Cold Regions Science and Technology 30 (1999) 115–123

cold regions science and technology

www.elsevier.com/locate/coldregions

The stuffblock snow stability test: comparability with the rutschblock, usefulness in different snow climates, and repeatability between observers

Karl W. Birkeland [a,b,*], Ron F. Johnson [b,1]

[a] Department of Earth Sciences, Montana State University, Bozeman, MT 59717, USA
[b] Gallatin National Forest Avalanche Center, P.O. Box 130, Bozeman, MT 59771, USA

Received 24 August 1998; accepted 29 March 1999

Abstract

The stuffblock is a new snow stability test developed by the Gallatin National Forest Avalanche Center and used operationally since 1993. The test involves stressing an isolated column of snow 0.30 m^2 by dropping a nylon sack filled with 4.5 kg onto the column from 0.10 m increments until weak layer failure occurs. Results over several winters correlate the stuffblock test with the more widely used rutschblock test, and validate the usefulness of the test for evaluating snow stability in several different climates. Further, the test provides results that are consistent between observers, a favorable attribute for regional avalanche forecasting operations which use numerous observers. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Avalanches; Avalanche forecasting; Snow avalanches; Snow stability

1. Introduction

The stuffblock snow stability test involves stressing an isolated column in a snow pit by dropping a nylon stuff sack with a known mass onto the column from various heights until the weak layer fails. Extensive testing of the stuffblock since its development in 1993 indicates that: (1) stuffblock results can be statistically correlated to the widely used rutschblock test; (2) though developed in the intermountain snow climate of Montana, USA, the stuffblock test is also applicable in Washington's coastal and Colorado's continental snow climate; and (3) stuffblock results are repeatable between different observers. This latter result is especially important since it demonstrates the usefulness of the stuffblock in backcountry avalanche forecasting operations that rely on observer networks and in scientific research that utilizes a number of observers to measure spatial changes in snowpack properties (e.g., Birkeland, 1997).

Testing snow stability on a given slope is difficult. There are numerous tests, and little guidance about employing and interpreting some of them. Due to the subjective nature of several tests, comparisons between various observers are often difficult. Further, locating a ''representative'' site for the test is

* Corresponding author. Fax: +1-406-587-6804; E-mail: kbirkela/r1_gallatin@fs.fed.us
1 Fax: +1-406-587-6804; E-mail: rjohnson/r1_gallatin@fs.fed.us.

PII: S0165-232X(99)00015-4

complex, as emphasized by recent field studies which have documented spatial variations in snowpack properties (Conway and Abrahamson, 1984; Föhn, 1988; Logan, 1992; Jamieson and Johnston, 1993; Birkeland et al., 1995). In spite of the difficulties, stability tests are critical tools for avalanche workers attempting to evaluate the stability of a particular slope (LaChapelle, 1980; McClung and Schaerer, 1993), and for scientists attempting to test various aspects of the snowpack.

All currently available stability tests have shortcomings. The shear frame is the most quantifiable test, and is used to measure the shear strength of weak layers which are responsible for avalanche release. Combining results with calculations of stresses applied by the snow above the weak layer, and additional loads applied by skiers or others, allows the derivation a variety of stability indices (i.e., Föhn, 1987a; Jamieson and Johnston, 1994; Jamieson, 1995). Some stability indices are being used for operational avalanche forecasting in Canada, but shear frames are predominantly used only in scientific work because they are time consuming and difficult to use (Perla and Beck, 1983).

Tests more commonly used by avalanche workers and backcountry skiers also have their drawbacks. Simple shovel shear tests have been used widely, probably because they are fast, easy, and only require an avalanche shovel (although many people prefer to also use a snow saw). Though effective at identifying the location of weak interfaces, shovel shear results are not easily communicable between various observers (a ‘‘moderate’’ shovel shear often means different things to different observers), and it may take one person several tests to reliably rate the shear strength (Schaerer, 1988). A slightly more time consuming, but still relatively quick test is the ‘‘loaded column’’ test whereby blocks of snow are placed on top of an isolated column until the column fails (McClung and Schaerer, 1993). An advantage of this test is a better ability to communicate results (i.e., ‘‘the column failed when loaded with 0.25 m of old snow with a density of 300 kg/m^3’’). Still, block size may not be uniform and cutting reasonable blocks out of the snow when it is cohesionless (such as with new or faceted snow) is difficult. The compression test is another quick stability test that is widely used in North America. In this test, a 0.30 m by 0.30 m column is isolated, a shovel is put on top of the column, and the column is stressed by alternately tapping on the shovel five to 10 times with a motion beginning at the wrist, the elbow, and then shoulder (Jamieson and Johnston, 1996). Though results are more easily compared than the shovel shear test, ample room for error exists between observers who might apply different forces to the column.

Increasingly popular among North American researchers and backcountry skiers is the Swiss rutschblock test (Föhn, 1987b). This test involves completely isolating a column 2 m long and 1.5 m wide which is then progressively loaded by a skier until the block fails. Several studies have utilized rutschblocks (Föhn, 1988; Jamieson and Johnston, 1993; Birkeland, 1997), and work indicates that rutschblocks can be roughly correlated to slope stability (Jamieson and Johnston, 1992). Still, rutschblock results depend on how well the large block is isolated, the weight of the person jumping on the block, and how hard they jump. Results, which are given a value between 1 and 7 on scale of increasing difficulty to failure, are easier to compare than the ‘‘easy, moderate, or hard’’ values given to shovel shears, but they are still biased.

2. Conducting the stuffblock test

To address concerns about various stability tests we developed the stuffblock test during the 1992–1993 winter, and have used it since then as part of our avalanche forecasting operation in the mountains around Bozeman, Montana. Methods are more exhaustively described by Birkeland et al. (1996). Equipment required for the test includes a shovel, a snow saw, a small spring scale, and a nylon stuff sack with a string hanging off the bottom of it marked in 0.10 m increments. The stuffblock test is conducted as follows: (1) a vertical column of snow 0.30 m wide by 0.30 m deep is completely isolated from the snow pit wall using a snow saw; (2) a nylon stuff sack is packed firmly with snow until it weights 4.5 kg (measured by the spring scale) and the top of the sack is tied off; (3) the shovel is put on the top of the isolated column and the stuff sack is gently placed on the shovel blade (failure of the weak layer

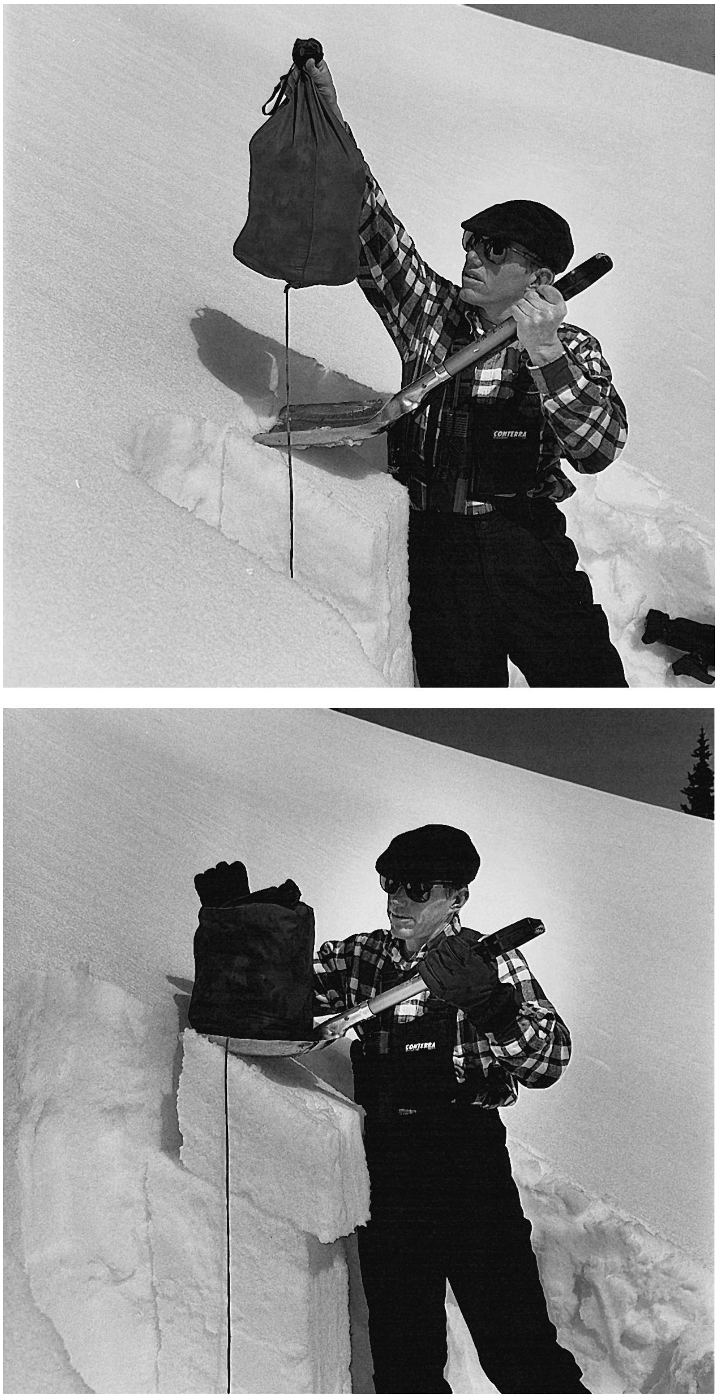

Fig. 1. The stuffblock snow stability test consists of dropping a nylon sack filled with 4.5 kg of snow from known heights onto a 0.30-m^2 isolated column until weak layer failure.

Table 1
Qualitative ratings of shear quality used with the stuffblock test (and subsequently used with compression and rutschblock tests)

Shear quality	Description
Q1	Unusually clean and smooth shear plane, weak layer may noticeably collapse during failure. Slab typically slides easily into the snow pit after weak layer failure on slopes steeper than 35°, and sometimes on slopes as gentle as 25°.
Q2	''Average'' shear, shear plane appears mostly smooth, but slab does not slide as readily as Q1. Shear plane may have some small irregularities, but not as irregular as Q3. Shear failure occurs through the whole block being tested, and slab may or may not slide into snowpit.
Q3	Shear plane is uneven, irregular or rough. Shear failure may not occur through the whole block being tested. After weak layer failure, slab moves little, or may not move at all, even on slopes steeper than 35°.

at this point indicates a stuffblock drop height of zero); and (4) the column is loaded dynamically by dropping the stuff sack from 0.10 m, and increasing that height by 0.10 m increments (measured with the string tied under the stuff sack) until shear failure in the weak layer occurs (Fig. 1). If more than one weak layer is present in the snowpack, we remove the first layer after it fails and continue dropping the stuff sack from increasing heights until the next layer of interest fails. In addition to noting the drop height necessary to get weak layer failure, over the past two seasons we have begun recording shear quality on a qualitative scale from one (Q1; clean, smooth shear) to three (Q3; irregular or uneven shear) (Table 1). We feel shear quality gives us additional information about the relationship between snowpack layers, and may provide important clues about slope stability. We have also begun assessing shear quality for other stability tests such as the rutschblock and compression tests.

3. Correlation between stuffblock and rutschblock tests

During the 1993–1994 winter, we conducted 54 stuffblock and rutschblock tests adjacent to each other to control for variations in aspect, elevation and slope angle, although small-scale variability in snow strength may still have been present (Johnson and Birkeland, 1994). Comparisons were performed on a variety of avalanche starting zones, and an analysis of the snow profile insured that both tests failed on the same weak layer. We considered only dry slab conditions. Test slopes faced all aspects at elevations ranging from about 2300 to 3200 m, slope angles varying from 28° to 38° (average of 33°), and slab depths ranging from about 0.10 to 1.0 m (average of about 0.40 m). Weak layers tested included new snow interfaces, surface hoar, depth hoar, and near-surface faceted crystals (Birkeland, 1998). We used the nonparametric Spearman rank order correlation (Zar, 1984) to statistically compare stuffblock drop heights and rutschblock numbers.

The data clearly indicate a strong, positive correlation between stuffblock and rutschblock results (Spearman rank order correlation = 0.77; $p < 0.0001$). Since rutschblock results can be roughly correlated to slope stability (Jamieson and Johnston, 1992), this relationship implies that stuffblock results are similarly useful, an implication backed up by our observations while avalanche forecasting. It is no surprise that the two tests are not exactly comparable. The stuffblock scale translates to a linear increase in the impact energy imparted to the snow column (Johnson and Birkeland, 1994) while recent research suggests that each rutschblock step between two and five approximately doubles the force applied to the weak layer (Camponovo and Schweizer, 1996). Another difference is the way stresses are transmitted to the weak layer since with a rutschblock the forces applied by the skier testing the block can be approximated by a line load (Föhn, 1987a) while the stuffblock imparts force more evenly on the isolated column.

4. Comparisons in different snow climates

Snowpack characteristics in the western United States are classified into three predominant snow climates: coastal, intermountain, and continental

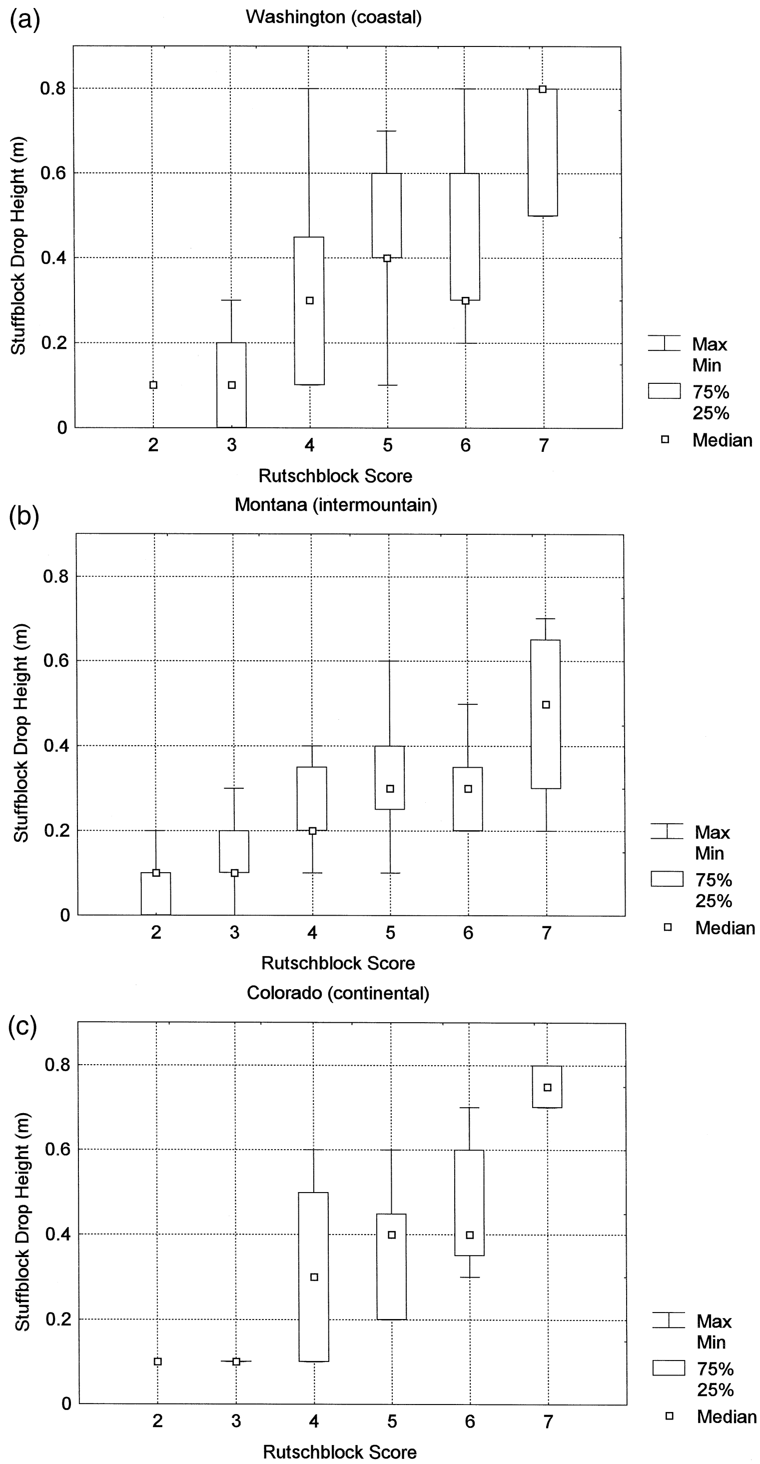

Fig. 2. Box–whisker plots for data comparing stuffblock drop heights to rutschblock scores for (a) Washington's coastal, (b) Montana's intermountain, and (c) Colorado's continental snow climate. The number of observations in each climate is listed in Table 2.

(LaChapelle, 1966; Armstrong and Armstrong, 1987; Mock, 1995; Mock and Kay, 1992). Coastal snow climates are characterized by generally warmer temperatures, more snowfall, higher snow density, and less faceted snow crystal growth than areas farther inland. Continental snow climates have colder temperatures, less snowfall, and more faceted crystal growth, while intermountain snow climates are intermediate between these two extremes. During the 1995–1996 winter, experienced avalanche workers tested the stuffblock in Washington's coastal snow climate, Montana's intermountain snow climate, and Colorado's continental snow climate (Birkeland et al., 1996). Investigators collected data on slopes representative of avalanche starting zones in their particular region.

Avalanche workers in all three snow climates agreed that, qualitatively, the stuffblock provided effective snow stability information. To further validate the stuffblock's usefulness, we investigated the relationship between the stuffblock and rutschblock tests in each snow climate using the same methods used previously in Montana. A comparison of the median, upper and lower quartile, and range for stuffblock results associated with each rutschblock step shows that the relationship between the two tests is roughly similar in all three snow climates (Fig. 2). Further, in all cases, the Spearman rank order correlation was highly significant and positive (Table 2), indicating that increasing stuffblock drop heights are correlated with increasing rutschblock scores, and further validating the usefulness of the stuffblock test for assessing snow stability.

Combining results for all three snow climates allows a closer examination of the relationship between rutschblock and stuffblock results (Table 3, Fig. 3). Föhn (1987b) concludes that when rutschblocks are between 1 and 3 the snowpack is mostly unstable. Median stuffblock results associated with rutschblock scores of 2 and 3 are 0.10 m, with upper and lower quartiles of 0 to 0.10 m for rutschblocks of 2 and 0.10 to 0.20 m for rutschblocks of 3. Our observations in Montana indicate that these stuffblock results are typically associated with mostly unstable snowpack conditions, in agreement with the rutschblock conclusions of Föhn (1987b). More intermediate stability conditions are represented by rutschblock scores of 4 or 5, which are associated with median stuffblock drop heights of 0.30 and 0.40 m, respectively. Rutschblock values of 6 and 7 show mostly stable conditions (Föhn, 1987b) and increasing stuffblock drop heights are associated with those values. Finally, we emphasize the cautionary notes of Föhn (1987b) about stability tests such as the rutschblock and stuffblock: (1) there are substantial variations in weak layer strength over a given slope; (2) slope stability is determined by the number and linkages between these weaker areas, and these may not be located by a small number these tests; and (3) effective snow stability evaluation requires experienced observers with a wide variety of available data.

Table 2
Spearman rank order correlation coefficients comparing adjacent stuffblock and rutschblock results from data collected during the 1995–1996 winter in Washington, Montana, and Colorado. N is the number of adjacent tests and p is the probability that the relationship is due to chance

Area	Snow climate	Spearman r	p	N
Washington	Coastal	0.72	0.0000	57
Montana	Intermountain	0.71	0.0000	64
Colorado	Continental	0.69	0.0001	27
All sites	–	0.73	0.0000	148

Table 3
Stuffblock drop heights associated with rutschblock scores for all data from the winter of 1995–1996. The number of times a given rutschblock score was observed is represented by N

Rutschblock score	Stuffblock drop height (m)			
	Median	Lower quartile	Upper quartile	N
2	0.10	0	0.10	13
3	0.10	0.10	0.20	27
4	0.30	0.20	0.40	44
5	0.40	0.30	0.50	28
6	0.40	0.30	0.60	20
7	0.80	0.60	0.80	16

5. Repeatability between observers

Our primary goal in stuffblock development was to increase the consistency of stability test results

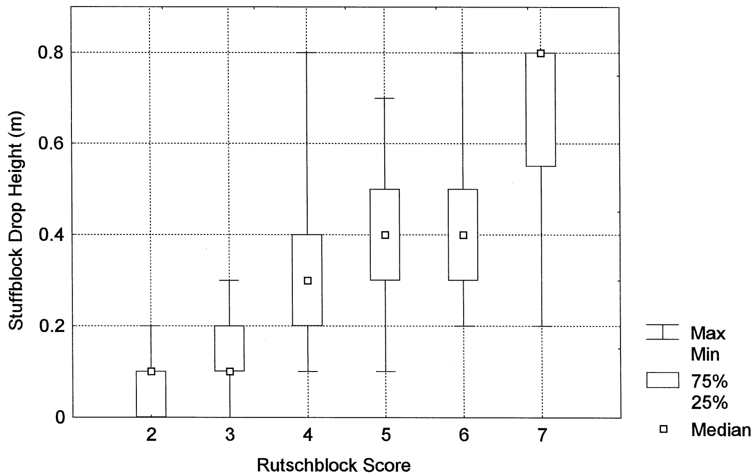

Fig. 3. Box–whisker plot comparing stuffblock drop heights to rutschblock scores for all data from Washington, Montana, and Colorado ($N = 148$).

between different observers. To determine whether or not results were repeatable, four different observers conducted stuffblock tests, and we analyzed results using the non-parametric Wilcoxson matched pairs test. The study utilized a fairly uniform, north facing, 35° slope at 2315 m in elevation. Excavating a series of wide snowpits allowed us to conduct 15 different sets of four adjacent stuffblock tests. Two distinctive failure planes existed in the snowpack, and the shear quality for both was rated Q2. The first was at 0.08 m below the surface and consisted of a density change between snow that had fallen the day before our test and the older snow. The second shear was at a depth of 0.45 m and consisted of a weak interface between snow that had fallen over the past 6 days and a hard, wind affected layer beneath that snow.

Data demonstrate the repeatability of the stuffblock between different observers. No variation between observers or between tests existed at 0.08 m,

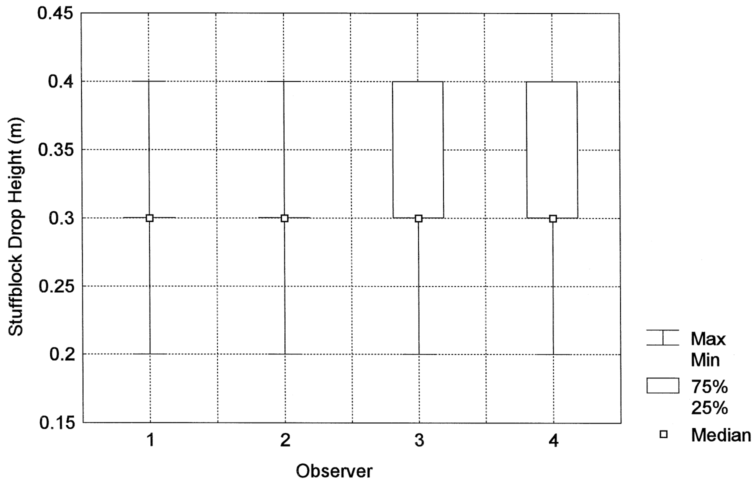

Fig. 4. Box–whisker plot for four observers conducting 15 side-by-side stuffblock tests showing the range, upper and lower quartile, and median for each observer. Weak layer failure occurred 0.45 m below the snow surface.

Table 4
p-values for differences in stuffblock results between observers at a failure depth of 0.45 m computed using a Wilcoxson matched pairs test. There are no significant differences ($p < 0.05$) between any observer pairs

Observer	1	2	3	4
1	–	0.59	0.94	0.96
2	0.59	–	0.58	0.69
3	0.94	0.58	–	0.86
4	0.96	0.69	0.86	–

with every one of the 15 tests for each of the four observers having a stuffblock drop height of zero. Although more variability was evident at 0.45 m, results were still remarkably uniform, especially considering the small scale variations in stability test results that can occur over a given slope. Drop heights ranged from 0.20 to 0.40 m, and the median was 0.30 for each observer (Fig. 4). A Wilcoxson matched pairs test showed no significant difference between any observer pair at $p < 0.05$ (Table 4). This result matches well with our experience with a number of different observers which indicates that the stuffblock test provides remarkably consistent results.

Consistency is the main advantages of the stuffblock test over other available stability tests. Reliable and consistent results are important for regional avalanche forecasting, which typically relies on a large number of observers who may have varying avalanche skills. In addition, consistency is critical for avalanche research which utilizes several observers to analyze spatial variations in snow stability, and the stuffblock works well for these projects (Birkeland, 1997).

6. Summary

The stuffblock snow stability test provides valuable and quantifiable information about the strength and location of snowpack weak layers. Stuffblock results are statistically correlated to adjacent rutschblock results, and the stuffblock test works well in coastal, intermountain, and continental snow climates. Further, results can be readily compared between observers. Such comparability is especially useful for regional avalanche forecasters who must compare the results of several different observers with differing avalanche skills, and for scientists that rely on several observers when making spatial measurements of snow stability. Though an improvement over many previous snow stability tests, the stuffblock is not perfect. A particular shortcoming is that stuffblocks test a much smaller area of the slope than rutschblocks (approximately 0.09 m^2 for the stuffblock vs. 3.0 m^2 for the rutschblock), and therefore more stuffblock tests are required to reliably estimate slope stability. Still, stuffblock data are useful in combination with other data when assessing snow avalanche potential.

Acknowledgements

We thank the US Forest Service National Avalanche Center, the US Forest Service Missoula Technology and Development Center, and the National Science Foundation (Dissertation Improvement Grant #SBR-962830) for providing financial support for this research. B. Doubek, A. Gleason, A. Horwitz, N. Logan, G. Murphy, D.S. Schmidt, G. Urie, and K. Williams helped collect data. The Gallatin National Forest allowed us to experiment with this test while working for the Gallatin National Forest Avalanche Center.

References

Armstrong, R.L., Armstrong, B.R., 1987. Snow and avalanche climates of the western United States: a comparison of maritime, intermountain and continental conditions. Avalanche Formation Movement and Effects. IAHS Publ. No. 162, pp. 281–294.

Birkeland, K.W., 1997. Spatial and temporal variations in snow stability and snowpack conditions throughout the Bridger Mountains, Montana. PhD dissertation, Department of Geography, Arizona State University, Tempe, AZ, 205 pp.

Birkeland, K.W., 1998. Terminology and predominant processes associated with the formation of weak layers of near-surface faceted crystals in the mountain snowpack. Arct. Alp. Res. 30 (2), 193–199.

Birkeland, K.W., Hansen, K.J., Brown, R.L., 1995. The spatial variability of snow resistance on potential avalanche slopes. J. Glaciol. 41 (137), 183–190.

Birkeland, K.W., Johnson, R.F., Herzberg, D., 1996. The stuffblock snow stability test. US Forest Service Tech. Rep. 9623-2836-MTDC, 20 pp.

Camponovo, C., Schweizer, J., 1996. Measurements on skier triggering. Proceedings of the 1996 International Snow Science Workshop, Banff, Alberta, Canada, pp. 100–103.

Conway, H., Abrahamson, J., 1984. Snow stability index. J. Glaciol. 30 (106), 321–327.

Föhn, P.M.B., 1987a. The stability index and various triggering mechanisms. Avalanche Formation Movement and Effects. IAHS Publ. No. 162, pp. 195–214.

Föhn, P.M.B., 1987b. The ''Rutschblock'' as a practical tool for slope stability evaluation. Avalanche Formation Movement, and Effects. IAHS Publ. No. 162, pp. 223–228.

Föhn, P.M.B., 1988. Snowcover stability tests and the areal variability of snow strength. Proceedings of the 1988 International Snow Science Workshop, Whistler, BC, Canada, pp. 262–273.

Jamieson, J.B., 1995. Avalanche prediction for persistent snow slabs. PhD dissertation, Department of Civil Engineering, University of Calgary, Calgary, Alberta, 258 pp.

Jamieson, J.B., Johnston, C.D., 1992. Experience with rutschblocks. Proceeding of the 1992 International Snow Science Workshop, Breckenridge, CO, pp. 150–159.

Jamieson, J.B., Johnston, C.D., 1993. Rutschblock precision, technique variations and limitations. J. Glaciol. 39 (133), 666–674.

Jamieson, J.B., Johnston, C.D., 1994. Monitoring a shear frame stability index and skier-triggered slab avalanches involving persistent snowpack weaknesses. Proceedings of the 1994 International Snow Science Workshop, Snowbird, UT, pp. 14–21.

Jamieson, J.B., Johnston, C.D., 1996. The compression test for snow stability. Proceedings of the 1996 International Snow Science Workshop, Banff, Alberta, Canada, pp. 118–125.

Johnson, R.F., Birkeland, K.W., 1994. The stuffblock: a simple and effective snow stability test. Proceedings of the 1994 International Snow Science Workshop, Snowbird, UT, pp. 518–526.

LaChapelle, E.R., 1966. Avalanche forecasting — a modern synthesis. International Symposium on the Scientific Aspects of Snow and Ice Avalanches. IAHS Publ. No. 69, pp. 350–356.

LaChapelle, E.R., 1980. The fundamental processes in conventional avalanche forecasting. J. Glaciol. 26 (94), 75–84.

Logan, N., 1992. Snow temperature patterns and artificial avalanche release. Proceedings of the 1992 International Snow Science Workshop, Breckenridge, CO, pp. 37–46.

McClung, D., Schaerer, P., 1993. The Avalanche Handbook. The Mountaineers, Seattle, WA, 271 pp.

Mock, C.J., 1995. Avalanche climatology of the continental zone of the southern Rocky Mountains. Physical Geography 16 (3), 165–187.

Mock, C.J., Kay, P.A., 1992. Avalanche climatology of the western United States, with an emphasis on Alta, Utah. Prof. Geogr. 44 (3), 307–318.

Perla, R.I., Beck, T.H.M., 1983. Experience with shear frames. J. Glaciol. 29 (103), 485–491.

Schaerer, P., 1988. Evaluation of the shovel shear test. Proceedings of the 1988 International Snow Science Workshop, Whistler, BC, Canada, pp. 274–276.

Zar, J.H., 1984. Biostatistical Analysis. Prentice Hall, Englewood Cliffs, NJ, 718 pp.