# THE EXTENDED COLUMN TEST: TEST EFFECTIVENESS, SPATIAL VARIABILITY, AND COMPARISON WITH THE PROPAGATION SAW TEST

Karl W. Birkeland[1,*] and Ron Simenhois[2,3]

[1]USDA Forest Service National Avalanche Center, Bozeman, Montana, USA
[2]Copper Mountain Ski Area, Copper Mountain, Colorado, USA
[3]Mount Hutt Ski Area, Canterbury, New Zealand

ABSTRACT: The Extended Column Test (ECT) is a new stability test that aims to assess the fracture propagation potential across a 90 cm wide isolated column. Initial results with a dataset consisting of over 300 tests collected by one observer demonstrated the ECT's effectiveness for differentiating between stable and unstable slopes. Further, we have received positive feedback on the test from a world-wide network of observers. This paper: 1) presents new recording standards for the test, 2) uses the SnowPilot dataset to further assess the effectiveness of the test by analyzing over 300 tests performed by several observers in different snow climates, 3) looks at the spatial variability of ECT results from several test grids, and 4) compares side-by-side results between the ECT and the Propagation Saw Test on stable and unstable slopes. Our results indicate that the ECT is an effective stability test, with a false stability ratio generally 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 the ECT has a lower false stability rate, but a higher false instability rate. No test is perfect and all tests must be used in conjunction with additional data, but our results show the ECT is valuable additional tool for assessing snow stability.

KEYWORDS: ECT, extended column test, fracture propagation, stability evaluation, avalanches

## 1. INTRODUCTION

Avalanche forecasting relies of the collection of 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 whumpf.

Other snowpack data might not be so unambiguous. For example, people dig snowpits and do stability tests to try to ascertain whether the snowpack is unstable. However, interpreting stability tests is not always straightforward, and most existing snowpit tests have false stability ratios around 10% (Birkeland and Chabot, 2006). In other words, when conducting such tests on unstable slopes, observers can expect to get a test results typically associated with stable slopes about 10% of the time. This value is unacceptably high and is the reason 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, 2006) 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. 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) and snow surface warming (Simenhois and Birkeland, 2008b).

This paper updates the work by Simenhois and Birkeland (2006; 2007) and presents additional data. The purpose of the paper is to do the following: 1) Document the changes to recording standards for the ECT, 2) Investigate the test's effectiveness for discriminating between stable and unstable slopes, 3) Examine the spatial variability of ECT results, and 4) Compare ECT results with results from the PST.

* Corresponding author address: Karl Birkeland, USDA Forest Service Nat'l Avalanche Center, P.O. Box 130, Bozeman, MT 59771 USA, tel: 406-587-6954, email: kbirkeland@fs.fed.us

*Figure 1: Though the ECT is typically loaded with the compression test loading steps, it can also be loaded using the stuffblock steps, 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.*

## 2. RECORDING STANDARDS FOR THE ECT

Discussions with a number of individuals led us to the conclusion that the original recording standards for the ECT presented by Simenhois and Birkeland (2006) needed to be simplified and updated. Our main goal in establishing a new standard was to try to emphasize what the test results are telling the user. Our results (Simenhois and Birkeland, 2006; 2007; and below) emphasize the importance of whether or not a fracture propagates across the entire column (now coded as ECTP) or not (now ECTN), and this needed to be reflected in the way the test results were recorded. In the end we came up with:

**ECTPV** – fracture propagates across the entire column during isolation,
**ECTP##** - fracture initiates and propagates across the entire column in ## or ##+1 taps,
**ECTN** – fracture doe not propagate across the entire column, or there are 2 or more taps between the initiation and propagation of the fracture, and
**ECTNR** – no fracture occurs during the test

Test interpretation is straightforward. ECTPV and ECTP## results suggest unstable conditions, while ECTN or ECTNR are generally indicative of stable conditions. Though the ECT is typically loaded with taps identical to the compression test, it can also be loaded with the

same loading steps as the stuffblock test (Birkeland and Johnson, 1999) (Figure 1). Test interpretation is the same no matter which way the block is loaded.

## 3. ASSESSING ECT EFFECTIVENESS

Simenhois and Birkeland (2006) provided the first assessment of the ECT using a dataset of over 300 tests conducted by Simenhois during the winters of 2005/06 near Copper Mountain Ski Area in Colorado and Mount Hutt Ski Area in New Zealand. This work defined "unstable" slopes as those with obvious signs of instability like cracking, collapsing, or recent avalanche activity, while "stable" slopes are steep enough to slide (≥30°) and were tested by skiers or explosives, but did not present any of the above signs of unstable slopes. The analyses compared ECT test data from stable and unstable slopes, and showed the ECT demonstrated strong promise as a stability evaluation tool with very few misclassifications.

Simenhois and Birkeland (2007) built on these data using the SnowPilot database (Chabot et al., 2004) and a dataset from avalanche forecasters in the Pyrenees. Those data showed higher false stable and false unstable ratios, but the false stable ratios were still well below the ratios for other stability tests. This paper builds on that work by combining those data with the SnowPilot data from 2007-08 and analyzing the complete dataset.

*Table 1: Side-by-side comparisons of the ECT results for our original dataset (Simenhois and Birkeland, 2006) and for the SnowPilot data presented in this paper (SnowPilot data, 2006-2008). False-stable and false-unstable rates are higher with the more diverse, but less tightly controlled SnowPilot data. Unstable test results are ECTPV or ECTP.*

|  | Tests on unstable slopes | | Tests on stable slopes | |
|---|---|---|---|---|
|  | Simenhois and Birkeland (2006) (n = 68) | SnowPilot data, 2006-2008 (n = 125) | Simenhois and Birkeland (2006) (n = 256) | SnowPilot data, 2006-2008 (n = 186) |
| Unstable test result | 68 (100%) | 118 (94%) | 4 (2%) | 33 (18%) |
| Stable test result | 0 (0%) | 7 (6%) | 252 (98%) | 153 (82%) |

*3.1 Methods:*

The growing acceptance of the ECT as well as the use of SnowPilot allowed us to collect more diverse data from different observers and mountain ranges. We investigated the past two winter's collection of pits in the SnowPilot database and identified pits with ECT observations. Overall we found 311 ECT tests from nearly 20 different observers widely scattered throughout many different snow climates. We believe this dataset offers a good comparison to the better controlled (though not as diverse) dataset used previously (Simenhois and Birkeland, 2006).

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 stability rating was good or higher, 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 or had been skied with localized signs of instability as stable. Otherwise they rate as unstable. For the 2007-08 season the ECT had to be on the "problematic layer or interface" identified by the user in SnowPilot, and ECTNRs were not considered in our analyses since we were trying to see if fractures would propagate rather than initiate. We ended up manually viewing many of the pit graphs, and we used 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. The slope rating is not as definitive as the techniques we used to separate out stable from unstable slopes in our previous ISSW paper (Simenhois and Birkeland, 2006). 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 and our colleagues in the Pyrenees, 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:*

As we reported previously, our first season's data (collected by R. Simenhois) demonstrated the effectiveness of the ECT at discriminating between stable and unstable slopes (Simenhois and Birkeland, 2006). With these highly controlled data collected by one observer the false stability ratio was 0% and the false instability ratio was only 3% (Table 1).

Our more diverse SnowPilot dataset also demonstrated the effectiveness of the ECT for identifying unstable slopes. 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 1). 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 show a higher rate of false-instability than our original data. Of the 186 tests on stable slopes, the fracture propagated across the entire column (ECTP) in 33 tests (18%) (Table 1). This false-instability ratio was about six times higher than in our original data.

The low rates of false-stability and false-instability emphasize the usefulness of the ECT as an additional tool for avalanche professionals. 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 ECTNRs, or where an ECT results in no fracture. As expected, our data suggest that ECTNR results are more likely to occur on stable slopes. Out of 45 ECTNR 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 ECTNR results on unstable slopes suggests that an ECTNR should not be considered an absolute sign of stability. In these cases other stability tests should also be conducted, and in many of our unstable SnowPilot pits with ECTNR results these tests (compression and/or rutschblock tests) did indicate unstable conditions.

## 4. SPATIAL VARIABILITY OF ECT RESULTS

The low false stability ratios of the ECT suggest that ECT results may be relatively spatially uniform. Further, observers report a strong relationship between shear quality (and closely related fracture character) and ECT results, and Campbell and Jamieson (2007) report relatively spatially uniform fracture character in stability tests. Consistent with these results, Simenhois and Birkeland (2006) reported results from a slope with 21 ECT results. The result of every test on the slope was ECTN, suggesting that, at least for this particular slope, ECT results were spatially uniform.

During the 2006/07 winter we conducted another spatial array of ECTs, this time on Tucker Mountain in Colorado. These results are described in Simenhois and Birkeland (2007), but will be briefly discussed here as well. The array consisted of a 24 pit grid spanning an area 30 m across the slope by 20 m down the slope. 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 (Figure 2). 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



*Figure 2: An overview of a grid of 24 pits on Tucker Mountain in Colorado. The black line in the photo at left marks the lower boundary of the hard slab involved in avalanches on similar slopes two days before our sampling. At right, 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"). An active slab existed only at the upper left part of the grid, which is clearly reflected in the ECT results.*

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 actual stability conditions on this particular slope.

During the 2007/08 winter we collected an additional four spatial grids of ECTs in Montana, and two grids in New Zealand in 2008. These grids and their analyses are more thoroughly described in Hendrikx and Birkeland (2008). In the Montana grids we found some cases where slopes had a sizable percentage of ECTN results mixed in with ECTP results, with no clear pattern existing for the results. This suggests that on some slopes we may have a significant percentage (up to 50%) of false unstable results. 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, our two New Zealand datasets were quite consistent, with nearly all results in those grids being ECTNs. We are trying to sort out the inconsistencies between these data and previous data by collecting and analyzing additional spatial datasets of ECTs.

Our slope-scale spatial variability 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 of the two tests side-by-side. In this part of the paper we compare the two tests on both unstable and stable slopes.

### 5.1 Methods

During the 2007/08 winter the second author conducted numerous side-by-side tests on both unstable and stable slopes. Like our previous work, unstable slopes were slopes that had recently avalanched, or they showed other obvious signs of instability like cracking or collapsing. 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 ECTV indicated unstable conditions, while ECTN or ECTNR were stable. With the PST, cut lengths less than half the column length of 1 m that resulted in self-propagation along the entire column indicated unstable conditions, while longer cut length or fractures that did not self-propagate indicated stable conditions (Gauthier and Jamieson, 2007).

### 5.2 Results and discussion

Of the 45 tests on unstable slopes, in all cases the ECT fully propagated across the column, resulting in a false stability ratio of 0% for the test (Table 2). This result is similar to those presented by Simenhois and Birkeland (2006; 2007) where false stability ratios were less than 5%. In the 45 PSTs, 25 of them propagated with a cut length of 50 cm or less. This resulted in 20 PSTs on these slopes that indicated stable snowpack conditions, or a false stability ratio of 44% (Table 2). Our rate for the PST is higher than

Table 2: *Side-by-side comparisons of the ECT and PST indicate that the ECT has a lower false-stable ratio while the PST has a lower false-unstable ratio. Unstable test results are ECTPV or ECTP for the ECT, or PST cut lengths greater than 50 cm for a 1 m column.*

|  | Tests on unstable slopes (n = 45) | | Tests on stable slopes (n = 33) | |
|---|---|---|---|---|
|  | ECT | PST | ECT | PST |
| Unstable test result | 45 (100%) | 25 (56%) | 3 (9%) | 0 (0%) |
| Stable test result | 0 (0%) | 20 (44%) | 30 (91%) | 33 (100%) |

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 a single avalanche from 12 January, 2008, about two hours after it slid. Perhaps something about this particular slope or slab configuration made it particularly likely to produce false stable PST results. The weak layer on this slope consisted of 1mm of 1F+ near-surface facets buried under a P hard slab that varied from 40-82 cm in thickness. In all the PSTs on this slope, the saw cut length was more than 90 cm and compression test results in those pits were above 20 taps. On the other hand, in our dataset, we also observed false stabiles when the slab above the weak layer was soft (4 fingers or less) as described by Gauthier and Jamieson (2006). Other tests, such as the compression, stuffblock, and rutschblock tests have false-stability ratios more in the neighborhood of about 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 50 cm. In other words, the false instability ratio for the PST in our results was 0% (Table 2). Gauthier and Jamieson (2008) report a false instability ratio 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 ratio of about 9% (Table 2).

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 the Columbia Mountains (where they did their work) the PST is more effective for depths greater than 70 cm. In their work, the ECT worked well for depths from about 30 to 70 cm. In our side-by-side tests, slab depths averaged 60 cm on our unstable slopes and 54 cm on our stable slopes, with slabs varying from 27 to 104 cm. 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 or perhaps a little more in the Colorado snowpacks where we did our comparative tests.

Taken together, our results suggest that the false-stability ratio for the ECT is quite low, and is lower than that of other stability tests. With stability tests a primary goal is to have a low false stability ratio 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 ratio 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, the ECT appears to have a lower false-stability ratio than any of the 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 at this conference 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).

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.

## References

Birkeland, K.W. and Chabot, D., 2006. Minimizing "false stable" stability test results: Why digging more snowpits is a good idea. In: A. Gleason (Editor), International Snow Science Workshop, Telluride, Colorado, pp. 498-504.

Birkeland, K.W. and Johnson, R.F., 1999. The stuffblock snow stability test: comparability with the rutschblock, usefulness in different snow climates, and repeatability between observers. Cold Regions Science and Technology, 30(1-3): 115-123.

Campbell, C. and Jamieson, J.B., 2007. Spatial variability of slab stability and fracture characterististics within avalanche start zones. Cold Regions Science and Technology, 47(1-2): 134-147.

Chabot, D., Kahrl, M., Birkeland, K.W. and Anker, C., 2004. SnowPilot: A "new school" tool for collecting, graphing, and databasing snowpit and avalanche occurrence data with a PDA. In: K. Elder (Editor), International Snow Science Workshop, Jackson Hole, Wyoming, pp. 476.

Fredston, J.A. and 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. and Jamieson, J.B., 2006. Evaluating a prototype field test for weak layer fracture and failure propagation. In: A. Gleason (Editor), 2006 International Snow Science Workshop, Telluride, Colorado, pp. 107-116.

Gauthier, D. and Jamieson, J.B., 2007. The propagation saw test. Avalanche.ca - Journal of the Canadian Avalanche Community, 83: 60-63.

Gauthier, D. and Jamieson, J.B., 2008. Fracture propagation propensity in relation to snow slab avalanche release: Validating the propagation saw test. Geophysical Research Letters, 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.

Hendrikx, J. and Birkeland, K.W., 2008. Slope scale spatial variability across time and space: Comparison of results from continental and maritime climates, 2008 International Snow Science Workshop, Whistler, British Columbia.

LaChapelle, E.R., 1980. The fundamental process in conventional avalanche forecasting. Journal of Glaciology, Vol. 26, No. 94: 75-84.

Landry, C.C. et al., 2004. Variations in snow strength and stability on uniform slopes. Cold Regions Science and Technology, 39(2-3): 205-218.

Ross, C. and Jamieson, J.B., 2008. Comparing fracture propagation tests and relating test results to snowpack characteristics. In: C. Campbell (Editor), 2008 International Snow Science Workshop, Whistler, British Columbia, this volume.

Simenhois, R. and Birkeland, K.W., 2006. The extended column test: A field test for fracture initiation and propagation. In: A. Gleason (Editor), 2006 International Snow Science Workshop, Telluride, Colorado, pp. 79-85.

Simenhois, R. and Birkeland, K.W., 2007. An update on the Extended Column Test: New recording standards and additional data analyses. The Avalanche Review, 26(2): 17-18.

Simenhois, R. and Birkeland, K.W., 2008a. The effect of changing slab thickness on fracture propagation, 2008 International Snow Science Workshop, Whistler, British Columbia.

Simenhois, R. and Birkeland, K.W., 2008b. Extended column test results during surface warming: Does surface warming affect fracture propagation potential? 2008 International Snow Science Workshop, Whistler, British Columbia.