://doi.org/10.3189/S0022143000010601, 1980.

LeFebvre, T., Mathewson, M., and Hansen, T.: The Rapid Prototype Project, in: 19th Int. Conf. on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, vol. 12.4, 2003.

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., and Zimmerli, M.: SNOWPACK model calculations for avalanche warning based upon a new network of weather and snow stations, Cold Reg. Sci. Technol., 30, 145–157, https://doi.org/10.1016/S0165-232X(99)00022-1, 1999.

Maguire, L. and Percival, J.: Sensemaking in the snow: Exploring the cognitive work in avalanche forecasting, in: Proceedings of the 2018 International Snow Science Workshop, 2018.

McClung, D.: The elements of applied avalanche forecasting, Part I: The human issues, Nat. Hazards, 26, 111–129, https://doi.org/10.1023/A:1015665432221, 2002.

McClung, D.: Dimensions of dry snow slab avalanches from field measurements, J. Geophysi. Res. – Earth, 114, https://doi.org/10.1029/2007JF000941, 2009.

Milbrandt, J. A., Bélair, S., Faucher, M., Vallée, M., Carrera, M. L., and Glazer, A.: The Pan-Canadian High Resolution (2.5 km) Deterministic Prediction System, Weather Forecast., 31, 1791–1816, https://doi.org/10.1175/waf-d-16-0035.1, 2016.

Monti, F., Schweizer, J., and Fierz, C.: Hardness estimation and weak layer detection in simulated snow stratigraphy, Cold Reg. Sci. Technol., 103, 82–90, https://doi.org/10.1016/j.coldregions.2014.03.009, 2014.

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Ližar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J., 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, 102 910, https://doi.org/10.1016/j.coldregions.2019.102910, 2020.

Munzner, T.: A nested model for visualization design and validation, IEEE T. Vis. Comput. Gr., 15, 921–928, https://doi.org/10.1109/TVCG.2009.111, 2009.

Munzner, T.: Visualization Analysis and Design, Taylor and Francis, Boca Raton, FL, USA, https://doi.org/10.1201/b17511, 2014.

Parks Canada: Avalanche Bulletin Glacier National Park Mon Jan 8 2018, https://avalanche.pc.gc.ca/bulletin-eng.aspx?r=3&d=2018-01-08, last access: 31 August 2019, 2018.

Rautenhaus, M., Böttinger, M., Siemen, S., Hoffman, R., Kirby, R. M., Mirzargar, M., Röber, N., and Westermann, R.: Visualization in meteorology — A survey of techniques and tools for data analysis tasks, IEEE T. Vis. Comput. Gr., 24, 3268–3296, https://doi.org/10.1109/TVCG.2017.2779501, 2018.

Schweizer, J. and Jamieson, 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., Bellaire, S., Fierz, C., Lehning, M., and Pielmeier, C.: Evaluating and improving the stability predictions of the snow cover model SNOWPACK, Cold Reg. Sci. Technol., 46, 52–59, https://doi.org/10.1016/j.coldregions.2006.05.007, 2006.

Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations, in: Proceedings 1996 IEEE Symposium on Visual Languages, pp. 336–343, https://doi.org/10.1109/VL.1996.545307, 1996.

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

Stauffer, R., Mayr, G. J., Dabernig, M., and Zeileis, A.: Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations, B. Am. Meteorol. Soc., 96, 203–216, https://doi.org/10.1175/BAMS-D-13-00155.1, 2015.

Sterchi, R., Haegeli, P., and Mair, P.: Exploring the relationship between avalanche hazard and run list terrain choices at a helicopter skiing operation, Nat. Hazard Earth Sys., 19, 2011–2026, https://doi.org/10.5194/nhess-19-2011-2019, 2019.

Szafir, D. A., Haroz, S., Gleicher, M., and Franconeri, S.: Four types of ensemble coding in data visualizations, J. Vision, 16, 11–11, https://doi.org/10.1167/16.5.11, 2016.

Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M., and DiMarco, S. F.: True colors of oceanography: Guidelines for effective and accurate colormap selection, Oceanography, 29, 9–13, https://doi.org/10.5670/oceanog.2016.66, 2016.

Ware, C.: Information visualization: Perception for design, Elsevier, Waltham, MA, USA, https://doi.org/10.1016/C2009-0-62432-6, 2012.

Winstral, A., Magnusson, J., Schirmer, M., and Jonas, T.: The bias detecting ensemble: A new and efficient technique for dynamically incorporating observations into physics-based, multi-layer, snow models, Water Resour. Res., https://doi.org/10.1029/2018WR024521, 2018.