CONSISTENCY AND ACCURACY OF PUBLIC AVALANCHE FORECASTS IN WESTERN CANADA

Grant Statham[1, 2]*, Stephen Holeczi[1] and Bret Shandro[2]

[1]Parks Canada Agency, Banff, AB, Canada
[2]Simon Fraser University, Burnaby, BC, Canada

ABSTRACT: Forecasting avalanche problems and avalanche danger is a judgmental assessment that is highly susceptible to interpretation and bias. Thus, congruency between forecasters is a significant challenge and much effort has been expended to harmonize assessment methods between different forecasters, regions, and even countries. While recent studies have helped to identify bias and inconsistency in avalanche danger ratings (Lazar et al. 2016; Techel et al. 2018), in-house feedback directly to forecasters is sometimes absent. Accurate, well-summarized feedback could provide the primary basis for avalanche forecasters to improve their judgmental forecasts. Using data from historic forecasts, this paper looks first at the consistency of how avalanche problems are applied by different agencies with adjacent regions in the Canadian Rockies. We show inconsistency in the distribution of avalanche problems published for adjacent regions with similar snowpacks, and by different forecasters within the same region. Although definitions exist for different types of avalanche problems (Statham et al. 2018), insufficient guidance for forecasters on how to apply avalanche problems consistently can lead to conflicting information and confusion for backcountry users. Next we look at the accuracy of 24, 48 and 72 hour forecasts of danger ratings when compared against real-time assessments. Drawing from 3,752 avalanche bulletins over seven seasons, we show an overall accuracy of 73%. Forecasts of *Low* danger are the most accurate (84%) and they become progressively less accurate as the forecast danger levels rise. We conclude by offering recommendations on the application of avalanche problems, enhanced forecaster training and encouragement for other agencies to analyze their own forecasting data. Feedback can have a pronounced effect on bias if incorporated more routinely into professional activities (Vick 2002) and with it, forecasters can become better calibrated.

KEYWORDS: avalanche problems, danger, forecasting, consistency, accuracy

1. INTRODUCTION

Public avalanche bulletins are published once per day for many regions in western Canada. While each region is distinct from the others, in reality there is overlap of avalanche conditions between adjacent neighbors. This overlap extends also to the users of the information who travel between the regions and use different avalanche bulletins produced by different agencies. Consistency between forecasters is important to ensure a coherent public message across boundaries.

Canadian avalanche bulletins are usually valid for 24-hours and are comprised of four distinct parts: 1) headline, 2) danger ratings, 3) avalanche problems, and 4) forecast details. Our objective was to examine the consistency between adjacent regions in how avalanche problems are presented, and to determine the accuracy of avalanche danger predictions.

* *Corresponding author address:*
Grant Statham, Parks Canada Agency
Box 900, Banff, Alberta, Canada, T1L 1K2
T: +01 403 760 0510, E: grant.statham@pc.gc.ca

1.1 Forecasting avalanche problems

In Canada, forecasters choose from nine different types of avalanche problems (Statham et al. 2018). Public forecasts display between zero and three avalanche problems, prioritized so the most consequential problems are presented first (i.e.: Problem 1, 2 and 3). Some types of problems are short-term in nature (e.g.: storm slab), while others change slowly and linger for long periods of time (e.g.: persistent slab).

1.2 Forecasting avalanche danger

The Avalanche Danger Scale is an ordinal, five-level rating system used internationally to rank the level of avalanche danger. Forecasters compare their assessment of the avalanche danger to a set of definitions to determine the rating. In Canada, danger ratings are forecast for 24, 48 and 72 hours ahead for three different vegetation bands (alpine, treeline and below treeline).

2. DATA AND METHODS

2.1 Forecast study regions

For our avalanche problem analysis we chose three regions: Jasper National Park (JNP), Banff, Yoho and Kootenay National Parks (BYK) and Kananaskis Provincial Park (KPP). These regions share borders over a 400km stretch of the Rocky Mountains where the continental snowpack conditions are generally analogous between the regions.

For our avalanche danger analysis, we used 3,752 public avalanche bulletins from five Parks Canada regions (BYK, Little Yoho, JNP, Glacier National Park and Waterton Lakes National Park) from 2011/12 to 2017/18. These bulletins contained 27,550 forecasts of avalanche danger.

2.2 February and March 2017

Our avalanche problem investigation spans a two month period from February 1 to March 31, 2017 where we analyzed 59 days of avalanche bulletins for three regions (JNP, BYK, KPP) to determine what avalanche problems were listed, in what order and by what forecaster. Our goal was to make comparisons between regions and forecasters. During this period, the lower snowpack was consistently weak (facets, depth hoar) across all three regions with several storms producing short-term problems in the upper half of the snowpack.

2.3 Forecast compared to nowcast

True verification of avalanche danger is not possible because avalanche danger cannot be measured, only estimated. In Canada, all avalanche forecasts contain a real-time assessment of avalanche danger (nowcast), assumed to be the most accurate assessment of the danger available. For this analysis we ranked danger levels from 1-5 and compared the forecast danger ratings to the subsequent nowcast ratings to determine their accuracy. Positive differences indicate forecasts where the predicted danger rating was too high, while negative differences represent forecasts that were too low.

3. ANALYSIS AND RESULTS

3.1 Number of avalanche problems per bulletin

Despite the proximity of these regions to each other and their similar avalanche conditions, there is substantial variation in the number of problems published per bulletin for each region (Figure 1).

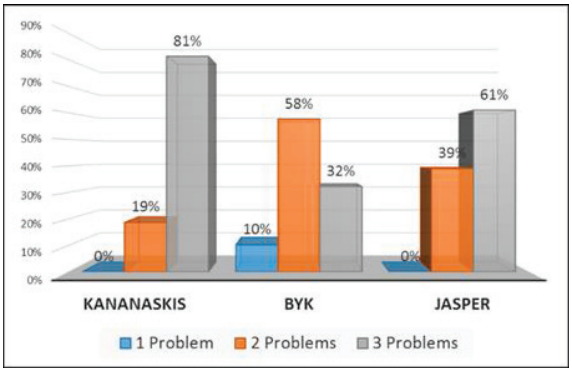

Figure 1: Proportion of forecasts showing one, two or three avalanche problems.

One reason for these inconsistencies is that often, two avalanche problems were listed that described the same condition. This redundancy occurred most often with *Persistent Slabs* and *Deep Persistent Slabs* (Figure 2), but also *Wind Slabs* and *Storm Slabs* were often listed together with little distinction given between them.

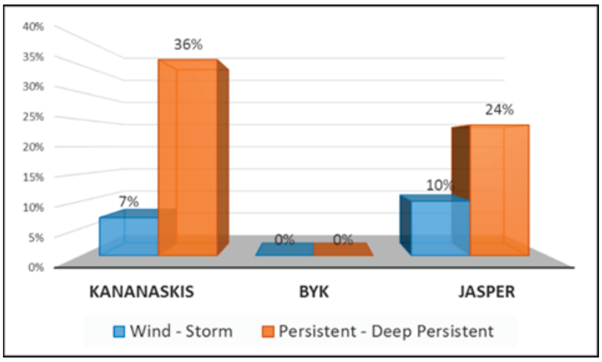

Figure 2: Proportion of forecasts with *Wind/Storm Slabs* and *Persistent/Deep Persistent Slabs* shown concurrently.

3.2 Prioritizing and switching avalanche problems

Prioritizing avalanche problems is crucial because the order in which they are listed warns of the most consequential avalanche issues. Switching problems is significant because it alerts people that important changes have occurred in the conditions. For neighboring regions with similar conditions, changes should be somewhat in-sync and happening at about the same time. Our analysis showed variation in the frequency that each region switched Problem 1 (JNP 27% of the time, KPP 19%, BYK 10%) and in many cases we were unable to correlate the change with anything other than a different forecaster. For example, over the two month period the priority of the *Deep Persistent Slab* problem was inconsistent between all three regions (Figure 3).

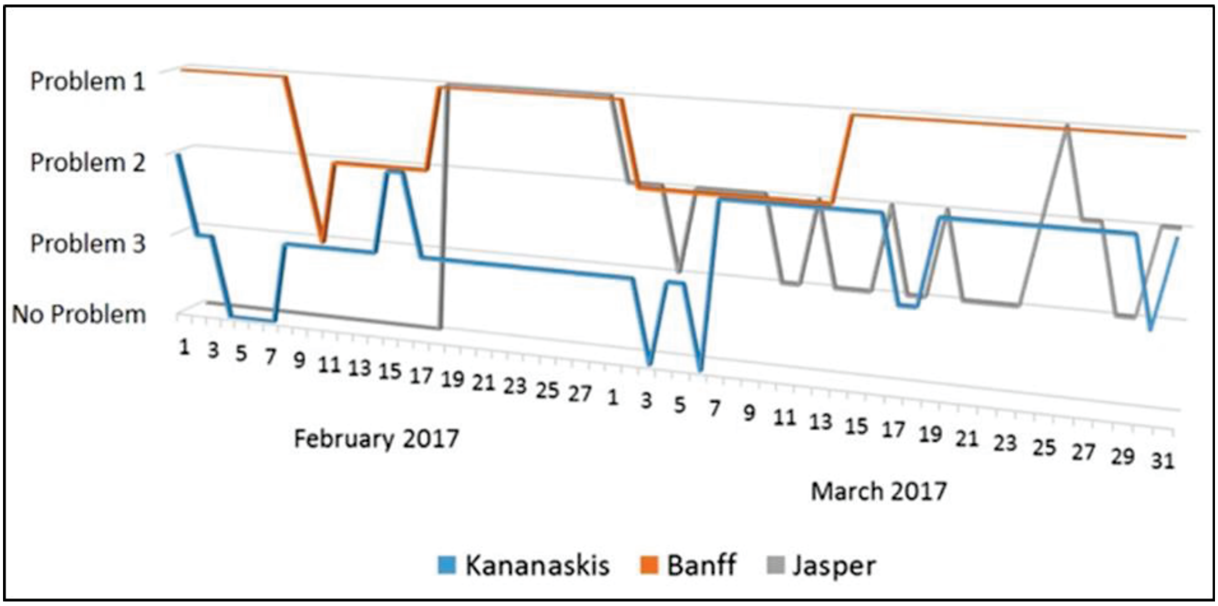

Figure 3: Prioritizing the 2017 *Deep Persistent Slab* problem. From March 14-21, a climax avalanche cycle occurred where this problem was producing widespread 1:30-year avalanches.

3.3 Regional danger rating accuracy

Regional danger ratings include all of Parks Canada's five regions. 24-hour forecasts are shown to be the most accurate at 81% with accuracy then deteriorating as time scales become longer. 73% of all danger rating forecasts were correct (Table 1).

Table 1: Number of forecasts that matched the nowcast for all Parks Canada regions.

Forecast	*Total Forecasts*	*Correct Forecasts*	*Accuracy*
24 Hour	9,932	8,017	81%
48 Hour	9,975	7,113	71%
72 Hour	7,643	5,073	66%
Overall	27,550	20,203	73%

Of the remaining 27% of errors, the majority of them occurred at the mid to upper-end of danger levels by over-forecasting one-step (Figure 4). Interestingly, the highest accuracy occurs with forecasts of *Low* (84%) and then deteriorates progressively, in-order to *Extreme* (11%). Lower danger appears to correlate with lower uncertainty for forecasters.

Overall		Under Forecast ←					Over Forecast →			
Danger Rating	**n**	**-4**	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	**+4**
Low	7681	0%	0%	1%	15%	84%				
Moderate	10306		0%	1%	14%	71%	14%			
Considerable	8190			0%	7%	69%	22%	1%		
High	1364				3%	55%	39%	3%	0%	
Extreme	9					11%	89%	0%	0%	0%

Figure 4: Accuracy of danger rating forecasts for 24, 48 and 72 hours for all Parks Canada regions (n=27,550).

3.4 <u>BYK 24-hour danger rating accuracy</u>

In BYK, the overall 24-hr danger rating forecasts were accurate 84% of the time. For specific danger levels, accuracy ranged from 93% at *Low* to 61% at *High*, while *Extreme* was never issued in a 24-hour forecast (Figure 5).

BYK: 24 hr Forecasts		Under Forecast					Over Forecast			
Danger Rating	n	-4	-3	-2	-1	0	+1	+2	+3	+4
Low	830	0%	0%	0%	7%	93%				
Moderate	993		0%	0%	7%	84%	9%			
Considerable	781			0%	4%	79%	16%	1%		
High	170				6%	61%	31%	2%	0%	
Extreme	0					0%	0%	0%	0%	0%

Figure 5: Accuracy of danger rating forecasts for 24-hours for the Banff, Yoho and Kootenay region (n=2,774).

3.5 <u>Individual forecaster accuracy</u>

All avalanche forecasters are interested in the accuracy of their own predictions. The individual forecaster accuracies for 24 and 48-hour forecasts in BYK ranged from 91% to 69%, and the 24-hour forecast is shown to be the most accurate.

Table 2: BYK individual forecaster accuracy for 24 and 48-hour danger rating forecasts.

Forecaster	*Forecasts*	*24h accuracy*	*48h accuracy*
TH	342	91%	81%
IJ	755	87%	80%
SS	1968	87%	75%
CJ	609	81%	72%
AB	990	83%	72%
SH	771	82%	74%
LP	570	82%	71%
BW	651	80%	69%
GS	792	77%	69%

4. RECOMMENDATIONS

Our analysis illustrates patterns in the way avalanche problems were published and the accuracy of avalanche danger forecasts. The following recommendations are made to improve the consistency and accuracy of avalanche forecasts.

4.1 <u>Refer to definitions</u>

When determining avalanche danger ratings, forecasters refer to the avalanche danger definitions (Statham et al. 2010). However, this has not been the case for avalanche problems, where forecasters often make assumptions about what constitutes a particular type of avalanche problem. When assessing avalanche problems, forecasters are encouraged to reference the avalanche problem definitions described in detail by Statham et al. (2018) in the same way they reference danger rating definitions.

4.2 <u>Minimize the number of problems and overlap</u>

When choosing avalanche problems, forecasters should start with the previous assessment and ask themselves what has changed, and what message do they want to communicate? Too many avalanche problems makes public interpretation difficult and defeats the purpose of giving only the most salient information – rather than all of the information. It is easy to confuse someone by publishing two problems that say almost the same thing. Most often, two problems can communicate the situation (deep snowpack and upper snowpack). Avalanche problems should help people to understand the conditions *so they can make effective terrain and risk management choices.*

4.3 <u>Prioritize problems and minimize switching</u>

The priority of an avalanche problem ranks its importance in terms of risk management and terrain choices. Switching the priority of avalanche problems should not be taken lightly, and ought to be based on the conditions, rather than a different forecaster’s perspective. Frequent switching of problems is confusing and shows a need for more communication within forecasting teams. The avalanche bulletin should present a unified approach that represents the conditions in the region.

4.4 <u>Consensus-based changes and change rules</u>

Non-persistent avalanche problems such as *Dry Loose*, *Storm Slabs* or *Wind Slabs* are dynamic in nature; they come and go as weather systems affect the surface of the snowpack. Persistent problems are the opposite and once established, usually exist for weeks or months. The decision to add a persistent problem to the forecast is important, as it will likely remain there for an extended period. When possible, forecasters should seek consensus from within their team before making major changes such as this to the structure of their forecast's avalanche problems. In-house rules can assist with this by setting guidelines for how and when to add or remove persistent problems, or when to transition a *Storm Slab* into a *Wind Slab*.

4.5 Forecaster feedback and training

Subjective, judgmental assessments are difficult, thus mitigating bias and striving for consistency requires close coordination. One of the best ways for forecasters to improve their accuracy and consistency is through direct feedback, which can have a pronounced effect on bias if incorporated routinely into professional activities (Vick 2002). Past avalanche forecasts offer fertile ground for this kind of evaluation, and the topics presented here could form a baseline for multi-agency forecaster training.

5. CONCLUSION

Our analysis shows inconsistency in the application of avalanche problems between JNP, BYK and KPP. While weather and avalanche conditions are not identical between these regions, they are similar enough that these differences cannot be solely attributed to avalanche conditions. We observed a link between different public messages and different forecasters. Avalanche problems were switched, added, removed then added back again over short periods of time. These changes can be attributed to forecasters coming on and off shift.

We also found that combined, all of the forecasts of avalanche danger were 73% accurate when compared to the subsequent nowcast. Our analysis reaffirmed that forecasting accuracy is dependent upon temporal scale, with predictions having increasing uncertainty farther into the future and at higher levels of danger. Interestingly, over seven years and 27,550 forecasts, danger level *Extreme* was only forecast 0.03% of the time and was incorrect 89% of those times.

In summary, we believe there are important lessons to be learned from this data. First and foremost, public avalanche forecasting is risk communication, where the information must be explicit, coherent and brief (Young and Lovvoll 1999). Just because a particular condition exists does not mean it's important enough to publish. Forecasters must filter out extraneous information and publish only the most salient messages. In addition, teamwork and internal communication are critical for producing a consistent forecast. The bias of individual forecasters can be tempered through reference to definitions and consensus-based decisions. Finally, forecaster training that includes direct feedback from previous forecasts is of immense value and interest, and helps everyone to become better calibrated.

6. REFERENCES

Lazar, B., Trautman, S., Cooperstein, M., Greene, E., Birkeland, K., 2016. North American Avalanche Danger Scale: Are Public Backcountry Forecasters Applying it Consistently? Proceedings of the 2016 International Snow Science Workshop. Breckenridge, CO, USA.

Statham, G., Haegeli, P., Birkeland, K., Greene, E., Israleson, C., Tremper, B., Stethem, C., McMahon, B., White, B., Kelly, J., 2010. The North American Public Avalanche Danger Scale. Proceedings of the 2010 International Snow Science Workshop, Lake Tahoe, CA, USA.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israleson, C., Tremper, B., Stethem, C., McMahon, B., White, B., Kelly, J., 2018. A conceptual model of avalanche hazard. Natural Hazards, Vol. 90, Issue 2, pp 663–691.

Techel, F., Ceaglio, E., Coléou, C., Mitterer, C., Morin, S., Purves, R., Rastelli, F., 2018 (in review). Spatial consistency and bias in avalanche forecasts – a case study in the European Alps. Nat. Hazards Earth Syst. Sci. https://doi.org/10.5194/nhess-2018-74.

Vick, S., 2002. Degrees of belief—subjective probability and engineering judgment. American Society of Civil Engineers, Reston, VA.

Young and Lovvoll (1999). Intermediate Processing Stages: Methodological Considerations for Research on Warnings. In M. Wogalter, D. DeJoy & K. Laugherty (Eds.), *Warnings and Risk Communication* (pp 27-52). Philadelphia, PA: Taylor & Francis.