COMPARISON OF ECT, RB, THRESHOLD SUM AND RELATIVE THRESHOLD SUM APPROACH TO TO SNOWPACK INSTABILITY EVALUATION - NEW RESULTS

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

ABSTRACT: the Italian regional avalanche forecasting services (AINEVA) collect snowpack observations such as snowpack profiles or stability tests on a daily or weekly base both itinerant or at fixed study plots. Snow stratigraphy data need to be interpreted by the avalanche forecasters in order to estimate the actual snow stability. Reducing the interpretation subjectivity is the main goal of both the Threshold Sum (TSA) or Relative Threshold Sum (RTA) approaches.
During the period 2010-2015 the AINEVA snow observers collected a dataset consisting of more than 800 snow profiles coupled with stability tests (ECT and/or RB).
In this work, the snow stability forecasts obtained by snow stratigraphy analyses and snow stability tests are compared integrating the results of previous studies. Moreover, side-by-side ECT and RB tests have been compared in order to asses the ECT versus RB effectiveness for discriminating the main weak layers, their propensity to crack initiation and propagation and finding relations, through the integration with data from snowpack profiles analysis, between load steps, fracture character and depth, quality shear and weak layer characteristics. The analysis made it possible to better characterize the most common weak layers of snow cover on the Italian Alps in terms of thickness, grain type, depth and mechanical response to stability tests.

KEYWORDS: RB, ECT, Structural instability indices, TSA, RTA

1. INTRODUCTION

The snow cover, usually, exhibits a complex layered structure as each snowfall accumulates through bonding a layer onto the previous snow cover surface. Its observation allows to depict such structure and to follow, during the winter season, its evolution as metamorphic processes change, through time, the characteristics of those layers often developing new layers and type of grains. Each layer shows distinct mechanical properties which control failure initiation (crack nucleation and growth) inside the snowpack and fracture propagation (possible avalanche release) along the weakest layer or interface between layers when an additional stress, locally, overcome the strength. Further 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, and might develop into a catastrophic failure (avalanche) when the fracture toughness is overcome.

* *Corresponding author address:*
Chiambretti Igor, AINEVA,
Trento (TN), Vicolo dell'Adige 18, Italy 38122;
tel: +39 0461266427; fax: +39 0461230305;
email: igor.chiambretti@aineva.it

The execution of a snow profile and associated stability tests (rutschblock – RB - Föhn, 1987 and extended column test – ECT - Simenhois and Birkeland, 2006) allows to record the following parameters: snowpack structure (layering), fracture initiation (strength) and fracture propagation (toughness) - McCammon and Sharaf (2005). The potential weak layers are located and described (position, grain type, grain size, hardness); the stability test identifies the failure layers (test score, fracture character or release type or shear quality – sensu van Herwijnen and Jamieson, 2004) and structural instability indices based on threshold sums such as the lemons or yellow flags (Jamieson and Schweizer, 2005; McCammon and Schweizer, 2002) which can be derived from such observations.
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.(Schweizer, McCammon, and Jamieson, 2006).
Moner et al (2008), Monti (2008), Monti et al (2009) applayed the threshold sum approach (TSA) to several snowcover types and to the most frequent weak layers on the Pyrenees and Italian

Alps slopes. They considered also as unstable layers the ones formed by precipitation grains (PP and DF).

Monti et al (2012) and Monti and Schweizer (2013). further refined the TSA approach linking it to the type of layers rather than to the surfaces of separation and 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 most 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, Piedmont 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 806 side-by-side stability tests (ECT and RB). The data were collected during the period 2010-2015. Unfortunately, not all the samples show complete structured information and some samples 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), but also 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 and 2014; Greene, 2004; Green et al 2010). The elevations at the profile site range from 1530 m to 3490 m ASL with a median elevation of 2488 m ASL for the first data set and from 1600 m to 2300 m ASL with a median elevation of 2200 m ASL for the second one. For the first data set, profiles were performed both on shady slopes (NW, NE and N) and sunny ones (E, SE, S, SW, W) where more frequently poor snow stability can be found and a large part of the avalanche accidents occur (see Valt and Pivot, 2013). For the second one, profiles were performed only on shady slopes (NW, NNW and N) and were accompanied by stability tests (RB and ECT).

The rutschblock test (RB – Föhn, 1987) was performed onto an isolated block of snow (2.0 m cross-slope x 1.5 m upslope) and test score or loading step (#RB from 1 to 7) was recorded as well as the release type: whole block - W, part of the block - P, edge only - E (sensu Schweizer, 2002). Close to the RB, the extended column test (ECT – Simenhois and Birkeland, 2006) was performed onto an isolated block of snow (90 cm cross-slope x 30 cm upslope) and test score or loading step (#ECT from 1 to 31) was 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 and 2014). 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).

Various categorical statistics scores were applied, to the first data set, to evaluate the performance of predictors (Wilks, 1995; Jamieson, Schweizer, Haegeli, and Campbell, 2006) by confronting the relative performance of the two test in term of similar results. Due to the variability of the first data set, in fact, it was not possible to 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).

Test scores (#) were subdivided into stable and unstable as follow (following Winkler and Schweizer, 2008; Schweizer and Jamieson, 2010; Chiambretti, Monti and Valt, 2013):

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 show 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, the number of layers considered unstable in a profile increases.

4. RESULTS

4.1 Results of the first dataset

For scores ≤ 4, 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 almost an equal subdivision between edge only – E release type and 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. 1). Note that RB scores = 1 are absent from this data set.

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 - N.

For scores = 1 to 4, 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 – (Fig. 2).

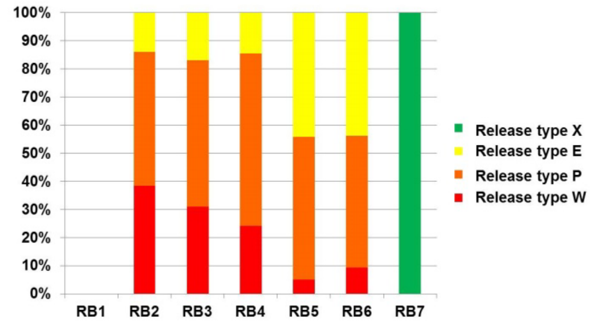

Figure 1. RB score vs release type

For scores = 5, ECT tests show an unusual strong prevalence of the release type: propagation of fracture across the entire column at tap # or #+1 – P but located at 35 cm of depth (in average). For scores = 6 to 18, 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 (Fig. 2).

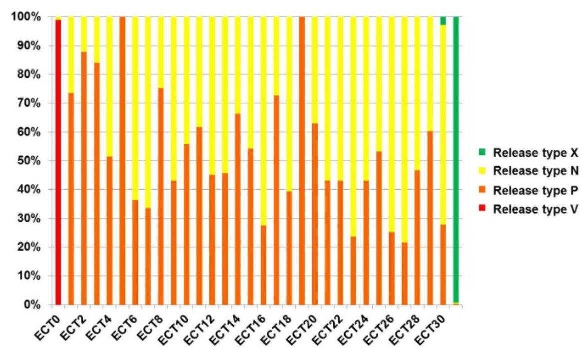

Figure 2. ECT score vs release type.

For scores = 19, ECT tests show an unusual strong prevalence of the release type: propagation of fracture across the entire column at tap # or #+1 – P but located at 57 cm of depth (in average). For scores = 20 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. 2).

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) or there is nucleation of fracture but absence of propagation at the same tap (#) or one additional (#+1) tap as initiation (N release type) if the weak layer is < 70 - 80 cm from the snow cover surface (Fig. 3).

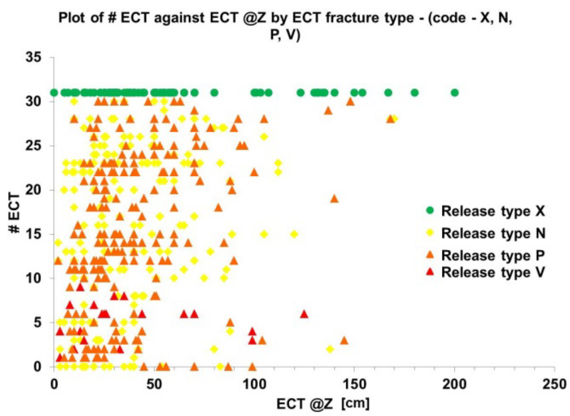

Figure 3. ECT score vs depth of the weak layer by fracture type

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.

N=a+b+c+d=**806**		**RB** Rutschblock Test	
		Stable	Unstable
ECT Extended Column Test	Stable	**a** - Correct stable	**b** - Misses (false stable)
		517	**93**
		64,14%	11,54%
	Unstable	**c** – False alarms (false unstable)	**d** – Hits (correct unstable)
		102	**94**
		12,66%	11,66%

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 – 0,75) but as the dataset is unbalanced (517 stable / 94 unstable) the unweighted average accuracy (UAA = 0,67 – 0,69) gives a far unbiased estimation.

	EQUATION	RB vs ECT	RB vs release type	ECT vs release type	Ect vs fracture type
Probability of correct detection	PCD $= \frac{a+d}{N}$	0,76	0,81	0,75	0.67
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	0.69	0.50
Sensitivity	POD $= \frac{d}{b+d}$	0.50	0.51	0.50	0.46
False alarm rate	FAR $= \frac{c}{a+c}$	0.16	0.15	0.11	0.22
True skill score	TSS $= POD - FAR$	0,34	0,36	0.39	0.24
Specificity	PON $= \frac{a}{a+c}$	0,85	0,84	0.89	0.78
Critical success index	CSI $= \frac{a}{a+b+c}$	0,73	0,79	0.69	0,82
Bias	B $= \frac{a+b}{a+c}$	0.99	0.93	1.19	1.07
Heidke and Kuipers skill score	KSS $= \frac{\left(\left(\frac{a+d}{N}\right) - \left(\frac{(a+b)(a+c)+(b+d)(c+d)}{N^2}\right)\right)}{\left[1 - \left(\frac{((a+c)^2)+((b+d)^2)}{N^2}\right)\right]}$	0.73	0,81	0.75	0.67
Odds ratio	$\Theta = \frac{a \times d}{b \times c}$	5.07	5.44	9.62	3.53
False alarm ratio	FAR $= \frac{b}{a+b}$	0.15	0.08	0.25	0.27
Probability of false detection	POFD $= \frac{b}{b+d}$	0.50	0.49	0.50	0.54
Heicke skill score	HSS $= \frac{[2 \times ((a \times d) - (b \times c))]}{[(a+c) \times (c+d) + (a+b) \times (b+d)]}$	0.33	0.31	0.42	0.25
Peirce skill score	PSS $= \frac{(a \times d) - (b \times c)}{(a+c) \times (b+d)}$	0,34	0,36	0.39	0.24
Clayton skill score	CSS $= \frac{(a \times d) - (b \times c)}{(a+b) \times (c+d)}$	0.33	0.27	0.48	0.26
Gilbert skill score	GSS $= \frac{a - \left(\frac{(a+b)(a+c)}{N}\right)}{a - \left(\frac{(a+b)(a+c)}{N}\right) + b + c}$	0.20	0.18	0.26	0.14
Odds ratio skill score	Q $= \frac{\theta - 1}{\theta + 1} = \frac{ad - bc}{ad + bc}$	0.67	0.71	0.79	0.56

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

The probability of correct detection (PCD) is slightly better, considering the release type, for RB (0,81) than for ECT (0,75) but the unweighted average accuracy (UAA) shows a different ratio (RB = 0,68; ECT = 0,69).
The probability of correct detection (PCD) for ECT test, considering the fracture type, is 0,67 and the unweighted average accuracy (UAA) is 0,50.
The probability of a false alarm (POFD) is a medium value (0,50) for both tests, slightly lower considering the release type, 0,49 for RB.
The sensitivity (probability of detection – POD) for both tests is medium (0,50) and slightly above medium value (0,51) for RB tests considering the release type, whereas the specificity is quite high (probability of null events – PON – 0,84) for both tests, slightly higher considering the release type (RB=0,85; ECT=0,89) and high (0,78) considering the fracture type of ECT.

4.2 Results of the second dataset

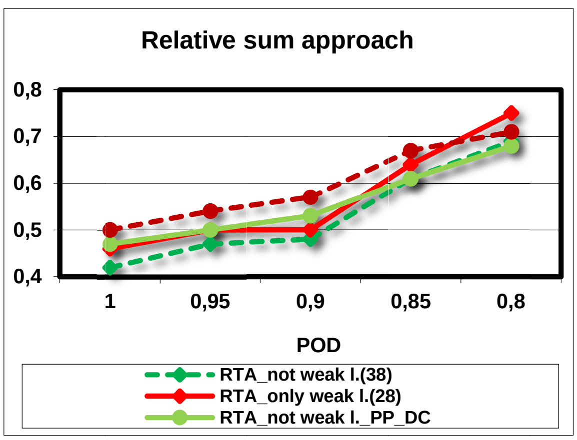

Figure 4. relative threshold sum approach POD

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.4).

5. CONCLUSIONS

The first data set analyzed has allowed a better representation (in space and time) of Italy's snowpack mechanical characteristics although it is not yet fully satisfactory.

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 data suggests that ECT test is already an excellent and fast using aid tool for forecasters as good as RB test provided that at least two ECTs are executed side by side.

The analysis performed onto the second data base shows 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 data set of the Italian Alps, have improved the performance of the two methods. Such results indicate 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.

CAA, 2014, Observational Guidelines and Recording Standards for Weather, Snowpack and Avalanches: Canadian Avalanche Association (CAA), Revelstoke BC, 94 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.

Greene, E., Atkins, D., Birkeland, K., Elder, K., Landry, C., Lazar, B., McCammon, I., Moore, M., Sharaf, D., Sternenz, C., Tremper, B., and Williams, K., 2010. Snow, Weather and Avalanches - Observation Guidelines for Avalanche Programs in the United States: American Avalanche Association, Pagosa Springs, CO, Second Printing Fall 2010.

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.

Monti, F., and Schweizer, J., 2013, A relative difference approach to detect potential weak layers within a snow profile; International Snow Science Workshop ISSW, Grenoble, France, 7th – 11th October 2013, P1-30, pp. 339–343.

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.