Cold Regions Science and Technology 59 (2009) 210–216

Contents lists available at ScienceDirect

Cold Regions Science and Technology

journal homepage: www.elsevier.com/locate/coldregions

The Extended Column Test: Test effectiveness, spatial variability, and comparison with the Propagation Saw Test

Ron Simenhois [a,b], Karl W. Birkeland [c,*]

[a] Copper Mountain Ski Area, Copper Mountain, Colorado, USA
[b] Mount Hutt Ski Area, Canterbury, New Zealand
[c] USDA Forest Service National Avalanche Center, Bozeman, Montana, USA

ARTICLE INFO

Article history:
Received 7 November 2008
Accepted 8 April 2009

Keywords:
ECT
Extended Column Test
Fracture propagation
Stability evaluation
Avalanches
Avalanche forecasting

ABSTRACT

The Extended Column Test (ECT) is a new stability test that aims to assess the fracture propagation potential across a 0.90 m wide isolated column. This paper: 1) describes the test procedure and presents new recording standards for the test, 2) uses two independent datasets (each consisting of over 300 tests) to assess the effectiveness of the test, 3) looks at the spatial variability of ECT results from several test grids, and 4) compares adjacent results between the ECT and the Propagation Saw Test (PST) on stable and unstable slopes. Our results indicate that the ECT is an effective stability test, with a false-stability rate less than other standard snow stability tests. Results are sometimes quite spatially uniform, though occasionally slopes may exhibit variable ECT results. In comparison to the PST, our data suggest that the ECT has a lower false-stability rate, but a higher false instability rate. Overall, the ECT is better at discriminating between stable and unstable slopes in our dataset. No test is perfect and all tests must be used in conjunction with additional data, but our results show that the ECT is valuable additional tool for assessing snow stability.

Published by Elsevier B.V.

1. Introduction

Avalanche forecasting relies on collecting diverse data, including data from the snowpack. The most highly prized snowpack data are what LaChapelle (1980) termed “low entropy” data or Fredston and Fesler (1994) call “bulls-eye” data. These are data that unambiguously inform the observer about the state of the snowpack, and include things like observing avalanches or hearing the snow collapse with a whumpfing sound (Johnson et al., 2004).

Other snowpack data might not be so unambiguous. For example, avalanche forecasters dig snowpits and do stability tests to help to ascertain whether the snowpack is unstable. However, interpreting stability tests is typically not straightforward, and most existing snowpit tests have false-stability rates around 10% (Birkeland and Chabot, 2006). In other words, when conducting such tests on slopes with clear signs of instability, observers can expect to get test results typically associated with stable slopes about 10% of the time. This value is unacceptably high and is why avalanche practitioners must use much more data than simply stability tests. Clearly, a need for better field stability tests exists.

The last few years have seen the development of two new tests. The Extended Column Test (ECT) (Simenhois and Birkeland, 2006) and the Propagation Saw Test (PST) (Gauthier and Jamieson, 2006a,b; Sigrist and Schweizer, 2007) both aim to investigate the fracture propagation potential of the snowpack. This is a critically important part of the avalanche puzzle since avalanche release requires both fracture initiation and fracture propagation along the weak layer (Schweizer et al., 2003; Gauthier and Jamieson, 2006b). Not only are these tests useful for stability evaluation, but they allow us to better investigate some of the factors associated with fracture propagation in the field, such as changes in slab depth (Simenhois and Birkeland, 2008a), snow surface warming (Simenhois and Birkeland, 2008b), and fracture propagation mechanics in weak snowpack layers (van Herwijnen et al., 2008).

The motivation for developing the PST and the ECT differed. Investigators developed the PST primarily as a fracture propagation test. On the other hand, we developed the ECT as a stability test. As with all stability tests, the primary goal of the ECT is to discriminate between stable and unstable slopes. Although not a pure fracture propagation test, we believe that the ECT does help to index the fracture propagation propensity of buried weak layers.

This paper synthesizes several recent papers on the ECT that have been written for practitioners (i.e., Simenhois and Birkeland, 2006; Simenhois and Birkeland, 2007; Birkeland and Simenhois, 2008) for the scientific community. The purpose of the paper is to: 1) describe the test procedure and document recent changes to recording standards for the ECT, 2) investigate the test's effectiveness for discriminating between stable and unstable slopes, 3) conduct a

* Corresponding author. Tel.: +1 406 587 6954.
E-mail address: kbirkeland@fs.fed.us (K.W. Birkeland).

0165-232X/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.coldregions.2009.04.001

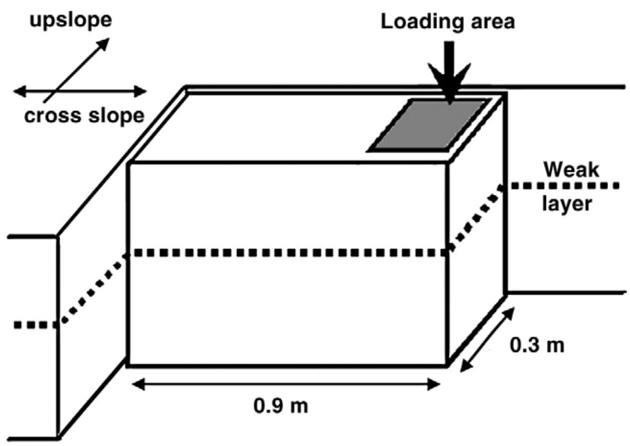

Fig. 1. The preparation of the ECT involves isolating a column 0.90 m across the slope by 0.30 m upslope. The column is then loaded from one side using the same technique as the compression test.

preliminary investigation of the spatial variability of ECT results, and 4) compare ECT results with results from the PST.

2. Extended Column Test procedure and recording standards

The Extended Column Test involves isolating a vertical column 0.9 m wide in the cross-slope dimension and 0.3 m deep in the upslope dimension that is deep enough to expose potential weak layers (Fig. 1). Depth should not exceed about 1.3 m since the loading steps rarely affect deeper layers. In fact, in our data the deepest test that propagated across the column was 1.04 m, and the deepest test that did not propagate across the column was 1.31 m. To conduct the test one end of the column is dynamically loaded using the loading steps of the compression test, whereby the tester taps a shovel ten times from the wrist, ten times from the elbow and then ten times from the shoulder (Greene et al., 2004). The observer notes the number of taps required to initiate a fracture in the weak layer below the shovel and whether or not the fracture propagates through the weak layer across the entire column.

The original recording standards for the ECT presented by Simenhois and Birkeland (2006) needed to be simplified and updated. The new standard better emphasizes what the test results are telling the user. Our findings, discussed later in this paper, emphasize the importance of whether or not a fracture propagates across the entire column, and this is reflected in the recording standards:

ECTPV—fracture propagates across the entire column through the weak layer or interface during isolation,
ECTP##—fracture initiates and propagates across the entire column through the weak layer or interface on the ## tap *or* the fracture initiates on the ## tap and propagates across the column on the ## + 1 tap,
ECTN##—fracture initiates on the ## tap but does not propagate across the entire column through the weak layer or interface on either the ## or the ## + 1 tap, and
ECTX—no fracture occurs in the weak layer during the test.

An advantage of the ECT is that test interpretation is straightforward. ECTPV and ECTP## results suggest unstable conditions because fracture propagation propensity is relatively high, while ECTN is generally indicative of stable conditions. With ECTX there is no fracture initiation, so we cannot evaluate the fracture propagation propensity for that layer. While an ECTX generally indicates stable conditions because fracture initiation is unlikely at the test location, previous spatial variability research shows that the force needed for fracture initiation can be widely variable across slopes (e.g., Campbell and Jamieson, 2007). Thus, we recommend using a different snowpack test when a user gets an ECTX result. Though the ECT is typically loaded with taps identical to the compression test, the same loading steps as the stuffblock test (Birkeland and Johnson, 1999) have been used by some researchers (Hendrikx and Birkeland, 2008) (Fig. 2).

3. Assessing ECT effectiveness

We use two independent datasets to test the effectiveness of the ECT in discriminating between stable and unstable slopes. Our first dataset consists of 324 tests conducted by the senior author during the winters of 2005/06 near Copper Mountain Ski Area in Colorado (202 tests) and Mount Hutt Ski Area in New Zealand (122 tests). These Colorado–New Zealand data cover two distinctly different snow climates and, since they are collected by a single observer, they are more consistent in data collection and in rating the slope stability. In addition, each pit includes all the typical snowpit observations described by Greene et al. (2004). Our second dataset comes from the SnowPilot database (Chabot et al., 2004) from 2006 to 2008 and is augmented with a number of tests from avalanche forecasters in the Spanish Pyrenees (Moner, pers. comm., 2007). Overall we found 311

Fig. 2. Though the ECT is typically loaded with the compression test loading steps (Greene et al., 2004), it can also be loaded using the stuffblock steps (Birkeland and Johnson, 1999), as shown above. From left to right: loading the extended column, column fractures (ECTP), examining the fractured slab and weak layer. Photos by Jordy Hendrikx.

ECT tests from nearly 20 different observers widely scattered throughout many different snow climates. Though the quality is more difficult to control with this SnowPilot dataset, we believe it offers an excellent comparison to our better controlled (though not as diverse) Colorado–New Zealand dataset.

3.1. Methods

We evaluated the performance of the ECT using contingency tables (Table 1) and the Hanssen–Kuipers discriminant (Hanssen and Kuipers, 1965), as has been done in previous snow research (e.g., Purves et al., 2003). The Hanssen–Kuipers discriminant, also called the True Skills Score (TSS), compares test accuracy to what one would expect from a set of random, unbiased predictions. If the test is a perfect predictor then TSS = 1, and if the test is only as good as random predictions then TSS = 0. A TSS < 0 means the test is performing worse than random predictions. With the contingency tables, we computed the following:

$$\text{Probability of detection (POD)} = \frac{d}{b+d} \tag{1}$$

$$\text{Probability of False Detection (or false instability rate)(POFD)} = \frac{c}{a+c} \tag{2}$$

$$\text{False stability rate} = 1 - \text{POD} \tag{3}$$

$$\text{True Skills Score (TSS)} = \text{POD} - \text{POFD} \tag{4}$$

A challenging part of this work is classifying whether a particular slope where a test has been conducted is stable or unstable. For our first dataset, the Colorado–New Zealand data, we defined "unstable" slopes as those with obvious signs of instability like observing shooting cracks or whumpfing, or triggering an avalanche (either with explosives or ski cuts) within a day or less from the time of testing. "Stable" slopes were tested by skiers or explosives, but did not present any of the above signs of unstable slopes. Of the 324 tests in this dataset, we conducted 256 tests (79%) on stable slopes and 68 tests (21%) on unstable slopes. Slope angles at the test locations varied from 25 to 41° on our stable slopes and from 25 to 42° on our unstable slopes, though portions of all the slopes in the Colorado–New Zealand dataset were greater than 30° and therefore steep enough to avalanche. For the unstable slopes we performed 12 tests (18%) within a day of an avalanche, another 12 (18%) within 4 h of the avalanche, 3 tests (4%) within an hour of the slide, and we performed 38 tests (56%) either shortly before or within 10 min from the time when a sign of instability (like cracks, a whumpf or an avalanche) was observed.

Our second dataset consisted of data from SnowPilot, a free software program that allows many different users to enter, graph, and database their snowpits at snowpilot.org (Chabot et al., 2004). The primary disadvantage of the SnowPilot database is the lack of quality control. Many different users with unknown levels of experience and training input data into SnowPilot, in sharp contrast to the Colorado–New Zealand dataset where the senior author carefully collected all the data. Also, in the SnowPilot dataset the slopes stability rating is subjective, is at the discretion of the observer, and sometimes conflicts with decisions on whether or not to ski a particular slope. However, those disadvantages are balanced by the ability to collect a great deal of data from diverse sources in many snow climates at a low cost. Our main intent with the SnowPilot data is to see whether or not it confirms the results in our Colorado–New Zealand dataset.

Table 1
Contingency table for assessing stability test performance.

		Observed slope conditions	
		Stable	Unstable
Test result	Stable	*a:* correct stables	*b:* misses (false stables)
	Unstable	*c:* false alarms (false unstables)	*d:* hits (correct unstables)

Table 2
Contingency tables comparing ECT results for our Colorado–New Zealand and SnowPilot datasets.

		Observed slope conditions			
		Colorado–New Zealand dataset ($n = 324$)		SnowPilot dataset ($n = 311$)	
		Stable	Unstable	Stable	Unstable
Test result	Stable (ECTN)	252	1	153	7
	Unstable (ECTP, ECTPV)	4	67	33	118

False-stable and false-unstable rates are higher with the more diverse, but less tightly controlled SnowPilot data.

To decide if a pit is on a stable or an unstable slope in SnowPilot we relied on the observer's similar slopes stability rating, comparable to the methods used by Birkeland and Chabot (2006) for their analysis of false-stable stability tests. If the observer rates the stability as good or better, we rate the slope as stable, while ratings of poor or very poor put the slope in the unstable category. If the stability was rated as fair or there was no stability rating, we rate those slopes that had no signs of instability as stable, while slopes with signs of instability such as cracking or whumpfing were rated unstable. The only exception is that slopes rated fair that had only localized collapsing and were skied by the observer were rated stable. For the 2007–08 season the ECT had to be on the "problematic layer or interface" identified by the user in SnowPilot, and ECTXs were not considered in our analyses since we were trying to see if fractures would propagate rather than initiate. For the SnowPilot data prior to the 2007–08 season we manually viewed pit graphs to attempt to identify the most problematic layer, a process that required some subjective judgment to try to better interpret a few cases. Clearly there are some flaws in this system since in some cases it relies on incomplete, subjective and inconsistent data. Further, the slope rating is not as definitive as the techniques we used to separate stable from unstable slopes in the Colorado–New Zealand dataset, and an observer's slope stability rating will be affected by the stability test results they observe. Still, we feel the diversity of these data make them valuable, and that our technique is reasonable for our analyses. Out of the 311 tests from SnowPilot, 186 tests (60%) were on slopes rated stable and 125 tests (40%) were done on slopes rated unstable. Over 40 tests were not used in our analysis because we could not determine the slope stability due to unclear, missing, or incomplete data.

3.2. Results and discussion

In our Colorado–New Zealand data the ECT effectively discriminated between stable and unstable slopes. Of 68 tests on unstable slopes, 67 tests (99%) were ECTP. On the other hand, out of 256 tests on stable slopes, 252 tests were ECTN and only four tests (2%) produced an ECTP result (Table 2). The TSS for the ECT for this dataset is 0.97, which is nearly perfect. Thus, for these data the ECT showed strong promise as a tool to discriminate between unstable and stable slopes with few misclassifications.

Our more diverse SnowPilot dataset also demonstrated the effectiveness of the ECT for identifying unstable slopes, though with an increase in misclassifications. Of the 125 tests on unstable slopes, 118 tests resulted in an ECTP, while in only seven cases (6%) did the fracture fail to fully propagate across the column (ECTN) (Table 2). This low false-stability rate is encouraging and is nearly half that reported for stability tests such as the compression test or the rutschblock (Birkeland and Chabot, 2006). The SnowPilot dataset does

Fig. 3. ECT results from 21 pits on a stable 32° slope at Mount Hutt range in New Zealand. None of the tests propagated across the entire column (ECTN).

show a higher rate of false-instability than the Colorado–New Zealand data. Of the 186 tests on stable slopes, the fracture propagated across the entire column (ECTP) in 33 tests (18%) (Table 2). Though the TSS for the ECT with these data dropped to 0.76, it still does a reasonably good job at discriminating between stable and unstable slopes.

Two other recent studies also report encouraging results for the ECT. Winkler and Schweizer (2008) compare the ECT to compression and rutschblock tests on a number of stable and unstable slopes in the Swiss Alps. They conclude that the ECT does a reasonably good job of differentiating between stable and unstable slopes, and that its performance in their dataset was clearly better than the compression test. They also suggest that reducing the number of taps to 21 or less might improve the ECT's performance, though this suggestion does not hold for our data. For example, in the Colorado–New Zealand dataset, it took more than 21 taps for ECTP results in nine cases out of the 68 unstable pits (13%), so adjusting our classification scheme would have greatly increased the number of false-stable results. Moner et al. (2008) report on 47 pits (29 on stable slopes and 18 on unstable slopes) in the Spanish Pyrenees. Their data resulted in a TSS of 0.85 for the ECT, quite a bit higher than the TSS of 0.59 they report for the rutschblock for their dataset. These two studies reinforce our data which show relatively low misclassification rates for the ECT.

Our results suggest the ECT is a useful tool for avalanche practitioners. However, no test is perfect. The presence of some misleading results highlights the necessity for avalanche workers to continue to use a variety of snow stability tests and combine those test results with avalanche, snowpack and weather observations for effective avalanche assessments.

One final note of caution about the ECT involves ECTXs, or where an ECT results in no fracture. As expected, our data suggest that ECTX results are more likely to occur on stable slopes. Out of 45 ECTX tests in the 2007–08 SnowPilot dataset, 37 (82%) occurred on stable slopes. However, the other eight (18%) tests were on slopes rated unstable, though some of these pits had contradictory stability test results. Still, this relatively high percentage of ECTX results on unstable slopes suggests that an ECTX should not be considered an absolute sign of stability. We believe this is because fracture initiation conditions are often quite variable across slopes (e.g., Campbell and Jamieson, 2007). For this reason we feel that in these cases other stability tests should also be conducted, and in many of our unstable SnowPilot pits with ECTX results these tests (compression and/or rutschblock tests) did indicate unstable conditions.

4. Spatial variability of ECT results

Observers report a strong relationship between shear quality (and closely related fracture character) and ECT results, and Campbell and Jamieson (2007) report reasonably spatially uniform fracture character in stability tests. As such, we hypothesize that ECT results may also be relatively spatially uniform. Though we do not have extensive data, here we report and discuss the results of two separate grids of ECT tests, and compare those results to grids conducted by Hendrikx and Birkeland (2008).

We sampled our first grid on 27 June 2006 on a slope near Mount Hutt ski area in New Zealand. Consisting of 21 tests, the grid spanned an area 30 m across the slope (5 m between tests) by 15 m down the slope (7.5 m between pits). Spatially uniform ECT results existed on the relatively planar 32° slope, with all 21 tests resulting in an ECTN (Fig. 3).

During the 2006/07 winter we conducted another spatial array of ECTs, this time on Tucker Mountain in Colorado. The array consisted of a 24 pit grid spanning an area 25 m across the slope (5 m between tests) by 15 m down the slope (5 m between tests). We rated slope stability as fair, with the same aspect and elevation as other slopes that avalanched two days earlier with explosives and ski cuts. However the slab that avalanched was confined to the top 15 m of the ridge tops. In our grid we found similar conditions, with a slab similar to the slab that produced avalanches in the location of the upper 17 pits and a softer slab at the other 7 pits (Figs. 4 and 5). ECT results on this grid were spatially uniform within the top 17 pits (ECTP) and within the other 7 pits (ECTN). There is a clear and explainable reason for the observed spatial variability, which is not always the case for the variability observed for some other tests which focus on fracture initiation (e.g., Landry et al., 2004). Indeed, the variability in ECT results observed appears to reflect the stability conditions on this particular slope.

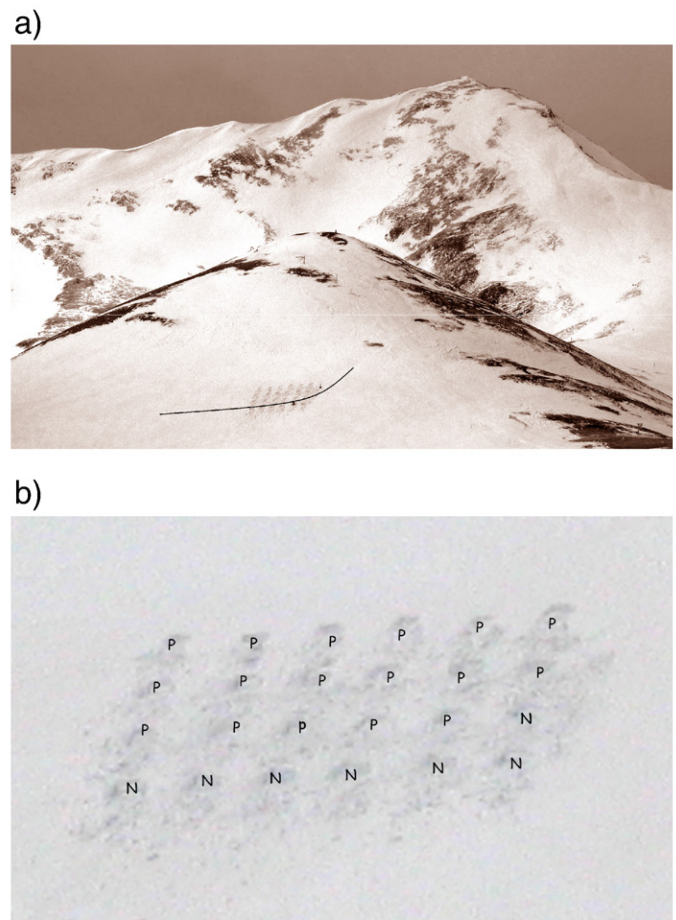

Fig. 4. An overview of a grid of 24 pits on Tucker Mountain in Colorado. The black line in a) marks the lower boundary of the hard slab involved in avalanches on similar slopes two days before our sampling. Photo b) is, a close-up of the grid showing the locations with ECTP results (shown as "P") and locations where the result was ECTN (shown as "N"). A reactive slab existed only at the upper left part of the grid, which is clearly reflected in the ECT results.

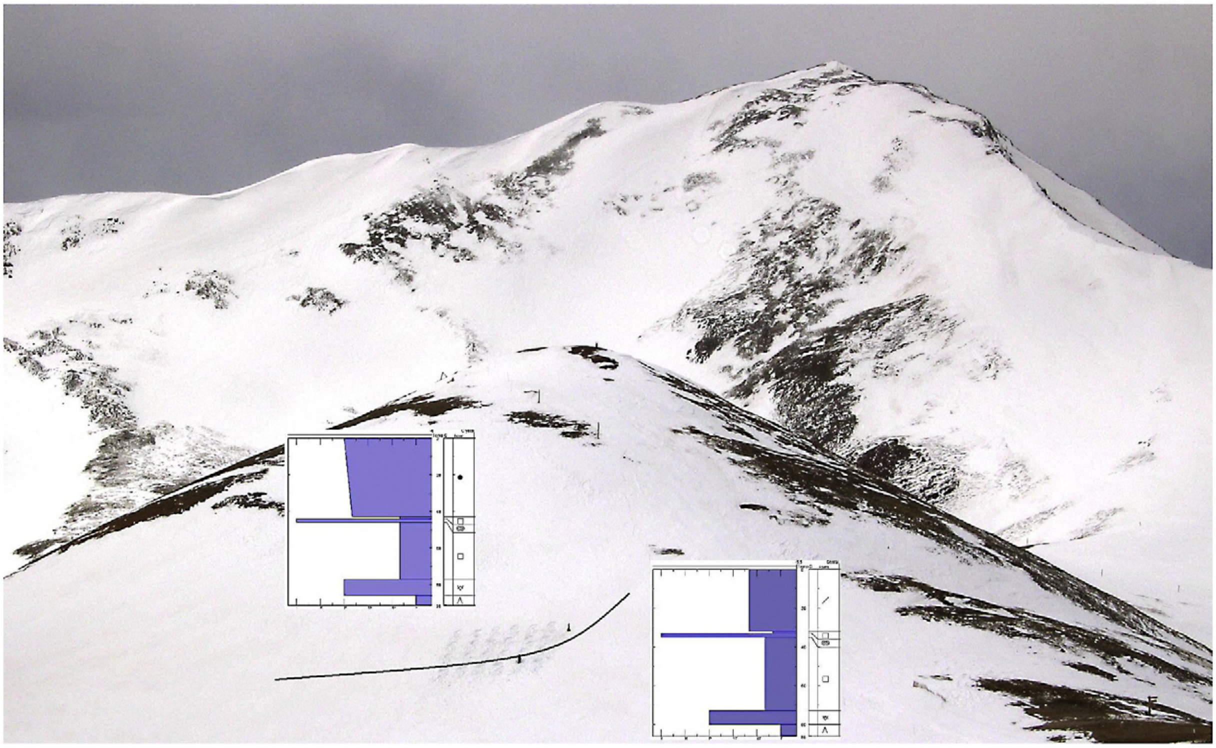

Fig. 5. An overview of the same 24 pits grid on Tucker Mountain, Colorado from Fig. 4. Representative pits on both sides of the black line are shown where the reactive slab existed (above the line) and where it did not (below the line).

During the 2007/08 winter Hendrikx and Birkeland (2008) collected an additional four spatial grids of ECTs in Montana, and two grids in New Zealand. The Montana grids exhibited some cases where slopes had a sizable percentage of ECTN results mixed in with ECTP results, resulting in no clear spatial pattern. This suggests that on some slopes we may have a significant percentage (up to 50%) of misclassifications. Some of these slopes also had relatively variable shear quality; on one slope most shears were Q1, but 31% were either Q2 or Q3. Conversely, the two New Zealand datasets showed consistent results, with all tests in those grids being ECTNs.

Clearly, additional work is required to better understand the slope-scale spatial variability of ECT results. Our results again emphasize that ECTs, like all stability tests, provide only one piece of the stability evaluation puzzle, and that a great deal of other data are necessary to accurately assess slope stability. Further, snowpit location is crucial and in some cases more than one snowpit may be required to improve the reliability of stability assessments.

5. Comparison of ECT and PST results

The ECT and PST are both useful tests for showing fracture propagation potential. However, few data exist on the two tests side-by-side. Though the PST was developed primarily as a fracture propagation test, in this part of the paper we treat both tests as stability tests, comparing their effectiveness in discriminating between unstable and stable slopes.

5.1. Methods

During the 2007/08 winter the senior author collected an additional dataset consisting of numerous side-by-side tests on both unstable and stable slopes adjacent to Copper Mountain ski area in Colorado. We grouped slope stability as we did for the Colorado–New Zealand dataset used to investigate the effectiveness of the ECT, classifying slopes as unstable if they had recently avalanched, or they showed other obvious signs of instability like cracking or whumpfing. Alternately, stable slopes were tested with explosives or heavy skier traffic and did not avalanche. In total, we did 45 sets of tests on unstable slopes and 33 sets of tests on stable slopes. In addition to the ECT and PST data, we collected standard snowpit observations following Greene et al. (2004). Test interpretation followed previous work; ECTP and ECTPV indicated unstable conditions, while ECTN and ECTX were stable. The PST involves isolating a block 0.3 m in the cross-slope direction and 1.0 m (or as long as the slab thickness if it is greater than 1.0 m) in the upslope direction. The observer locates the weak layer and uses the back of a snow saw to cut through the weak layer, noting the point at which the fracture takes off from the saw and either propagates up through the slab or to the end of the column along the weak layer. With the PST, cut lengths less than half the column length that resulted in self-propagation along the weak layer through the entire column indicated unstable conditions, while longer cut lengths or fractures that did not self-propagate indicated stable conditions (Gauthier and Jamieson, 2008). We compare the tests using contingency tables and the measures discussed previously.

5.2. Results and discussion

Of the 45 tests in 31 pits on unstable slopes, in all cases the ECT fully propagated across the column, resulting in a false-stability rate of 0% for the test (Table 3). This result is reasonably similar to the Colorado–New Zealand and the SnowPilot datasets where false-stability rates were less than about 6%. In the 45 PSTs, 25 of them propagated with a cut length of 0.50 m or less. This resulted in 20 PSTs on these slopes that indicated stable snowpack conditions, or a false-stability rate of 44% (Table 3). Our rate for the PST is higher than previously reported; Gauthier and Jamieson (2008) report a false-stable rate of 30% for 113 tests on unstable slopes in Canada. A reason for the high false-stability rate in our data may be due to the fact that nine of our 31 unstable pits were on a single slope on a crown wall and flanks of an avalanche from 12 January 2008, about 2 h after it slid. Perhaps the weak layer–slab combination on this particular made it particularly likely to produce false-stable PST results. The weak layer

Table 3
Contingency tables comparing ECT results to PST results for adjacent tests in Colorado.

		Extended Column Test		Propagation Saw Test	
		Stable slopes	Unstable slopes	Stable slopes	Unstable slopes
Test result	Stable	30	0	33	20
	Unstable	3	45	0	25

We conducted 33 pairs of tests on stable slopes and 45 pairs on unstable slopes. Unstable ECT results are ECTP and ECTPV and unstable PST results are cut lengths of 0.50 m or less self-propagate through the weak layer to the end of a 1 m long column.

on this slope consisted of 1 mm of 1F+ near-surface facets buried under a harder slab that varied in thickness from 0.40 to 0.82 m and that was P hardness directly above the weak layer. In all the PSTs on this slope, the saw cut length was more than 0.90 m and compression test results in those pits were above 20 taps. In our dataset we also observed false stables for the PST when the slab above the weak layer was soft (4 fingers or less), similar to descriptions by Gauthier and Jamieson (2006a, 2008). Other tests, such as the compression, stuffblock, and rutschblock tests have false-stability rates closer to 10% (Birkeland and Chabot, 2006).

The results for stable slopes are different. Of the 33 tests on stable slopes, in all cases the cut length for the PST was greater than 0.50 m. In other words, the false instability rate for the PST in our results was 0% (Table 3). Gauthier and Jamieson (2008) report a false instability rate of 5% for the 57 tests they conducted on stable slopes. Looking at ECT results reveals that the fracture propagated across the extended column completely in three cases, for a false instability rate of about 9% (Table 3).

Taking both unstable and stable slopes together, the TSS for the ECT for these side-by-side comparisons is 0.91, while the TSS for the PST is 0.56. Thus, in our data the ECT is doing a better job of discriminating between stable and unstable slopes, and also has the advantage of a lower false-stability rate.

Independent of and simultaneous to this work, Ross and Jamieson (2008) also evaluated the ECT and the PST. They collected a large dataset, but did not compare their results against independent observations of fracture propagation potential. They concluded that both tests work well, but that in Canada's Columbia Mountains the PST is more effective for depths greater than 0.70 m. In their work, the ECT worked well for depths from about 0.30 to 0.70 m. In our side-by-side tests, slab depths averaged 0.60 m on our unstable slopes and 0.54 m on our stable slopes, with slabs varying from 0.27 to 1.04 m. We did not observe a pattern of changing relationships between the two tests in our smaller dataset, but we acknowledge that a primary advantage of the PST is that it likely handles deeper weak layers—where fracture initiation with the ECT is difficult or impossible using standard techniques—more effectively. However, our work suggests that the ECT still works well for slab depths on the order of 1 m in the Colorado snowpacks where we did our comparative tests.

Taken together, our results suggest that the false-stability rate for the ECT is quite low, and is lower than that of other stability tests as reported by Birkeland and Chabot (2006). With stability tests a primary goal is to have a low false-stability rate since the most dangerous situation is when you are collecting data that indicate that a slope is stable when it is in fact unstable. Conversely, the false-instability rate is lowest for the PST, and is higher for the ECT.

6. Conclusions

The Extended Column Test offers a new way of testing the snow stability, with a focus on examining the fracture propagation potential of the slab/weak layer combination. Our work indicates that the test works well, and that it is a valuable addition to other tests. Interestingly, interpreting ECT results is relatively unambiguous and the test appears to have a lower false-stability rate than many other tests, which enhances its practical usefulness for slope stability evaluation. However, it also has a reasonably high rate of false instability (up to 18%). Still, the rapid acceptance of this test by practitioners around the world in only two seasons attests to its practical usefulness for field testing.

Besides the practical application of the test, the ECT offers scientists a tool to investigate changes in fracture propagation over space and time. In other papers we utilize the ECT to show changes in fracture propagation potential with changes in slab depth (Simenhois and Birkeland, 2008a), with surface warming (Simenhois and Birkeland, 2008b), and over space and time (Hendrikx and Birkeland, 2008; Hendrikx et al., 2009).

Acknowledgements

The American Avalanche Association provided partial support for this work through a practitioners grant to the senior author, and we would like to thank the Copper Mountain and Mount Hutt Ski Patrols for giving him the time and opportunity to collect data. Thanks to Ivan Moner Seira and his colleagues from the Centre De Prediccion De Lauegi Dera Val D'aran for sharing their data from the Pyrenees, to Doug Chabot and Mark Karhl for all their work on SnowPilot, and to all the people who submitted their ECT results to SnowPilot the last two seasons. Ethan Greene and Doug Chabot provided great comments and feedback on updating the ECT recording system, and Jordy Hendrikx participated in numerous discussions about the spatial variability of ECT results. We thank Kurt Winkler and an anonymous reviewer for providing extensive comments which greatly improved the clarity of this paper.

References

Birkeland, K.W., Chabot, D., 2006. Minimizing “false stable” stability test results: why digging more snowpits is a good idea. In: Gleason, A. (Ed.), Proceedings of the 2006 International Snow Science Workshop, Telluride, Colorado, pp. 498–504.
Birkeland, K.W., Johnson, R.F., 1999. The stuffblock snow stability test: comparability with the rutschblock, usefulness in different snow climates, and repeatability between observers. Cold Reg. Sci. Technol. 30 (1–3), 115–123.
Birkeland, K.W., Simenhois, R., 2008. The extended column test: test effectiveness, spatial variability, and comparison with the propagation saw test. In: Campbell, C., Conger, S., Haegeli, P. (Eds.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., pp. 401–407.
Campbell, C., Jamieson, J.B., 2007. Spatial variability of slab stability and fracture characteristics within avalanche start zones. Cold Reg. Sci. Technol. 47 (1–2), 134–147.
Chabot, D., Kahrl, M., Birkeland, K.W., Anker, C., 2004. SnowPilot: a “new school” tool for collecting, graphing, and databasing snowpit and avalanche occurrence data with a PDA. In: Elder, K. (Ed.), Proceedings of the 2004 International Snow Science Workshop, Jackson Hole, Wyoming, p. 476.
Fredston, J.A., Fesler, D., 1994. Snow Sense: A Guide to Evaluating Snow Avalanche Hazard. Alaska Mountain Safety Center, Anchorage, Alaska, U.S.A. 116 pp.
Gauthier, D., Jamieson, J.B., 2006a. Evaluating a prototype field test for weak layer fracture and failure propagation. In: Gleason, A. (Ed.), Proceedings of the 2006 International Snow Science Workshop, Telluride, Colorado, pp. 107–116.
Gauther, D., Jamieson, J.B., 2006b. Understanding the propagation of fractures and failures leading to large and destructive snow avalanches: recent developments. Proceedings of the 2006 Annual Conference of the Canadian Society for Civil Engineering, First Specialty Conference on Disaster Mitigation, 23/05/2006, Calgary, Alberta.
Gauthier, D., Jamieson, J.B., 2008. Fracture propagation propensity in relation to snow slab avalanche release: validating the propagation saw test. Geophys. Res. Lett. 35 (L13501). doi:10.1029/2008GL034245.
Greene, E.M., et al., 2004. Snow, Weather and Avalanches: Observational Guidelines for Avalanche Programs in the United States. American Avalanche Association, Pagosa Springs, Colorado. 136 pp.
Hanssen, A.W., Kuipers, W.J.A., 1965. On the relationship between the frequency of rain and various meteorological parameters. Koninklijk Netherlands Meteorologisch Institus, Meded. Verhand. 81, 2–15.
Hendrikx, J., Birkeland, K.W., 2008. Slope scale spatial variability across time and space: comparison of results from continental and maritime climates. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., pp. 155–162.
Hendrikx, J., Birkeland, K.W., Clarke, M., 2009. Assessing changes in the spatial variability of the snowpack fracture propagation propensity over time. Cold Reg. Sci. Tech. 56 (2–3), 152–160. doi:10.1016/j.coldregions.2008.12.001.
Johnson, B.C., Jamieson, J.B., Stewart, R.R., 2004. Seismic measurement of fracture speed in a weak snowpack layer. Cold Reg. Sci. Technol. 40 (1–2), 41–46.

LaChapelle, E.R., 1980. The fundamental process in conventional avalanche forecasting. J. Glaciol. 26 (94), 75–84.
Landry, C.C., et al., 2004. Variations in snow strength and stability on uniform slopes. Cold Reg. Sci. Technol. 39 (2–3), 205–218.
Moner, I., 2007. Personal communication. Avalanche forecaster, Centre De Prediccion De Lauegi Dera Val D'aran.
Moner, I., Gavaldà, J., Bacardit, M., Garcia, C., Martì, G., 2008. Application of field stability evaluation methods to the snow conditions of the eastern Pyrenees. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., pp. 386–392.
Purves, R.S, Morrison, K.W., Moss, G., Wright, D.S.B., 2003. Nearest neighbours for avalanche forecasting in Scotland — development, verification and optimization of a model. Cold Reg. Sci. Technol. 37 (3), 343–355.
Ross, C., Jamieson, J.B., 2008. Comparing fracture propagation tests and relating test results to snowpack characteristics. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., 376–385.
Schweizer, J., Jamieson, J.B., Schneebeli, M., 2003. Snow avalanche formation. Rev. Geophys. 41 (4). doi:10.1029/2002RG000123.
Sigrist, C., Schweizer, J., 2007. Critical energy release rates of weak snowpack layers determined in field experiments. Geophys. Res. Lett. 34 (3), L03502.
Simenhois, R., Birkeland, K.W., 2006. The extended column test: a field test for fracture initiation and propagation. In: Gleason, A. (Ed.), Proceedings of the 2006 International Snow Science Workshop, Telluride, Colorado, pp. 79–85.
Simenhois, R., Birkeland, K.W., 2007. An update on the Extended Column Test: new recording standards and additional data analyses. Avalanche Rev. 26 (2), 17–18.
Simenhois, R., Birkeland, K.W., 2008a. The effect of changing slab thickness on fracture propagation. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., 755–760.
Simenhois, R., Birkeland, K.W., 2008b. Extended column test results during surface warming: does surface warming affect fracture propagation potential? In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, B.C., pp. 26–35.
Winkler, K., Schweizer, J., 2008. Comparison of different snow stability tests including the extended column test. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, British Columbia, pp. 393–400.
van Herwijnen, A., Heierli, J., Schweizer, J., 2008. Field study on fracture propagation in weak snowpack layers. In: Campbell, C. (Ed.), Proceedings of the 2008 International Snow Science Workshop, Whistler, British Columbia, pp. 1–8.