ESTABLISHING THE LINK BETWEEN THE CONCEPTUAL MODEL OF AVALANCHE HAZARD AND THE NORTH AMERICAN PUBLIC AVALANCHE DANGER SCALE: INITIAL EXPLORATIONS FROM CANADA

Taylor Clark* and Pascal Haegeli

School of Resource and Environmental Management, Simon Fraser University, Burnaby, BC, Canada

ABSTRACT: In 2010, Statham and colleagues introduced the conceptual model of avalanche hazard (CMAH; Statham et al., 2018) to make avalanche bulletin production in North America more transparent and consistent. Since the CMAH did not provide a prescriptive link between hazard assessments and avalanche danger ratings, forecasters need to rely on their own judgment to assign danger ratings, which can lead to inconsistencies in public avalanche risk communication. The present paper aims to address this missing link by exploring the relationship between avalanche hazard assessments and danger ratings within Canadian avalanche bulletins since the introduction of the CMAH. Using conditional inference trees, key decision rules that forecasters utilize to assign avalanche danger rating in Canada are extracted. Our results offer insight on how Canadian public forecasters assigned danger ratings and provide an evidence-based platform for discussing inconsistency and addressing them through the development of meaningful decision-aids.

KEYWORDS: Danger Rating, Avalanche Hazard, Forecasting, Decision Trees

1. INTRODUCTION

Public avalanche bulletins are a key source of information for backcountry recreationalists planning trips into avalanche terrain. Canadian avalanche bulletins consist of avalanche danger ratings that succinctly describe current and future hazard conditions according to the North American public avalanche danger scale (Statham et al., 2010), a structured description of the avalanche problems responsible for the hazard, and a detailed description of observed and expected avalanche activity, snowpack characteristics and weather conditions.

It is well established that consistency is a key attribute of high-quality forecast products (Murphy 1993) and effective public warning messages (Mileti and Sorensen 1990). Several studies have recently examined the topic in the context of public avalanche bulletins. Lazar et al. (2016) examined inconsistencies in danger rating assignments by asking forecasters to evaluate avalanche condition scenarios in an online survey. The study found consensus at the extreme ends of the danger scale (low and extreme) but substantial differences in the middle. Techel et al. (in review) examined spatial consistency of forecast danger ratings in European alpine countries and found considerable differences and biases across national and forecast center boundaries.

* *Corresponding author address:*
Taylor Clark, Simon Fraser University
School for Resource & Environ. Management,
8888 University Drive, Burnaby, BC;
V5A 1S6, Canada
email: taylor_clark@sfu.ca
web: www.avalancheresearch.ca

To make avalanche bulletin production in North America more transparent and consistent, Stethem and colleagues developed the conceptual model of avalanche hazard (CMAH; Statham et al. 2018). The CMAH standardizes language and structures the workflow for assessing avalanche hazard according to four essential questions for avalanche risk mitigation: *what* type(s) of avalanche problems are present, *where* are they located in the terrain, how *likely* are associated avalanches and how *big* will these avalanches be. In Canada, Avalanche Canada and Parks Canada have used the CMAH for the production of public avalanche bulletins since the winter of 2010 and 2012 respectively.

Even though the CMAH structures the hazard evaluation process, Statham et al. (2018) deliberately chose not provide explicit guidance for assigning danger ratings to hazard situations similar to the Bavarian Matrix (EAWS 2018) used by the European avalanche warning services. There were two reasons for this choice. First, the CMAH describes the fundamental principles of avalanche hazard assessment independent of application and not all types of avalanche hazard assessments are used to produce public avalanche danger ratings. Second, the authors believed that the relationship between avalanche hazard and danger ratings to be too complex for a small group of experts to prescriptively define danger rating levels for different combinations of avalanche problems without broad consultation and extensive evidence. Instead, the idea was to use the CMAH in practice and use the systematically collected assessment data to establish the missing link empirically (Statham et al. 2018).

Haegeli, Falk, & Klassen (2012) were the first to use CMAH-compliant hazard assessment data to examine the association between avalanche hazard and danger ratings. Their ordinal logistic regression model revealed that maximum likelihood of avalanches and maximum expected destructive size had the strongest influence on ratings, but also highlighted considerable variation among forecasters. However, the model assumed a linear relationship between problem characteristics and danger rating and was limited to only two winters of data.

The objective of this research is to address the missing link between the CMAH and the danger scale and provide quantitative insights into danger rating assignments by systematically examining the combined operational avalanche bulletin dataset of Avalanche Canada and Parks Canada.

2. METHODS

2.1 Dataset

Avalanche Canada and Parks Canada produce daily avalanche bulletins for 22 regions in western Canada. The analysis dataset for this study consists of all CMAH compliant avalanche hazard assessments published by Avalanche Canada and Parks Canada between Dec. 1 and Apr. 15 since 2010 and 2012 respectively. Since each of the 14,265 bulletins in this period includes hazard assessments for three elevation bands (alpine, treeline, and below treeline), the total number of available hazard assessments is 42,589. Each hazard assessment consists of up to three avalanche problems described by their minimum, typical and maximum likelihood of avalanches and destructive size as well as sensitivity to triggering, spatial distribution and aspects the problem is present on. The hazard assessment records also include the danger rating for the time when the bulletin was published (i.e., nowcast), forecasted danger ratings for 24, 48 and 72 hours into the future, as well as identifiers for forecasting agency, forecast region and forecaster.

2.2 Conditional inference trees

We chose classification trees for our analysis as they are fully transparent and provide easily interpretable results that offer insights into the hazard assessment process that are closely linked to model the human decision-making process (e.g., Martignon et al, 2012). For our analysis, this is a decisive advantage over modern machine learning methods (e.g., random forests, neural networks) where the focus is on maximizing predictability rather than gaining insight. Further, decision trees are well suited to describe non-linear relationships.

Classification trees have previously been used in avalanche research to derive danger ratings and snow stability ratings from a variety of observed or modelled snowpack and weather variables. Recent examples include Schirmer et al. (2010), Bellaire et al. (2013) and Hendrikx et al. (2014). The splitting criteria used in traditional classification and regression trees (Breiman et al. 1998) is based on impurity of member nodes found in metrics such as the Gini index. However, trees built using this approach tend to overfit and require “pruning” to establish an effective model (Kuhn and Johnson, 2013).

For this study, we chose the alternative approach of conditional inference trees (CIT; Hothorn, Hornik, & Zeileis, 2006), which uses statistical hypothesis testing to identify meaningful splits in the dataset. At each node, the recursive partitioning algorithm evaluates all of the potential splitting rules by comparing the distributions of the dependent variable in the resulting child nodes using permutation tests (Hothorn et al. 2008). The split resulting in the biggest difference (i.e., lowest p-value) is then used to split the dataset, creating a decision node. The algorithm then repeats itself, creating further nodes until no statistically significant splits remain. The advantage of this approach is the statistical grounding of the splitting criteria, which avoids overfitting.

2.3 Analysis approach

We build several CIT trees to examine the relationship between characteristics of avalanche problems (likelihood of avalanches, destructive size) and the nowcast danger rating and explore its variability among avalanche problem types, elevation bands, and agencies. To maximize the interpretability of trees, we divided the dataset into assessments with only one problem (n = 15,020; 35% of dataset) and assessments with combinations of problems (n = 21,079; 49%). Assessments that did not include any problems were discarded (n = 6,490; 16%). We also combined high and extreme danger ratings as our complete dataset only included 27 cases of extreme.

We first created a single CIT tree for all situations with one avalanche problem. Cornice and wet slab avalanche problems were excluded due to their small number of cases. The final dataset for this tree consisted of 14,899 assessments (Tbl. 1).

Although the CMAH describes avalanche hazard with value triplets for likelihood and size, we only included typical values in our analysis. Comparisons between trees using typical or maximum values did

Tbl. 1: Number of hazard assessments

Aval. problem type	*Alpine*	*Treeline*	*Below treeline*	*Total*
Storm slabs	1,309	1,260	1,898	4,467
Wind slab	2,337	1,846	176	4,359
Persistent slab	293	495	2,009	2,797
Loose wet aval.	55	311	1,405	1,771
Deep pers. slab	207	339	429	975
Loose dry aval.	46	143	341	530
Total	4,247	4,394	6,258	14,899

not reveal any substantial differences in performance. Furthermore, personal communication with forecasters suggested that the typical values are used more consistently. Subsequently, we developed similar trees for individual problem combinations (e.g., storm and persistent slab avalanche problems).

To illustrate the results of our analysis, decision rules and associated distributions of danger ratings for different scenarios were plotted onto the hazard chart, which is commonly used in the CMAH to depict the nature of avalanche hazard as a function of likelihood of avalanches and destructive size. Predicted danger rating assignments were calculated with product sums from the danger rating distributions at the terminal nodes.

3. RESULTS/DISCUSSION

The following paragraphs provide illustrative examples of the results from our CIT tree for hazard assessments with one avalanche problem. A complete description of the results is currently in preparation for publication in a peer-reviewed journal.

Our initial visualizations of danger rating assessments for single avalanche problems revealed considerably variability. The example of Avalanche Canada danger rating assignments for storm slab avalanche problems below treeline (Fig. 1) illustrate that there are rarely any likelihood and size combinations where forecasters completely agree.

Our analysis of assessments with single avalanche problems resulted in a CIT tree with 201 decision nodes of which 101 were terminal nodes. The initial and therefore most dominant split in the tree separated storm slabs from all other avalanche problem

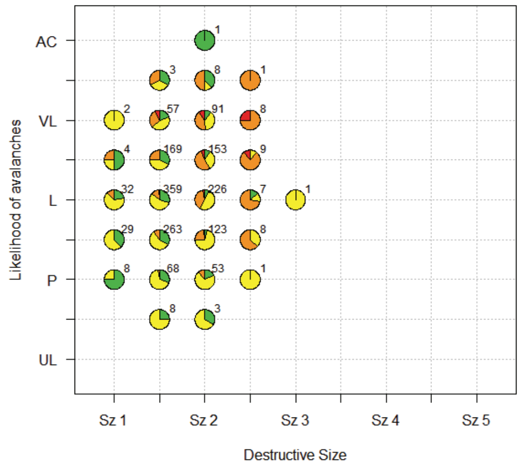

Fig. 1: Avalanche Canada storm slab hazard assessments below treeline with danger rating distribution and number of assessments (n = 1695).

types included in the analysis (Tbl. 1). This indicates that storm slab avalanches were assessed significantly more serious than other avalanche problems. This difference originates from inconsistencies in danger rating assignments as well as variation in the likelihood and size characteristics of avalanche problem types.

The storm slab section of the single avalanche problem tree (Fig. 2) further illustrates that elevation band had a significant effect on the danger level assignments as it represents the next splitting rule. Single storm slab avalanche problems were assessed significantly less severe below treeline than at treeline and in the alpine. Next, the tree separated the dataset according to size (< Size 2 and ≥ Size 2) in both elevation band branches. This indicates that for storm slabs, size had a bigger

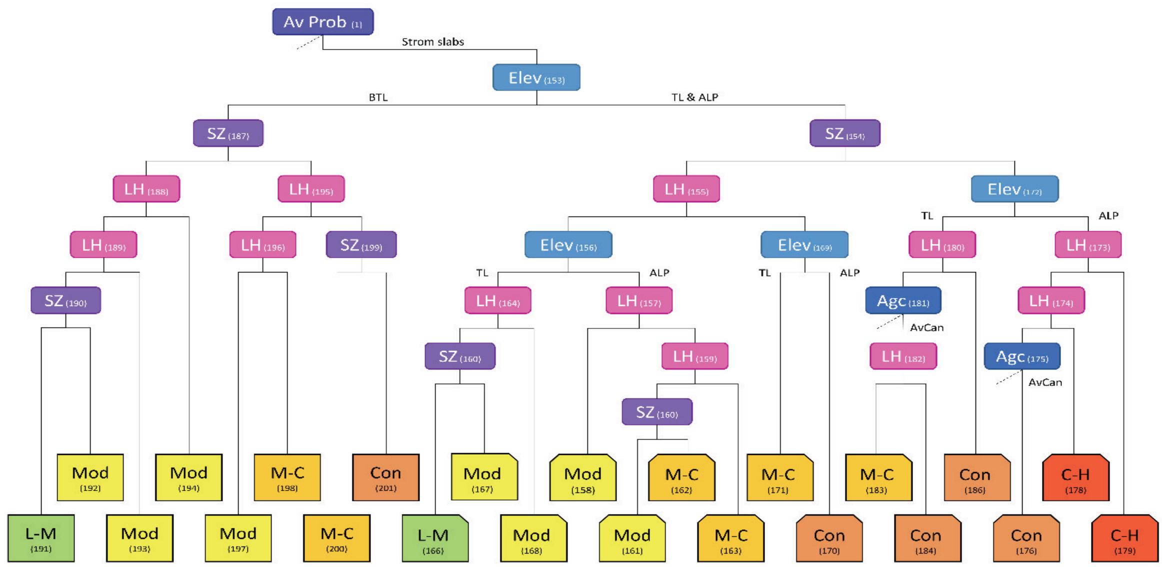

Fig. 2: Storm slab section extracted from single avalanche problem decision tree showing splits of typical likelihood (LH), typical size (SZ), elevation band (Elev) and agency (Agc) and danger rating assignments with half steps. Numbers indicate decision node.

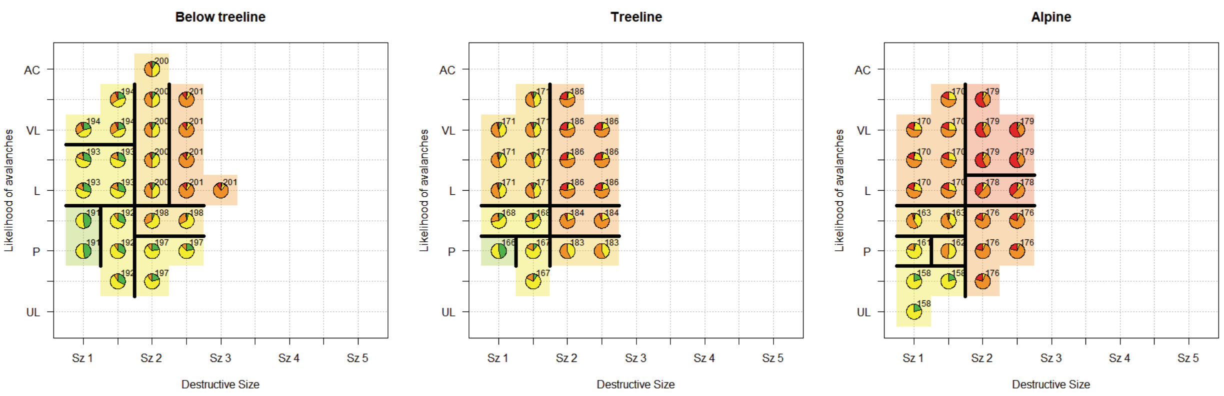

Fig. 3: Hazard assessment rules (with terminal note identifiers) and associated danger rating distributions for single storm slab avalanche problems for Avalanche Canada at likelihood and size combinations where observations exist. Background shading indicate danger rating predicted by CIT tree at intermediate steps.

effect on the danger rating assignment than likelihood, which was used further down the tree to fine-tune danger rating distributions. Only two splits were identified for agency suggesting considerable consistency in the assessments of storm slabs between Avalanche Canada and Parks Canada.

The visualization of the terminal nodes in the hazard chart (Fig. 3) more clearly highlights that single storm slab avalanche problem situations with similar likelihood and size assessments were assessed increasingly higher with increasing elevation. The illustrations further highlight the dominant split between < Size 2 and ≥ Size 2 (vertical black line) at all elevation bands. We attribute this strong rule to the fact that avalanches smaller than Size 2 are not considered harmful to backcountry travelers by definition. The likelihood split between Possible-Likely and Likely also emerged as a consistent splitting rule at all elevation bands even though it was not as substantial as the Size 2 rule.

A comparison of the alpine danger rating assignments for storm slab avalanche problems (Fig. 3, right panel) and other avalanche problem types (Fig. 4) highlight both differences and similarities. The charts clearly highlight how the different avalanche problems occupy different areas within the avalanche hazard chart. In addition, we can see that while the main split for storm and wind slabs is related to size, the splits in the other avalanche problem types are more dominated by likelihood of avalanches. This might indicate a different focus in the assessment of these avalanche problems. Furthermore, the illustrations highlight that storm slab avalanche problems were consistently rated more serious than any of the other three problems with the exception of a few likelihood and size combinations (e.g., Unlikely-Possible and Size 1.5; Possible and Size 1.5). However, a comparison of the hazard charts for wind, persistent and deep persistent slabs indicate considerable consistency in the individual assessment of these three problems.

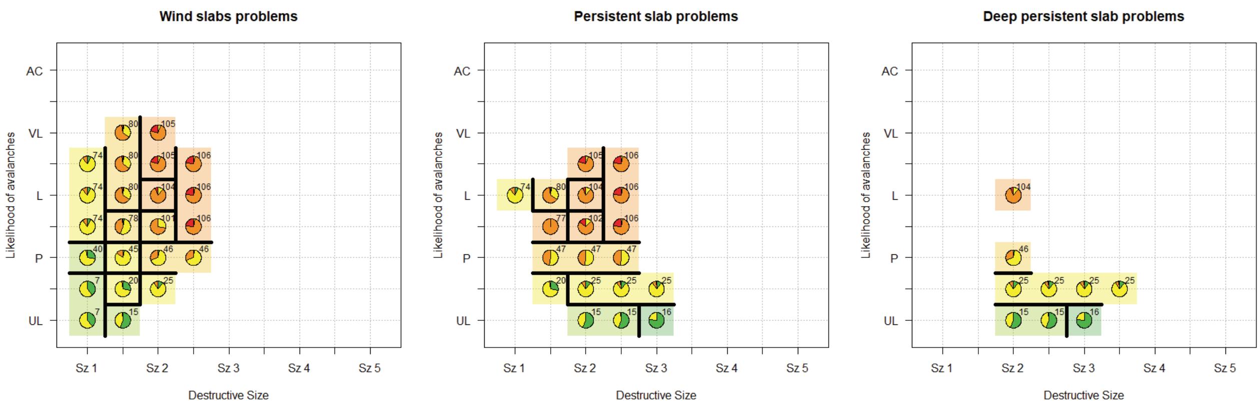

Fig. 4: Hazard assessment rules (with terminal note identifiers) and associated danger rating distributions for single wind slab, persistent slab and deep persistent slab avalanche problems for Avalanche Canada in alpine elevation band at likelihood and size combinations where observation exist. Background shading indicate danger rating predicted by CIT tree at intermediate steps.

4. CONCLUSION

To explore the missing link between the CMAH and avalanche danger ratings, we explored hazard assessments included in the operational bulletin datasets from Avalanche Canada and Parks Canada with conditional inference trees. Our results shed light on assessment rules of avalanche forecasters and the avalanche danger ratings they assign to single avalanche problem situations.

While we were only able to show illustrative examples of our results in this paper, our analysis revealed considerable variability in danger rating assignment. Single avalanche problems are assessed differently despite similar combinations of likelihood of avalanches and destructive size. In addition, our analysis showed that these ratings are different depending on elevation band and in some cases agency.

The next step in our analysis is to examine danger rating assessments including multiple avalanche problems to provide insight into the interplay between multiple avalanche problems and how assessment rules are affected by the presence of additional problem types. Together with the present analysis, this will provide a comprehensive picture of danger rating assignment practices of Canadian public avalanche forecasters.

While our analysis reveals considerable variability in danger rating assessments, it showcases that decision trees can provide valuable insights into how forecasters assign danger ratings in relation to their CMAH assessments. Despite the complexity of the extracted trees, they allow for a systematic examination of danger ratings associated with different combinations of contributing factors. This perspective offers a platform for avalanche forecasters to have an informed discussion about how and why danger rating assessments differ between problem types, elevation bands and agency. Together, the analysis and interpretation by forecasters will provide the necessary foundation for the development of meaningful rules allowing consistent danger rating assessments for public avalanche bulletins.

ACKNOWLEDGEMENT

We would like to thank Avalanche Canada and Parks Canada for the use of their bulletin datasets. Mitacs Accelerate, Avalanche Canada, the Avalanche Canada Foundation, HeliCat Canada, Canadian Pacific Railway and Simon Fraser University financially supported this research. A special thanks goes to Patrick Mair for his assistance with the statistical analysis.

REFERENCES

Bellaire, S., & Jamieson, B., 2013: On estimating avalanche danger from simulated snow profiles. *Proceedings of the 2013 International Snow Science Workshop, Grenoble, France*, 154–161.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., 1998: *Classification and regression trees.* CRC Press.

European Avalanche Warning Service., 2018: Bavarian matrix. Accessed 13 August 2018 [Available online at http://www.avalanches.org/eaws/en/main_layer.php?layer=basics&id=4]

Haegeli, P., Falk, M., & Klassen, K., 2012: Linking avalanche problems to avalanche danger - A first statistical examination of the conceptual model of avalanche hazard. *Proceedings of the 2012 International Snow Science Workshop, Anchorage, Alaska.*, 1–7.

Hendrikx, J., Murphy, M., & Onslow, T., 2014: Classification trees as a tool for operational avalanche forecasting on the Seward Highway, Alaska. *Cold Regions Science and Technology*, *97*, 113–120. https://doi.org/10.1016/j.coldregions.2013.08.009

Hothorn, T., Hornik, K., Wiel, M. A. van de, & Zeileis, A., 2008: Implementing a Class of Permutation Tests: The **coin** Package. *Journal of Statistical Software*, *28*(8). https://doi.org/10.18637/jss.v028.i08

Hothorn, T., Hornik, K., & Zeileis, A., 2006: Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651–674. https://doi.org/10.1198/106186006X133933

Kuhn, M., and Johnson, K., 2013: *Applied Predictive Modeling.* Springer, 369 - 411.

Lazar, B., Trautman, S., Cooperstein, M., Greene, E., & Birkeland, K., 2016: North American Avalanche Danger Scale: Do Backcountry Forecasters Apply It Consistently?, *Proceedings of the 2016 International Snow Science Workshop, Breckenridge, Colorado,* 457–465.

Martignon, L., Katsikopoulos, K. V., and Woike, J. K., 2012: Naïve, Fast, and Frugal Trees for Classification. P. M. Todd, G. Gigerenzer, and A. B. C. R. Group, Eds., Oxford University Press, 360-378.

Mileti, D.S., and Sorensen, J.H., 1990: Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. ORNL-6609, 161 pp.

Murphy, A.H., 1993: What is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, 8, 281 - 293

Schirmer, M., Schweizer, J., & Lehning, M., 2010: Statistical evaluation of local to regional snowpack stability using simulated snow-cover data. *Cold Regions Science and Technology*, *64*(2), 110–118. https://doi.org/10.1016/j.coldregions.2010.04.012

Statham, G., Haegeli, P., Birkeland, K. W., Greene, E., Israelson, C., Tremper, B., … Kelly, J., 2010: The North American Public Avalanche Danger Scale. *Proceedings of the 2010 International Snow Science Workshop*, *Squaw Valley, California.*, 80–87.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., … Kelly, J., 2018: A conceptual model of avalanche hazard. *Natural Hazards*, *90*(2), 663–691. https://doi.org/10.1007/s11069-017-3070-5

Techel, F., Ceaglio, E., Coléou, C., Mitterer, C., Morin, S., Purves, S., & Rastelli, F., in review: Spatial consistency and bias in avalanche forecasts - a case study in the European Alps. *Natural Hazards and Earth System Science Discussion*, https://doi.org/10.5194/nhess-2018-74,