Comparison of structural instability indices (Lemons), ECT and RB results

Chiambretti Igor,[1,*] Fabiano Monti,[2] and Mauro Valt,[3]
[1] AINEVA, Trento (TN), Italy
[2] SLF, Davos, Switzerland
[3] ARPA Veneto - Centro Valanghe di Arabba, Arabba (BL), Italy

ABSTRACT: Almost all avalanche forecasting services have a network of traditional stations where an array of daily or weekly snowpack measurement or observations are made (ram penetrometer, snow profile, stability tests). Ram penetrometer and snow profile data are classified as medium entropy type while stability test are a low entropy ones and they requires an interpretation by the avalanche forecaster. In recent years, several methods have been developed to analyze the profiles and to evaluate, more objectively, snowpack stability identifying the weak layers and their characteristics and properties. During the period 2010-2013 several AINEVA's regional offices have collected a dataset consisting of hundreds of snow profile each accompanied by side-by-side stability tests (ECT and/or RB). A sample of this data set was analyzed following the critical variables method and finding interesting relations between weak layer characteristics (structural instability indices or "lemons", grain type, layer thickness, weak layer properties) and stability test results. The preliminary data of this research project are here discussed and shows that the best correlation between the highest values of the structural instability indices and stability test results are detectable for weak layers developed due to medium to high temperature gradient metamorphism. Less good correlation have been detected for weak layers composed by new snow (PP and DF) or by various type of crust or smooth interface between layers. The complete data set has been, finally, analyzed to compare ECT versus RB effectiveness in discriminating the main weak layer and finding relations between load steps, fracture character and depth, quality shear and weak layer characteristics.

KEYWORDS: RB, ECT, Structural instability indices.

1 INTRODUCTION

The snowpack observation allows to depict the complex layered structure of the snow cover as each snowfall accumulates through a layer bonding, somehow, onto the previous snow cover surface. Metamorphic processes, developing during the winter season, change through time the characteristics of those layers often developing new layers and type of grains. Failure initiation (crack growth) inside the snowpack and fracture propagation (possible avalanche release) are controlled by several parameters related to the mechanical properties of each layer.

Failure initiation (crack nucleation) happens when an additional stress locally overcome the strength of the weakest layer or interface between layers. The spreading of such crack (local failure) along the layer or interface develops a fracture which can propagate, with different mechanisms, through the rest of the snowpack, which might develop into a catastrophic failure (avalanche) when the fracture toughness is overcome. Following McCammon and Sharaf (2005) snowpack observations focused onto stability evaluation should record the following parameters: snowpack structure (layering), fracture initiation (strength) and fracture propagation (toughness). The execution of a snow profile and associated stability tests (rutschblock – RB - Föhn, 1987 and extended column test – ECT - Simenhois and Birkeland, 2006) allows that type of observation. The potential weak layer are located and described (position, grain type, grain size, hardness), the stability test identifies the failure layer (test score, fracture character or release type or shear quality – *sensu* van Herwijnen and Jamieson, 2004). If all the observations are executed properly, structural instability indices based on threshold sums such as the lemons or yellow flags (Jamieson and Schweizer, 2005; McCammon and Schweizer, 2002) can be derived.

Schweizer, McCammon, and Jamieson (2006) suggest that threshold sum (TSA - corresponding to the release element layering), stability test score (corresponding to failure initiation) and stability test release type (corresponding to fracture propagation) are three variables which can be used as predictors of snow slope stability.

Moner et al (2008) have applayed the threshold sum approach (TSA) to snowcover

Corresponding author: Chiambretti Igor, AINEVA, Vicolo dell'Adige 18, 38122 TRENTO (TN), Italy;
tel: +39 0461266427; fax: +39 0461232225;
email: igor.chiambretti@aineva.it

types and the weak layers more frequent on the Pyrenees slopes. They've considered also as unstable layers the ones formed by precipitation grains (PP and DF).

Monti (2008) and Monti et al (2009) applayed the TSA also onto the Italian Alps.

Monti et al (2012) have proposed an approach of the threshold sum approach linked to the type of layers rather than to the surfaces of separation, creating new ways of graphical representation.

Monti and Schweizer (this issue) further refined the TSA by transforming each variable in a dimensionless quantity, standardized within the single snow profile (relative threshold sum approach - RTA). Such approach allows, considering relative differences and values, to better identify the location of potentially more unstable layers with less errors due to measures and their subjectivity.

2 DATA

In the present work, we used two different data sets, collected by several AINEVA's regional offices, each composed of snow profiles accompanied by side-by-side stability tests (ECT and/or RB). Profiles were done by several snow observers during their daily surveys for the forecasting offices in Aosta Valley Independent Region, Lombardia Region, Bolzano Independent Province and Friuli Venezia Giulia Independent Region (Italian Alps). Unfortunately the data sets lack indication whether the location of each profile was onto skier tested slopes (no avalanche released) or onto slopes where a recent avalanche occurred. Overall, the first data set included 652 side-by-side stability tests (ECT and RB). The data were collected during the period 2010-2013. Unfortunately, not all the samples show complete structural information and some show lack of details and accuracy.

The second data set is made by 40 snow profiles each one including side-by-side stability tests (ECT and RB), as for the first one also for this group lacks indication whether the location of each profile was onto stable or unstable snowcover. Onto this data set were made the analysis of TSA (following Moner et al., 2008) and RTA (following Monti et al. 2012).

3 METHODS

Standard methods were applied for snowpack observations (e.g., Cagnati, 2003; CAA, 2002; Greene, 2004). The elevations at the profile site range from 1550 m to 3490 m a.s.l. with a median elevation of 2447 m a.s.l.for the first data set and from 1600 m to 2300 m a.s.l. with a median elevation of 2200 m a.s.l.for the second one. Profiles were performed both on shady slopes (NW, NNW and N) and sunny ones (E, ESE, SE, SSE, S) for the first data set - (Fig. 1) where more frequently poor snow stability can be found and a large part of the avalanche accidents occur (see Valt and Pivot, this issue), and on shady slopes (NW, NNW and N) for the second one. Always, stability tests (RB and ECT) were performed along with a snow profile.

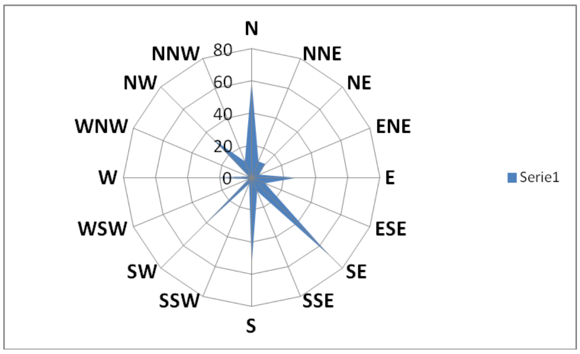

Figure 1. Slope aspect profile site distribution for the first data set.

The rutschblock test (RB – Föhn, 1987) is performed onto an isolated block of snow (2.0 m cross-slope x 1.5 m upslope) loaded in stages by a skier until eventually a weak layer failure. Test score or loading step (#RB from 1 to 7) is recorded as well as the release type: whole block - W, part of the block - P, edge only - E (*sensu* Schweizer, 2002). The extended column test (ECT – Simenhois and Birkeland, 2006) is performed onto an isolated block of snow (90 cm cross-slope x 30 cm upslope) loaded on one corner, in stages, by tapping onto a snow shovel until eventually a fracture initiates and propagates through the rest of the column. Test score or loading step (#ECT from 1 to 31) is recorded as well as the release type: fracture propagates across the entire column during isolation – V; propagation of fracture across the entire column at tap # or #+1 – P; fracture observed at # tap but does not propagate across the entire column at tap # or #+1 – N; no fracture observed during the test – X (*sensu* CAA, 2002) . For each stability test was recorded also the shear quality (Q1, Q2, Q3 – Johnson and Birkeland, 1998) and the fracture character: sudden collapse – SC; sudden planar – SP; progressive compression – PC; resistant planar – RP; non-planar break – B; no fracture – X (Jamieson, 1999; van Herwijnen and Jamieson, 2002, 2004).

For the first data set, performance of predictors was evaluated through various categorical statistics scores (Wilks, 1995; Jamieson, Schweizer, Haegeli, and Campbell, 2006). As it was not possible evaluate the performance of each snow stability test comparing the predicted stability (by the test) with the observed stability (avalanche activity or ski tested slope), we confronted the relative performance of the two test in term of similar results. Test scores (#) were subdivided into stable and unstable as follow (following Winkler and Schweizer, 2008; Schweizer and Jamieson 2010):

Test type and score	rather Unstable	rather Stable
RB#	≤ 3	> 4
ECT#	≤12	> 13

Table 1. Classification of test scores (#) into unstable or stable ongoings.

Then each test type was compared with its release type and test scores (#) were subdivided into stable and unstable (Tab 2a) and ECT was finally compared with its fracture character and test scores (#) were subdivided into stable and unstable as follow (Tab 2b):

Test type and score Tab. 2a	Release type	
	rather Unstable	rather Stable
RB#	W	P; E; X
ECT#	V; P	N; X
Test type and score Tab. 2b	**Fracture character**	
	rather Unstable	rather Stable
ECT#	SC; SP	PC; RP; B

Table 2a, 2b. Classification of test scores (#) into unstable or stable ongoings following release type or fracture character.

For the second data set, the RTA index issues from the sum of 6 related variables for each layer (grain size, difference in grain size, difference in hardness, layer hardness, grain shape, failure layer depth) derived from TSA. The relative value for each variable is the measured value for that layer minus the mean value along the profile divided its standard deviation. This relative value is then scaled to an index in the range between 0 and 1 and potentially unstable layers shows a value of 1 or greater than a threshold (0.95, 0.90, etc.) which can be fixed following the local conditions of the snowpack. Decreasing the threshold, increases the number of layers considered unstable in a profile.

4 RESULTS

4.1 Results of the first data set

For scores ≤ 3, RB tests show almost an equal subdivision between the three types of release: whole block - W, part of the block - P, edge only – E. For scores = 5 or 6, RB tests show a slight prevalence of edge only – E release type over part of the block – P, very few test recorded whole block – W type. For scores = 7, RB tests show, as obvious, only absence of fracture – X (Fig. 2).

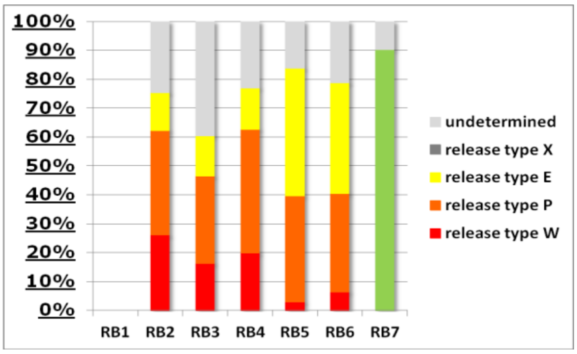

Figure 2. RB score vs release type

For scores = 0, ECT tests show almost all the release type: fracture propagates across the entire column during isolation – V and very few cases of fracture was observed but does not propagate across the entire column. For scores = 1 to 6, ECT tests show a strong prevalence of the release type: propagation of fracture across the entire column at tap # or #+1 – P; and in minor number the type: fracture ECT observed at # tap but does not propagate across the entire column at tap # or #+1 – N. For scores = 7 to 24, ECT tests show the two release types: propagation of fracture across the entire column at tap # or #+1 – P; and fracture observed at # tap but does not propagate across the entire column at tap # or #+1 – N. For scores = 25 to 30, tests show a prevalence of the release type: fracture observed at # tap but does not propagate across the entire column at tap # or #+1 – N; and slightly less of the release type: propagation of fracture across the entire column at tap # or #+1 – P. For score = 31, ECT tests show, as obvious, only absence of fracture – X (Fig. 3).

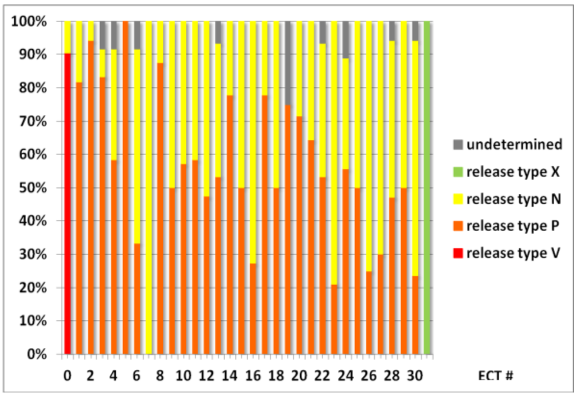

Figure 3. ECT score vs release type.

Plotting ECT test scores and depth of the failure layer shows that the fracture propagates across the full column during isolation (V release type) if the weak layer is < 50 cm from the snow cover surface. Usually, the fracture propagates across the full column at the same tap (#) or one additional (#+1) tap as initiation (P release type) if the weak layer is < 70 cm from the snow cover surface. Nucleation of fracture but absence of propagation at the same tap (#) or one additional (#+1) tap as initiation (N release type) does not show any trend related to the depth of the weak layer (Fig. 4).

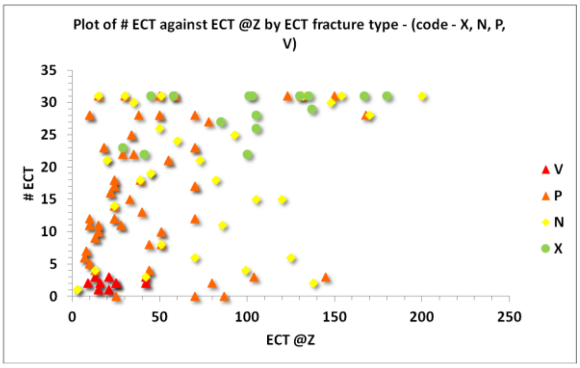

Figure 4. ECT score vs depth of the weak layer by fracture type.

Following the methods of Wilks (1995) we evaluated the relative performance of the tests applying the definitions used in contingency tables (Tab. 2), and calculated using the formulas displayed in Tab. 3.

<table>
<tr><td colspan="2" rowspan="2">N=a+b+c+d=652</td><td colspan="2">RB
Rutschblock Test</td></tr>
<tr><td>Stable</td><td>Unstable</td></tr>
<tr><td rowspan="6">ECT
Extended Column Test</td><td rowspan="3">Stable</td><td>a - Correct stable</td><td>b - Misses (false stable)</td></tr>
<tr><td>422</td><td>79</td></tr>
<tr><td>64,72%</td><td>12,12%</td></tr>
<tr><td rowspan="3">Unstable</td><td>c – False alarms (false unstable)</td><td>d – Hits (correct unstable)</td></tr>
<tr><td>75</td><td>76</td></tr>
<tr><td>11,50%</td><td>11,66%</td></tr>
</table>

Table 3. Contingency table comparing ECT results to RB results for adjacent tests.

The probability of correct detection (PCD - also known as overall accuracy) for both tests is quite high (0,76) but as the dataset is unbalanced (422 stable / 76 unstable) the unweighted

Table 3. Contingency analysis comparing ECT results to RB results for adjacent tests.

	EQUATION	RB vs ECT	RB vs release type	ECT vs release type	Ect vs fracture type
Probability of correct detection	$\text{PCD} = \frac{a+d}{N}$	0,76	0,81	0,78	0,72
Unweighted average accuracy	UAA $= 0,5 \times \left[\left(\frac{a}{a+c}\right) + \left(\frac{d}{b+d}\right)\right]$	0,67	0,68	0,71	0,50
Sensitivity	$\text{POD} = \frac{d}{b+d}$	0,49	0,52	0,52	0,50
False alarm rate	$\text{FAR} = \frac{c}{a+c}$	0,15	0,16	0,10	0,20
True skill score	$\text{TSS} = POD - FAR$	0,34	0,36	0,42	0,30
Specificity	$\text{PON} = \frac{a}{a+c}$	0,85	0,84	0,90	0,80
Critical success index	$\text{CSI} = \frac{a}{a+b+c}$	0,73	0,79	0,74	0,82
Bias	$\text{B} = \frac{a+b}{a+c}$	1,01	0,91	1,12	0,98
Heidke and Kuipers skill score	KSS $= \frac{\left(\left(\frac{a+d}{N}\right)-\left(\frac{(a+b)(a+c)+(b+d)(c+d)}{N^2}\right)\right)}{\left[1-\left(\frac{(a+c)^2+(b+d)^2}{N^2}\right)\right]}$	0,75	0,81	0,62	0,72
Odds ratio	$\Theta = \frac{a\times d}{b\times c}$	5,63	5,44	9,62	4,10
False alarm ratio	$\text{FAR} = \frac{b}{a+b}$	0,16	0,07	0,20	0,18
Probability of false detection	$\text{POFD} = \frac{b}{b+d}$	0,51	0,48	0,48	0,50
Heicke skill score	HSS $= \frac{[2\times((a\times d)-(b\times c))]}{[((a+c)\times(c+d)+(a+b)\times(b+d))]}$	0,34	0,28	0,46	0,30
Peirce skill score	$\text{PSS} = \frac{(a\times d)-(b\times c)}{(a+c)\times(b+d)}$	0,34	0,36	0,42	0,30
Clayton skill score	$\text{CSS} = \frac{(a\times d)-(b\times c)}{(a+b)\times(c+d)}$	0,35	0,24	0,51	0,29
Gilbert skill score	$\text{GSS} = \frac{a-\left(\frac{(a+b)(a+c)}{N}\right)}{a-\left(\frac{(a+b)(a+c)}{N}\right)+b+c}$	0,21	0,16	0,29	0,18
Odds ratio skill score	$\text{Q} = \frac{\theta-1}{\theta+1} = \frac{ad-bc}{ad+bc}$	0,70	0,69	0,80	0,61

average accuracy (UAA = 0,67) gives a far unbiased estimation.

The probability of correct detection (PCD) is slightly better, considering the release type, for RB (0,81) than for ECT (0,78) but the unweighted average accuracy (UAA) shows a different ratio (RB = 0,68 ; ECT = 0,71).

The probability of correct detection (PCD) for ECT test, considering the fracture type, is 0,72 and the unweighted average accuracy (UAA) is 0,50.

The probability of a false alarm (POFD) is a medium value (0,51) for both tests, slightly lowerconsidering the release type, both 0,48 for RB and for ECT.

The sensitivity (probability of detection – POD) for both tests is medium (0,49) and slightly above medium value (0,52) for both tests considering the release type, whereas the specificity is quite high (probability of null events – PON – 0,85) for both tests, slightly higher considering the release type (RB=0,84 ; ECT=0,90) and high (0,80) considering the fracture type of ECT.

4.2 Results of the second data set

The first processing performed on the second data set was the TSA to verify whether the identified weak layers were in relation with the performed stability tests or not (ECT). The first results were not very encouraging both on the detected weak layers (POD =0,32 over 38 layers) and on fracture propagation (POD =0,36 over 88 layers).

The second processing performed on the same data set was the RTA, with slightly better results: POD(1)=0,42; POD (0,95)=0,47; POD (0,90)=0,48; POD(0,80)=0,69 for the 38 weak layers detected and POD(1)=0,46; POD(0,95)=0,50; POD(0,90)=0,50; POD (0,80)=0,75 for the 88 layer with facture propagation (Fig.5).

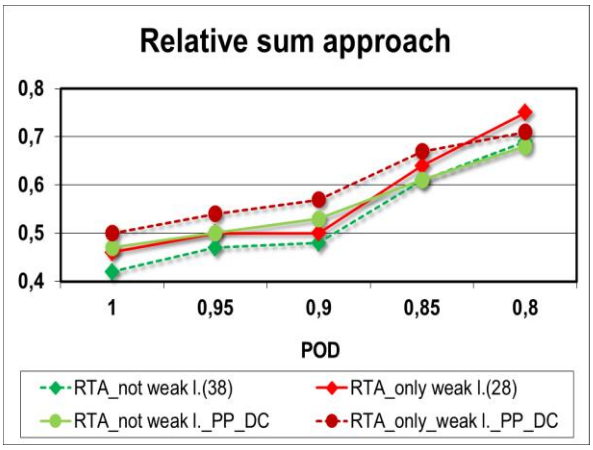

Figure 5. relative threshold sum approach POD

5 CONCLUSIONS

Despite its numbers, the population sample analyzed for the first data set is not yet large enough (in space and time) to adequately represent the complexity of Italy's areas.

The analysis of the relationship between the two stability tests is complicated by a greater snow cover variability (compared to other countries) and by the need for some technicians to achieve greater accuracy in test execution and recording.

However, these preliminary data suggests that ECT test could become, after a testing period and an improvement of the surveyor's technical skills of execution, an excellent and fast using aid tool for forecasters as good as RB test.

The analysis performed onto the second data base showed better results with RTA compared to the TSA. RTA can be a good way to eliminate the measurement errors related to subjectivities of observers or subtle differencies in the methodologies adopted by the regional forecasting offices.

The processing of RTA and TSA according to Moner et al (2008), for the data set of the Italian Alps, have improved the performance of the two methods. Such results indicates that specific TSA and RTA should be set for each climatic area (Southern Alps vs. northern Alps, Pyrenees, Ural, etc. ..).

6 REFERENCES

CAA, 2002, Observation guidelines and recording standards for weather, snowpack and avalanches: Canadian Avalanche Association (CAA), Revelstoke BC, Canada, 78 pp.

Cagnati, A., 2003, Strumenti di misura e metodi di osservazionenivometeorologici – manuale per i rilevatori dei Servizi di previsione valanghe: AINEVA, Rovereto (TN), Italy, 133 pp.

Föhn, P.M.B., 1987, The Rutschblock as a practical tool for slope stability evaluation: Symposium at Davos 1986 - Avalanche Formation, Movement and Effects, IAHS Publ., 162, International Association of Hydrological Sciences, Wallingford, Oxfordshire, U.K.: 223-228.

Greene, E. (Editor), 2004, Snow, weather and avalanches - Observational guidelines for avalanche programs in the United States: American Avalanche Association (AAA), Pagosa Springs CO, U.S.A., 136 pp.

Jamieson, J.B., 1999, The compression test - after 25 years: The Avalanche Review, 18(1): 10-12.

Jamieson, J.B. and Schweizer, J., 2005, Using a checklist to assess manual snow profiles: Avalanche News, 72: 57-61.

Jamieson, J.B., Schweizer, J., Haegli, P., and Campbell, A., 2006, Can stability tests help recreationists assess the local avalanche danger?: Proceedings ISSW 2006, International Snow Science Workshop, Telluride CO, U.S.A., 1-6 October 2006: pag. 1-10.

Johnson, J.B., and Birkeland, K.W., 1998, Effectively using and interpreting stability tests: International Snow Science Workshop, Sunriver, OR, U.S.A., 5 pp.

McCammon, I., and Sharaf, 2005, Integrating strength, energy, and structure into stability decisions: The Avalanche Review, 23(3), february, pag. 18-19.

McCammon, I. and Schweizer, J., 2002, A field method for identifying structural weaknesses in the snowpack. In: J.R. Stevens (Editor), Proceedings ISSW 2002. International Snow Science Workshop, Penticton BC, Canada, 29 September-4 October 2002: 477-481.

Monti, F., 2008, Modello SNOWPACK - verifica in area Dolomitica e applicazione al sistema di previsione valanghe: BSc Thesis, Università dell'Insubria, Como, 226 pp.

Monti, F., Cagnati, A., Fierz, C., Lehning, M., Valt, M., Pozzi, A., 2009, Validation of the SNOWPACK model in the Dolomites. In: Schweizer, J., van Herwijnen, A. (Eds.), International Snow Science Workshop ISSW, Davos, Switzerland, 27th September–2nd October 2009. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, pp. 313–317.

Monti, F, Cagnati, A., Valt, M., Schweizer, J,. 2012, A new method for visualizing snow stabilty profile: Cold Regions Science and Technology, Volume 78, pag. 64-72.

Schweizer, J., 2002, The Rutschblock test – procedure and application in Switzerland: The Avalanche Review, 20(5): 1, pag. 14-15.

Schweizer, J., and Jamieson, J.B., 2007, A threshold sum approach to stability evaluation of manual snow profiles: Cold Reg. Sci. Technol., 47(1-2), pag. 50-59.

Schweizer, J. and Jamieson, J.B., 2010, Snowpack tests for assesing snow-slope instability: Annals of Glaciology, 51(54), pag. 187-194.

Schweizer, J., McCammon, I., and Jamieson, J.B., 2006, Snow slope stability evaluation using concepts of fracture mechanics: Proceedings ISSW 2006, International Snow Science Workshop, Telluride CO, U.S.A., 1-6 October 2006: pag. ??-??.

Schweizer, J., McCammon, I., and Jamieson, J.B., 2008, Snowpack observations and fracture concepts for skier-triggering of dry-snow slab avalanches: Cold Reg. Sci. Technol., 51(2-3), pag. 112-121.

Simenhois, R. and Birkeland K.W., 2006, The extended column test - A field test for fracture initiation and propagation. In: Gleason J.A. (Editor), Proceedings ISSW 2006, International Snow Science Workshop, Telluride CO, U.S.A., 1-6 October 2006: pag. 79-85.

Simenhois, R. and Birkeland K.W., 2007, An upgrade on the extended column test - New recording standards and additional data analyses: The Avalanche Review, 26(2).

van Herwijnen, A. and Jamieson B., 2002, Interpreting fracture character in stability tests: International Snow Science Workshop, Penticton, B.C., Canada, 7 pp..

van Herwijnen, A. and Jamieson B., 2004, More results on fracture characterization in compression tests: Avalanche News, 68, pag. 38-41.

van Herwijnen, A. and Jamieson B., 2007, Fracture character in compression tests: Cold Reg. Sci. Technol. 47(1-2), pag. 60-68.

Winkler, K., and Schweizer, J., 2008, Comparison of different snow stability tests including the extended column test: Proceedings ISSW 2008, International Snow Science Workshop, Whistler, B.C., Canada, 21-27 September 2008, pag. 393-400.

Wilks, D.S., 1995, Statistical methods in the atmospheric sciences: an introduction. Vol. 59, International Geophysics, Academic Press, San Diego CA, U.S.A, 467 pp.