Natural Hazards
and Earth System
Sciences

Open Access

A new index combining weak layer and slab properties for snow instability prediction

J. Schweizer and B. Reuter

WSL Institute for Snow and Avalanche Research SLF, Flüelastrasse 11, 7260 Davos Dorf, Switzerland

Correspondence to: J. Schweizer (schweizer@slf.ch)

Abstract. Snow slope stability evaluation requires considering weak layer as well as slab properties – and in particular their interaction. We developed a stability index from snow micro-penetrometer (SMP) measurements and compared it to 129 concurrent point observations with the compression test (CT). The index considers the SMP-derived micro-structural strength and the additional load, which depends on the hardness of the surface layers. The new quantitative measure of stability discriminated well between point observations rated as either "poor" or "fair" (CT < 19) and those rated as "good" (CT ≥ 19). However, discrimination power within the intermediate range was low. We then applied the index to gridded snow micro-penetrometer measurements from 11 snow slopes to explore the spatial structure and possibly relate it to slope stability. Stability index distributions on the 11 slopes reflected various possible strength and load (stress) distributions that can naturally occur. Their relation to slope stability was poor, possibly because the index does not consider crack propagation. Hence, the relation between spatial patterns of point stability and slope stability remains elusive. Whereas this is the first attempt of a truly quantitative measure of stability, future developments should consider a better reference of stability and incorporate a measure of crack propagation.

1 Introduction

Snow stability data are among the key ingredients when establishing avalanche forecasts. Snow stability can either be assessed from observations of instability such as recent avalanching (Jamieson et al., 2009), by stability tests performed in the field (e.g., Schweizer and Jamieson, 2010) or

from stability indices derived from modeled snow stratigraphy (e.g., Durand et al., 1999; Schweizer et al., 2006). Whereas numerical modeling allows obtaining data with a high temporal and spatial resolution – though often of unknown accuracy – field tests are laborious and reveal partly subjective point information of low temporal and spatial resolution. Nevertheless, stability tests are presently the method of choice for estimating snow instability – despite the fact that the inherently variable nature of the mountain snowpack hinders extrapolation of point observations (Schweizer et al., 2008). One way to overcome the limitation of point observations is to perform many measurements in a given area within a couple of hours. This approach is only possible with a quick probing method, for example, with the snow micro-penetrometer (Schneebeli and Johnson, 1998) or potentially with remote sensing techniques.

The snow micro-penetrometer (SMP) is a probe with a high-resolution force sensor at its tip, driven into the snowpack at constant speed. It provides a penetration resistance (force)-depth signal that includes micro-structural information (Johnson and Schneebeli, 1999; Marshall and Johnson, 2009). Mechanical properties can be derived from the three basic micro-structural parameters: element length (L), deflection at rupture (δ) and rupture force (f). The micro-structural strength (σ_{m}) is assumed to scale as f/L^2.

In one of the first attempts to relate the SMP resistance to stability, Kronholm (2004) found increasing stability scores with increasing weak layer penetration resistance for four out of five investigated weak layers. Based on the characteristics found in manually observed snow profiles (Schweizer and Jamieson, 2007), Pielmeier and Schweizer (2007) tried to discriminate between unstable and stable observations based on SMP-derived characteristics of the weak layer and the ad-

jacent layers. Pielmeier and Marshall (2009) refined this approach and showed that the micro-structural strength of the weak layer (manually identified in the SMP profile) was the single best classifier to discriminate between unstable and stable Rutschblock test results. Classification accuracy improved to about 85 % when SMP-derived mean slab density (Pielmeier, 2003) was included in a 2-node classification tree. They pointed out the importance of signal quality control and showed the improvement in classification accuracy that can be obtained when several SMP measurements within an area of a few m² are performed. Lutz et al. (2009) and Bellaire and Schweizer (2011) also found the micro-structural strength of the weak layer to be related to stability. Floyer and Jamieson (2009) predicted the fracture character of compression tests (CT) from adjacent SABRE penetrometer profiles (Mackenzie and Payten, 2002), whereas van Herwijnen et al. (2009) found snow stratigraphy derived from microstructural properties of the SMP to be related to the fracture type in CTs.

Direct correlations of point measurements such as SMP penetration hardness to stability have proven challenging (Lutz et al., 2009). Previous studies of many point measurements at the slope scale using the SMP revealed, among other things, that typical weak layers are often continuously present but have clearly varying properties (e.g., Kronholm et al., 2004). However, relating spatial variations as derived from point measurements to slope stability has so far not been successful. For example, Bellaire and Schweizer (2011) stated that firm conclusions on the dependence of slope stability on spatial variations were not possible due to the limited range of snow conditions in the data set and the fact that the definition of slope stability is partly intangible. From a theoretical point of view, as supported by numerical modeling (e.g., Fyffe and Zaiser, 2004; Gaume et al., 2014; Kronholm and Birkeland, 2005), it seems clear that stability variations at the slope scale can either promote or hinder slope failure. Slope instability should increase with increasing coefficient of variation and increasing correlation length.

Whereas the above-mentioned studies indicate that considerable progress has been made towards objectively deriving snow stability information from the SMP resistance profile, a single measure of stability, combining slab and weak layer properties, is so far lacking. Also, relating point measurements to slope stability has not been successful. We will present a first attempt to directly derive an index of snow stability from the SMP signal and compare it to results of numerous small-scale stability tests (CT). The index will then be applied to the gridded SMP measurements collected on 11 slopes by Bellaire and Schweizer (2011) to explore the spatial structure of stability on these slopes and possibly relate it to slope stability. The index will at best be an estimate of the probability of initiating a failure in a weak layer, but it will not provide any information on the propensity of crack propagation.

2 Data

We primarily used the data collected and described by Bellaire and Schweizer (2011). They concurrently observed snow stability using the CT (Jamieson, 1999) and measured penetration resistance using the SMP on 15 slopes above Davos, Switzerland during the winters 2006–2007 to 2007–2008. During the winter 2008–2009 Bellaire (2010) sampled another eight slopes. On each of the slopes, one manually observed snow profile, nine pairs of CT and 45 SMP measurements were performed. In addition, other observations relevant for assessing slope stability such as signs of instability, snow surface conditions and ski penetration were recorded. For the analysis described below, we used three different data sets (a, b, c) out of the data from 23 slopes:

a. From the 23 concurrent observations of snow stratigraphy (snow profile) and penetration resistance (SMP) after quality control, 19 cases were retained for calculating the SMP-derived penetration depth ($N = 19$).

b. From the concurrent observations of point stability (CT) and penetration resistance (SMP), we analyzed, after quality control, in total 129 SMP profiles with corresponding CT score ($N = 129$).

c. From the gridded SMP data on 23 slopes, we analyzed 11 slopes in regard to their stability distribution. Most of the remaining data could not be used due to quality issues.

3 Methods

Our assumptions are tied to the CT as we aim at a stability criterion for failure initiation, which can be validated with previously collected field data. The stability index follows a simple strength to additional stress criterion in the weak layer, still accounting for slab layering. In the CT experiment, a snow column is loaded by dropping the hand, the forearm or the arm (Jamieson, 1999). For simplicity we consider a fixed weight corresponding to the weight of a forearm. Due to the impact, the surface layers are compacted, or rather crushed (van Herwijnen and Birkeland, 2014). The stress in the column is related to the braking (or decelerating) distance. The softer the surface layers, the larger the compaction (and also the braking distance) and hence the smaller the stress – and vice versa. We assumed idealized elastic-plastic behavior of the snow so that the initial potential energy (E_{pot}) of the dropping weight was completely dissipated over the braking distance:

$$E_{pot} = E_a = F_u u_{max} - F_u^2 / 2K, \tag{1}$$

where E_a is the dissipated energy equal to the area under the curve in the loading (force-displacement) diagram, F_u is the maximum impact force, u_{max} is the maximum displacement
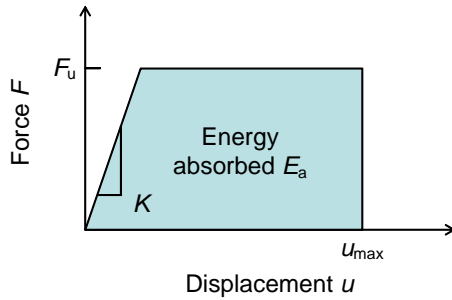
Figure 1. Schematic of elastic-plastic deformation describing the inelastic collision while loading the snow column in a compression test. K denotes the stiffness (equivalent to a spring constant) (after Wright, 2012).

(i.e., the braking distance) and K is the stiffness that in our 1-D approach is equivalent to a spring constant with units $\mathrm{N\,m^{-1}}$ (Fig. 1). As the elastic part of deformation is negligibly small compared to the plastic deformation, the second term in Eq. (1) can be neglected. With $E_{\mathrm{pot}} = \mathrm{mg}\Delta h$, the impact force can then be approximated:

$$F_u \approx \mathrm{mg}\Delta h / u_{\max}. \tag{2}$$

Dividing the impact force by the area of the column A ($0.3\,\mathrm{m} \times 0.3\,\mathrm{m}$) reveals the additional stress: $\Delta\sigma = F_u/A$. For the potential energy $\mathrm{mg}\Delta h$, we assumed a mass of $1.5\,\mathrm{kg}$ dropping from a height of $0.15\,\mathrm{m}$, resulting in an energy of about $2.2\,\mathrm{J}$; this roughly corresponds to the impact by a falling forearm.

   We assumed the braking distance $u_{\max}$ to be related to the penetration depth as measured with a penetrometer. In order to derive the penetration depth from the SMP signal, we cumulated the SMP penetration force over depth to an a priori unknown threshold of dissipated energy (e_{a}). This implies that the area under the penetration force-depth curve corresponds to the dissipated energy e_{a}. Using data set (a) as described above ($N = 19$), including observed penetration depth (PS) (either ram or ski penetration), we determined the dissipated energy up to the depth PS for each SMP profile:

$$e_{\mathrm{a}} = \int_0^{\mathrm{PS}} F \mathrm{d}h, \tag{3}$$

where F is the penetration force and h is the depth from the snow surface. The average energy e_{a}, absorbed up to the penetration depth PS, was $0.0036\,\mathrm{N\,m}$. In the following, we used this threshold value to calculate the SMP-derived penetration depth (or breaking distance). For the 19 cases, the median deviation between observed (PS) and modeled (ps) penetration depth was $1.5\,\mathrm{cm}$, with one outlier of $8.4\,\mathrm{cm}$ (standard error: $2.5\,\mathrm{cm}$) (Fig. 2).

   For calculating the stability index, we assumed that the additional stress (derived from Eq. 2) would not decrease
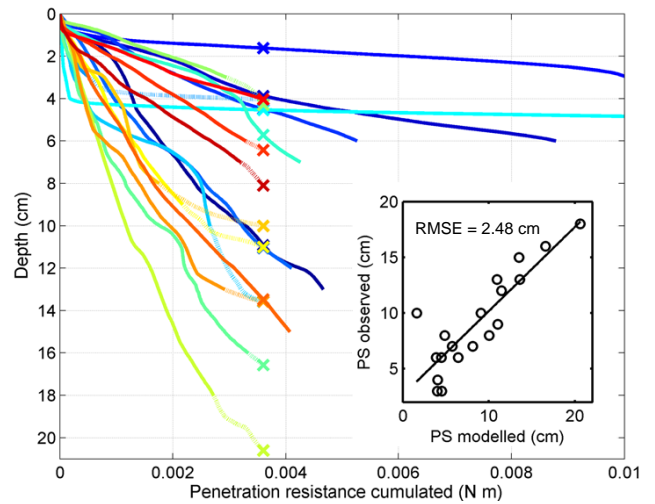

Figure 2. Cumulated penetration resistance ($=$ dissipated energy) vs. depth for 19 SMP penetration resistance-depth signals. Solid lines show the paths of cumulated penetration resistance up to observed penetration depth PS. Crosses denote corresponding depth ($=$ modeled penetration depth ps) for the average value of dissipated energy $e_{\mathrm{a}} = 0.0036\,\mathrm{N\,m}$. If observed PS is lower than modeled ps, a dashed line leads from the end of the solid line to the corresponding cross. Inset shows modeled vs. observed penetration depth.

strongly with depth as the snow column is uniformly loaded at the top. Furthermore, we neglected the weight of the overlying slab (which is e.g., considered in the skier stability index introduced by Föhn, 1987) because we suppose that the dynamic load (rather than the static load) is essential for initiating a failure due to the well-known deformation rate dependence of snow strength (e.g., Reiweger and Schweizer, 2010). Finally, we did not consider the effect of slope angle on either stress or strength as its effect is largely unknown in the case of a CT.

   The simple stability index was defined as

$$S = \frac{\sigma_{\mathrm{m}}}{\Delta\sigma}. \tag{4}$$

Hence we assume that SMP-derived stability S is simply proportional to the micro-structural strength σ_{m} and the SMP-derived penetration depth ps: $S \sim \sigma_{\mathrm{m}}\,\mathrm{ps}$. The above definition of the stability index (Eq. 4) yields values that are higher than (and thus not directly comparable to) the classical stability index where a value less than 1 (to 1.5) indicates instability (Jamieson and Johnston, 1998). Of course, the proposed stability index could easily be scaled so that it becomes comparable, but we prefer to show the actual values.

   To relate continuous ratio data such as the stability index and micro-structural strength, we used the Pearson correlation coefficient r_{p}. For ordinal data, we assessed the correlation with the Spearman rank order coefficient r_{s}, for example, to relate the newly developed stability index to the CT scores.

www.nat-hazards-earth-syst-sci.net/15/109/2015/

Nat. Hazards Earth Syst. Sci., 15, 109–118, 2015

Table 1. Point stability classification based on CT score and CT fracture type.

| CT score	CT fracture character	
	SP, SC	RP, PC, B
≤ 13	poor	fair
14–18	fair	fair
≥ 19	good	good

Table 2. Slope stability classification based on the slope median point stability and signs of instability (recent avalanching, whumpfs of shooting cracks).

Median CT score	Signs of instability	
	1: present	0: absent
≤ 13	POOR	FAIR
> 13	FAIR	GOOD

As suggested by Bellaire and Schweizer (2011), CT scores were classified into three point stability classes: "poor", "fair" and "good" (Table 1). "Poor" stability refers to CT scores ≤ 13, "fair" to 14–18 and "good" to ≥ 19, based on the conversion of CT scores to Rutschblock scores suggested by Schweizer and Jamieson (2003). For the stability class "poor", we also considered the CT fracture type (van Herwijnen and Jamieson, 2007) and required sudden fractures for this class, i.e., either sudden planar (SP) or sudden collapse (SC); if the CT fracture type was either resistant planar (RP), progressive compression (PG) or non-planar break (B), we classified these tests as "fair" despite a low score (≤ 13).

Similarly, all slopes were classified into one of three classes of slope stability "POOR", "FAIR" and "GOOD". The classification considered the presence or absence of signs of instability and the slope median CT score; in contrast to Bellaire and Schweizer (2011), we did not consider the profile classification because it is essentially based on a point observation (Table 2).

The stability distributions were characterized by the median, the interquartile range (IQR) and the quartile coefficient of variation (QCV). When comparing the distributions of the stability index from the three point stability classes, the non-parametric Kruskall–Wallis H test was used; pairwise comparison was performed with the Dwass–Steel–Critchlow–Fligner method (Critchlow and Fligner, 1991). Alternatively, differences between two samples were assessed with the Mann–Whitney U test. A level of significance $p = 0.05$ was chosen to decide whether the observed differences were statistically significant. We determined the split between two categories (e.g., to discriminate between two stability classes) with the classification tree method (Breiman et al., 1998). Classification accuracy was assessed by 10-fold cross validation. To describe the classification ac-
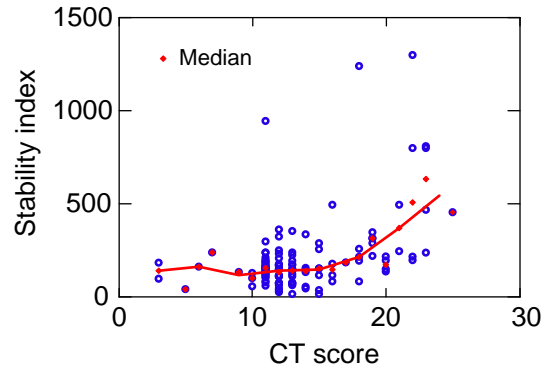
Figure 3. Stability index vs. CT score (eight cases with CT score 35 (no fracture) not shown, $N = 121$); moving average smoothing line between median values.

curacy, the probability of detection (POD), the probability of non-events (PON) and the true skill statistic (TSS) (i.e., the difference between POD and the false alarm rate) were calculated (Wilks, 2011).

To explore the spatial structure, the experimental semivariogram for a linear trend model of the Cartesian coordinates was calculated. By fitting a spherical model to the experimental semivariogram we determined the range, which is a measure of the correlation length. Details are given in Bellaire and Schweizer (2011). For contour plots, data were interpolated by ordinary kriging.

4   Results

The newly developed stability index was calculated for data set (b) of the 129 cases with SMP profile and CT score (Fig. 3). The Spearman rank correlation coefficient between the CT score and stability index S was $r_s = 0.42$ ($p < 0.001$) slightly higher than for the micro-structural strength ($r_s = 0.31$, $p < 0.001$). Correlating the median stability index for each CT score yielded $r_s = 0.77$ ($p < 0.001$).

Grouping the stability index values according to the three classes of point stability indicated that particularly the tests rated as "poor" or "fair" can be discriminated well from those rated as "good" ($p < 0.001$) (Fig. 4). In fact, the H test indicated that differences between all three classes, including between "poor" and "fair", were statistically significant based on the Dwass–Steel–Critchlow–Fligner test statistic ($p < 0.01$). However, a simple U test showed that there is no difference between the samples "poor" and "fair" ($p = 0.44$). If the classes of "poor" and "fair" are grouped, the classification simplifies and becomes comparable to previous studies (e.g., Pielmeier and Marshall, 2009). With a split value of 212, the classification accuracy (10-fold cross-validated) was 81 % ($N = 129$; POD: 78 %, PON: 89 %, TSS: 68 %).
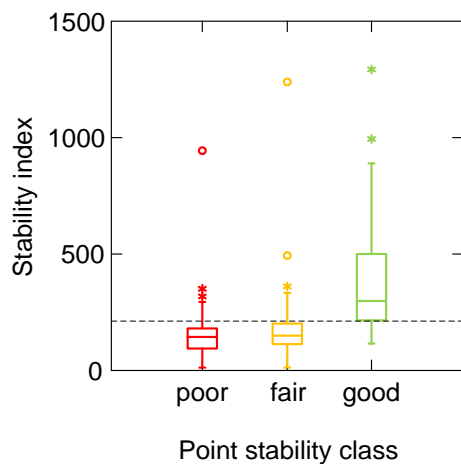
Figure 4. Stability index grouped into the three classes of point stability (based on CT score and fracture type) ($N = 129$; "poor": 70, "fair": 31, "good": 28). Dashed line indicates the split value (212) for the classification into two classes ("poor" and "fair" vs. "good").

4.1 Non-spatial analysis

For data set (c), including the measurements on the 11 slopes, we will first present the result for those locations where concurrent SMP measurements and CTs were performed so that the performance of the new stability index can be directly compared. We then proceed to analyze the slope-scale stability index distributions based on all SMP measurements performed on the 11 slopes.

Figure 5 shows the stability index values for the locations of SMP measurements where a CT was performed concurrently. The stability index distributions found on the 11 slopes were fairly different. The four slopes that were rated "POOR" had low stability indices and a low mean CT score. The two slopes rated as "FAIR" had similarly low stability indices and mean CT scores but were not rated as "POOR" since no signs of instability were observed. Three of the slopes rated as "GOOD" had relatively low stability index values and intermediate CT scores, whereas the other two slopes had high stability index values as well as high CT scores. Overall, per slope, the median stability index was still positively, but not significantly, correlated with the median CT score ($r_s = 0.47$, $p = 0.15$).

The different stability index distributions were the result of various different stress-strength (slab-weak layer) configurations (Fig. 6). For example, grid 0708_9 had rather low strength (53 kPa), but the stability index was relatively high (355) due to the low additional stress (151 Pa). On the other hand, grid 0607_6 had a rather low median value of stability (167) though the weak layer strength was intermediate (95 kPa), but the additional stress was high (561 Pa).

Considering all SMP measurements in the 11 grids (data set c) (Table 3), the grid median stability index tended to in-

crease with increasing median CT score, but the correlation was not significant ($r_s = 0.36$, $p = 0.28$). The stability index was only slightly higher for the slopes rated as "GOOD" (median stability: 145) compared to the slopes rated as either "POOR" or "FAIR" (median stability: 123). Most grids had a median stability index in the range of about 100 to 170, and the slope stability rating was mostly "POOR" or "FAIR" with 3 cases of "GOOD". In the latter three cases no signs of instability were observed, which explains the discrepancy. The two grids with a high median stability index were rated as "GOOD". One of these grids (0708_7) that showed a rather low median stability index (104), but was rated as "GOOD", had the largest variations in stability index values (QCV = 0.43). The large variations resulted from large variations in slab properties. As shown for the four grids in Fig. 6, the stability index depended on the slab and the weak layer properties. For example, the four grids 0708_1, 0708_3, 0708_7 and 0708_9 had all fairly low median strength in the range of 30 to 50 kPa (Table 3), but stability in terms of median CT score largely differed: the first two were rather unstable whereas the latter two were rather stable.

In all grids, the distribution of stability indices showed clear tendencies towards either primarily stable or primarily unstable values. The two grids 0708_6 and 0708_9 had 0 and 4 % "unstable" stability values, while in the other cases more than 75 % of the stability index values were below the stable-unstable threshold ($S \leq 212$). More even distributions of stability index values (i.e., about half of the stability index values is rather stable while the other half is rather unstable) were not observed.

The variation within a grid, expressed as the quartile coefficient of variation, was typically largest for the stability index (mean QCV = 0.28) and lowest for strength (mean QCV = 0.18). However, the differences were statistically not significant (H test, $p = 0.11$). The QCV and the range – a measure of autocorrelation that was determined by fitting a spherical model to the experimental semivariogram – were not related to the median stability index. The range tended to decrease with increasing QCV, but the trend was statistically not significant ($p = 0.43$).

The slope median stability index was positively related to the slope median strength of the weak layer ($r_p = 0.76$, $p = 0.02$), indicating that stability is in general largely influenced by strength and much less so by the stress (load) ($r_p = -0.47$, $p = 0.15$).

4.2 Spatial analysis

In most grids, the variogram indicated that the range was less than 5 m (Table 3). The values of the range for stress (load), strength and stability index varied on a given day. They were not significantly correlated, though the range of the stability index tended to increase with the range of strength ($r_p = 0.58$; $p = 0.06$). Furthermore, the range of the stability
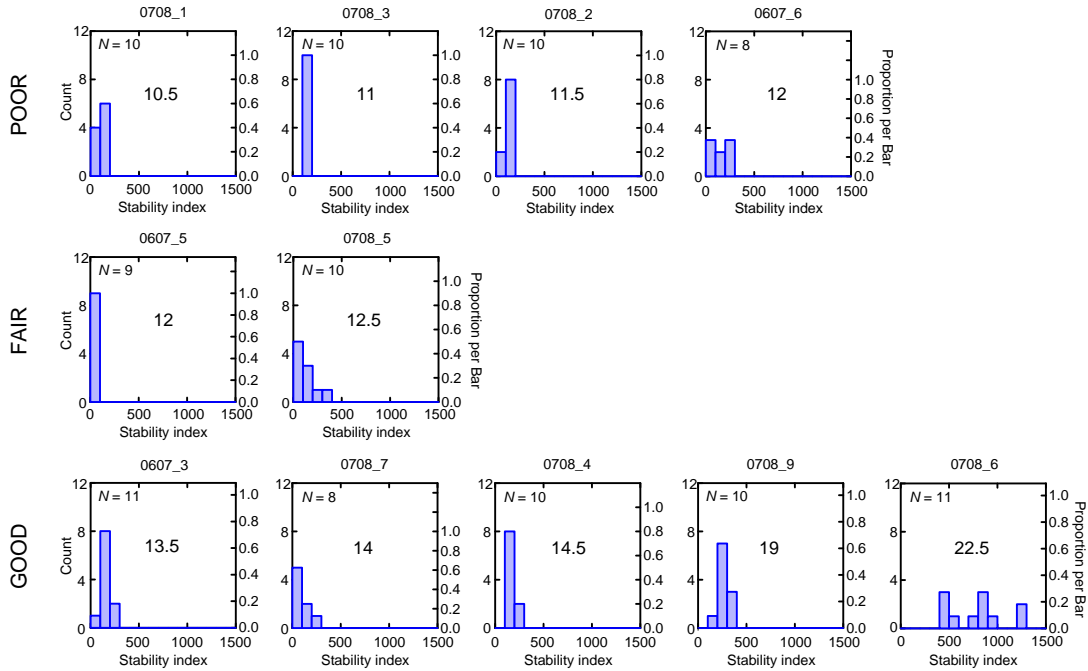
Figure 5. Distributions of stability index for the 11 slopes sampled during winters 2006–2007 and 2007–2008. Only stability derived from those SMP measurements with concurrent CT test are shown. Median CT score is indicated in the middle of each histogram. Slopes are ordered with regard to slope stability ("POOR", "FAIR", "GOOD") and within rows based on median CT score. N varies between 8 and 11.
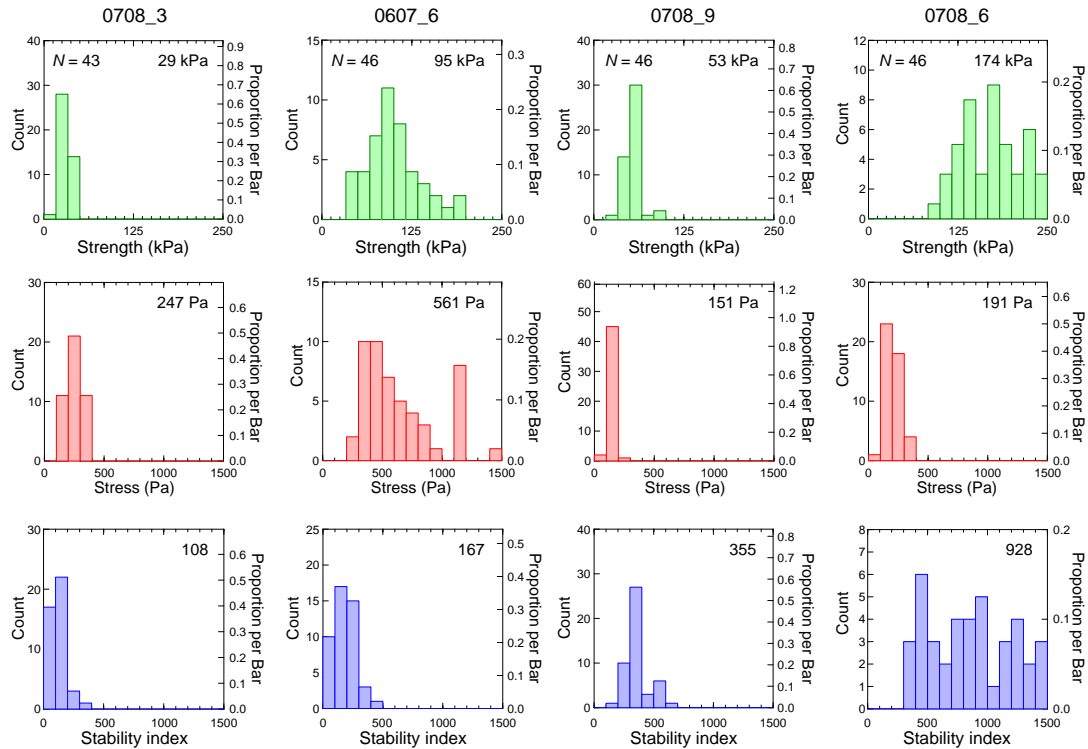

Figure 6. Exemplary distributions of strength, stress and stability index for four slopes (grids). Grids 0708_3 and 0607_6 are rather unstable whereas grids 0708_9 and 0708_6 are rather stable. Numbers indicate the median value. N varies between 43 and 46.

Table 3. Summary statistics for the 11 grids. For signs of instability, "1" indicates the presence and "0" the absence of whumpfs, shooting cracks or recent avalanches. For the stability index, the slope median, the interquartile range (IQR) and the quartile coefficient of variation (QCV) are given per grid. The range is a measure of autocorrelation and was determined by fitting a spherical model to the experimental semivariogram. Proportion weak describes the portion of point stability measurements with $S \leq 212$.

| Grid ID	Date	Median CT score	Signs of instability	Slope stability	Median strength (kPa)	Median stress (Pa)	Stability index Median	IQR	QCV	Range (m) Load	Strength	Stability	Proportion weak
0708_1	10 Jan 2008	10.5	1	POOR	40.9	281	142	89	0.28	2.3	1.6	1.8	0.78
0708_3	23 Jan 2008	11	1	POOR	29.0	247	108	75	0.31	3.2	8.8	9.1	0.95
0708_2	17 Jan 2008	11.5	1	POOR	76.6	623	124	30	0.12	1.7	5.3	1.8	0.96
0607_6	15 Mar 2007	12	1	POOR	95.2	561	167	110	0.33	1.8	0.2	2.7	0.74
0607_5	8 Mar 2007	12	0	FAIR	51.6	1530	35	17	0.24	1.2	1.7	1.4	1.0
0708_5	7 Feb 2008	12.5	0	FAIR	40.8	353	123	102	0.41	4.3	1.9	4.0	0.80
0607_3	16 Feb 2007	13.5	0	GOOD	110	729	145	55	0.18	1.1	5.5	5.6	0.78
0708_7	19 Feb 2008	14	0	GOOD	53.0	517	104	106	0.43	9.2	2.8	2.6	0.90
0708_4	31 Jan 2008	14.5	0	GOOD	91.0	878	122	78	0.28	1.3	9.6	3.5	0.87
0708_9	18 Mar 2008	19	0	GOOD	52.8	151	355	75	0.11	2.0	1.2	1.5	0.04
0708_6	15 Feb 2008	22.5	0	GOOD	174	191	928	691	0.37	2.3	7.4	2.9	0.0
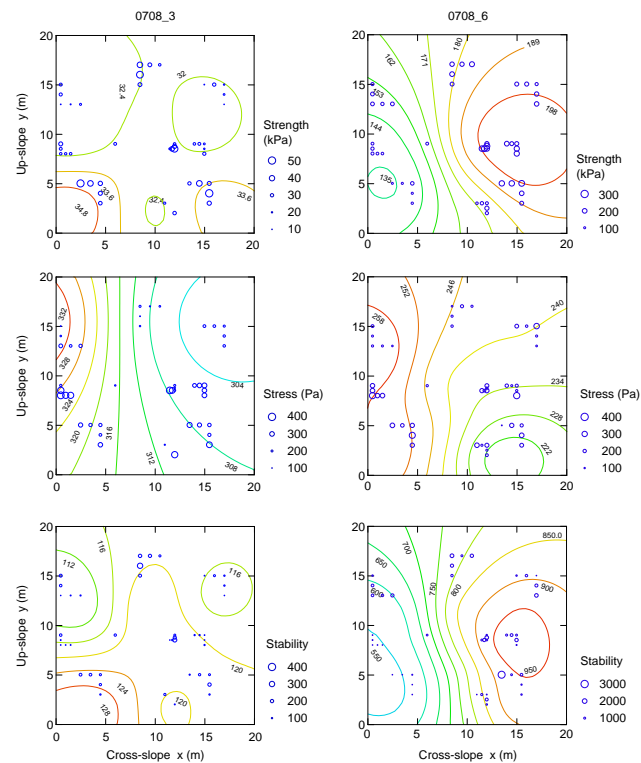
Figure 7. Contour plots of strength, stress (load) and stability index for two exemplary grids in winter 2007–2008. Grid 0708_3 (left) represents a rather unstable slope, whereas grid 0708_6 (right) represents a rather stable one.

index tended to be larger for the slopes rated as "GOOD" than for the slopes rated as either "POOR" or "FAIR", but the difference was small and statistically not significant (H test: $p = 0.7$).

Figure 7 illustrates for two grids the variable spatial structure of strength, stress and the resulting stability index. For grid 0708_3, when stability index values were low, the stress values did not show any particular trend or clustering; only the values towards the lower left and right corners tended to be slightly higher. For stress, a slope scale trend with some higher values towards the left can be observed. The stability index was somewhat higher in the lower left corner than in the higher left one. This observation can be explained by the trend for higher stress in the upper left corner and higher strength in the lower left corner. For the second grid (0708_6) in Fig. 7, some slope scale trends were observed for strength (higher values towards the right), stress (higher values towards the left) and accordingly for stability index (higher values towards the right).

5 Discussion

Based on concurrent SMP measurements and CTs we have derived a new stability index that may be used as a criterion to assess the propensity of failure initiation. The new index has then been applied to a previously published data set of spatially distributed SMP measurements on 11 slopes. The proposed stability index will at best be an estimate of the probability of initiating a failure in a weak layer, but it will not provide any information on the propensity of crack propagation.

As many SMP measurements with concurrent CT results were available, we used the CT as stability reference. Obviously this test is far from perfect (e.g., Winkler and Schweizer, 2009), but at least it is known that the CT score increases with decreasing probability of skier triggering (Jamieson, 1999). Some of the problems include the geometry (scale of length to width), unknown boundary effects and the stepwise loading with only three loading steps. Many

factors that probably play a role (how important is mostly un-known) were not considered in our simple model for deter-mining the additional stress acting at the depth of the weak layer. We considered the loading at the top of the column, the size of load according to the second loading step (tap-ping from the elbow) and the considerable compression of the surface layers. On the other hand, we did not consider the stratigraphy (apart from the surface layers), boundary ef-fects, possible stress waves in the column or load dissipation with depth. As the column is loaded over the whole area we assumed no load dissipation, which is obviously important in the case of a point or a line load. Some recent measure-ments by Thumlert and Jamieson (2014) may question this assumption.

Nevertheless, the derived stability index was clearly re-lated to the stability reference we had at hand. The micro-structural compressive strength has been shown to be related to stability (Pielmeier and Marshall, 2009), and our stability index is slightly better related to the CT score, further sug-gesting that our index is indicative of point snow stability. However, as has been previously shown by Jamieson (1999), the CT does not differentiate well within the intermediate range (CT scores 11 to 20) and hence our stability index is afflicted with the same problem. Accordingly, the correlation between the slope median CT score and the slope median stability index was rather poor, mainly since the sample size was small ($N = 11$) and most grids had a median CT score in the intermediate range.

Certainly, a better stability reference should allow devel-oping a more sophisticated index along the lines of the skier stability index (SK38) (Jamieson and Johnston, 1998). The comparison to the various slope classifications has clearly shown that the index lacks any information about crack prop-agation propensity. The slopes rated as "FAIR" had mostly low values of the stability index but no signs of instability were observed. In the future, the initiation index should be combined with a measure of crack propagation propensity, such as the critical crack length that can be derived from the SMP signal (Reuter et al., 2013).

Non-spatial variations of strength, stress and stability in-dex expressed as the QCV were similar to those found in pre-vious slope-scale studies (Schweizer et al., 2008). Whereas Bellaire and Schweizer (2011) separately related weak and slab layer properties to slope stability, we jointly considered both properties by introducing a simple measure of stability. However, we were still not able to resolve the influence of spatial patterns on slope stability. One of the reasons we did not find any relation between spatial characteristics of point stability measurements and the slope stability estimate may be the lack of slopes exhibiting strong variations in stability index, i.e., about equal shares of high and low stability in-dex values. In such situations one would expect that the spa-tial patterns of point stability characterized by its correlation length would control slope stability. A further reason may be

that the proposed stability index only considers failure initi-ation and does not include crack propagation.

6   Conclusions

We developed a new measurement-based snow instability index combining weak layer and slab properties to pre-dict snow instability in view of dry-snow slab avalanche forecasting. We analyzed a large data set of gridded field measurements acquired on snow slopes by Bellaire and Schweizer (2011) to first derive the index and then relate point stability to slope stability.

The new index combines weak layer strength (SMP-derived micro-structural strength) with a rough measure of the additional stress at the depth of the weak layer, depend-ing on the properties of the surface layers (i.e., slab layers) (SMP-derived penetration depth). The index was positively correlated with the results of CTs performed concurrently with the SMP measurements. It discriminated well between point stabilities rated as either "poor" or "fair" and those rated as "good" with a 10-fold cross-validated classification accuracy of about 80 %. A rich variety of stress, strength and stability scenarios was found, indicating that the index, de-spite its simplicity, seems to be able to mimic at least some of the complex interactions between slab and weak layer prop-erties. The well-known challenging problem of correlating variations in point stability to slope stability could not be solved, despite the fact that now at least a measure of sta-bility exists. However, the target variable – slope stability – is not even well defined either.

In a next step we will seek a data set with reference stabil-ity better suited than the CT, possibly the Rutschblock, and combine a more sophisticated stability index, rather an ini-tiation index, with a propagation propensity index, possibly the critical length from the propagation saw test (Reuter et al., 2014).

References

Bellaire, S.: Spatial variability of the snow cover and its effect on avalanche formation, PhD thesis, Department of Geosciences, University of Hamburg, Hamburg, Germany, 196 pp., 2010.

Bellaire, S. and Schweizer, J.: Measuring spatial variations of weak layer and slab properties with regard to snow slope stability, Cold Reg. Sci. Technol., 65, 234–241, 2011.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: Classification and regression trees, CRC Press, Boca Raton, USA, 1998.

Critchlow, D. E. and Fligner, M. A.: On distribution-free multiple comparisons in the one-way analysis of variance, Commun. Stat.-Theory Methods, 20, 127–139, 1991.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., and Martin, E.: A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting, J. Glaciol., 45, 469–484, 1999.

Floyer, J. and Jamieson, B.: Predicting the fracture character of weak layers from snowpack penetrometer signals, Cold Reg. Sci. Technol., 59, 185–192, 2009.

Föhn, P. M. B.: The stability index and various triggering mechanisms, in: Symposium at Davos 1986 – Avalanche Formation, Movement and Effects, IAHS Publ., 162, edited by: Salm, B. and Gubler, H., International Association of Hydrological Sciences, Wallingford, Oxfordshire, UK, 195–214, 1987.

Fyffe, B. and Zaiser, M.: The effects of snow variability on slab avalanche release, Cold Reg. Sci. Technol., 40, 229–242, 2004.

Gaume, J., Schweizer, J., van Herwijnen, A., Chambon, G., Reuter, B., Eckert, N., and Naaim, M.: Evaluation of slope stability with respect to snowpack spatial variability, J. Geophys. Res., 119, 1783–1799, 2014.

Jamieson, J. B.: The compression test – after 25 years, The Avalanche Review, 18, 10–12, 1999.

Jamieson, J. B. and Johnston, C. D.: Refinements to the stability index for skier-triggered dry slab avalanches, Ann. Glaciol., 26, 296–302, 1998.

Jamieson, B., Haegeli, P., and Schweizer, J.: Field observations for estimating the local avalanche danger in the Columbia Mountains of Canada, Cold Reg. Sci. Technol., 58, 84–91, 2009.

Johnson, J. B. and Schneebeli, M.: Characterizing the microstructural and micromechanical properties of snow, Cold Reg. Sci. Technol., 30, 91–100, 1999.

Kronholm, K.: Spatial variability of snow mechanical properties with regard to avalanche formation, PhD thesis, Faculty of Natural Sciences and Mathematics, University of Zurich, Zurich, Switzerland, 192 pp., 2004.

Kronholm, K. and Birkeland, K. W.: Relating spatial variability to snow stability using a cellular automaton model initialized with field data, Geophys. Res. Lett., 32, L19504, doi:10.1029/2005GL024373, 2005.

Kronholm, K., Schneebeli, M., and Schweizer, J.: Spatial variability of micropenetration resistance in snow layers on a small slope, Ann. Glaciol., 38, 202–208, 2004.

Lutz, E., Birkeland, K., and Marshall, H.-P.: Quantifying changes in weak layer microstructure associated with artificial load changes, Cold Reg. Sci. Technol., 59, 202–209, 2009.

Mackenzie, R. and Payten, W.: A portable, variable-speed, penetrometer for snow pit evaluation, in: Proceedings ISSW 2002. International Snow Science Workshop, Penticton BC, Canada, 29 September–4 October 2002, edited by: Stevens, J. R., 294–300, 2002.

Marshall, H.-P. and Johnson, J. B.: Accurate inversion of high-resolution snow penetrometer signals for microstructural and micromechanical properties, J. Geophys. Res., 114, F04016, doi:10.1029/2009jf001269, 2009.

Pielmeier, C.: Textural and mechanical variability of mountain snowpacks, PhD thesis, University of Berne, Berne, Switzerland, 127 pp., 2003.

Pielmeier, C. and Marshall, H.-P.: Rutschblock-scale snowpack stability derived from multiple quality-controlled SnowMicroPen measurements, Cold Reg. Sci. Technol., 59, 178–184, 2009.

Pielmeier, C. and Schweizer, J.: Snowpack stability information derived from the SnowMicroPen, Cold Reg. Sci. Technol., 47, 102–107, 2007.

Reiweger, I. and Schweizer, J.: Failure of a layer of buried surface hoar, Geophys. Res. Lett., 37, L24501, doi:10.21029/22010GL045433, 2010.

Reuter, B., Proksch, M., Loewe, H., van Herwijnen, A., and Schweizer, J.: On how to measure snow mechanical properties relevant to slab avalanche release, in: Proceedings ISSW 2013. International Snow Science Workshop, Grenoble, France, 7–11 October 2013, edited by: F. Naaim-Bouvet, F., Durand, Y., and Lambert, R., 7–11, 2013.

Reuter, B., Schweizer, J., and van Herwijnen, A.: A process-based approach to estimate point snow instability, The Cryosphere Discuss., 8, 5825–5856, doi:10.5194/tcd-8-5825-2014, 2014.

Schneebeli, M. and Johnson, J. B.: A constant-speed penetrometer for high-resolution snow stratigraphy, Ann. Glaciol., 26, 107–111, 1998.

Schweizer, J. and Jamieson, J. B.: Snowpack properties for snow profile analysis, Cold Reg. Sci. Technol., 37, 233–241, 2003.

Schweizer, J. and Jamieson, J. B.: A threshold sum approach to stability evaluation of manual snow profiles, Cold Reg. Sci. Technol., 47, 50–59, 2007.

Schweizer, J. and Jamieson, J. B.: Snowpack tests for assessing snow-slope instability, Ann. Glaciol., 51, 187–194, 2010.

Schweizer, J., Bellaire, S., Fierz, C., Lehning, M., and Pielmeier, C.: Evaluating and improving the stability predictions of the snow cover model SNOWPACK, Cold Reg. Sci. Technol., 46, 52–59, 2006.

Schweizer, J., Kronholm, K., Jamieson, J. B., and Birkeland, K. W.: Review of spatial variability of snowpack properties and its importance for avalanche formation, Cold Reg. Sci. Technol., 51, 253–272, 2008.

Thumlert, S. and Jamieson, B.: Stress measurements in the snow cover below localized dynamic loads, Cold Reg. Sci. Technol., 106–107, 28–35, 2014.

van Herwijnen, A. and Birkeland, K. W.: Measurements of snow slab displacement in Extended Column Tests and comparison with Propagation Saw Tests, Cold Reg. Sci. Technol., 97, 97–103, 2014.

van Herwijnen, A. and Jamieson, J. B.: Fracture character in compression tests, Cold Reg. Sci. Technol., 47, 60–68, 2007.

van Herwijnen, A., Bellaire, S., and Schweizer, J.: Comparison of micro-structural snowpack parameters derived from penetration resistance measurements with fracture character observations from compression tests, Cold Reg. Sci. Technol., 59, 193–201, 2009.

Wilks, D. S.: Statistical methods in the atmospheric sciences, Academic Press, San Diego CA, USA, 2011.

Winkler, K. and Schweizer, J.: Comparison of snow stability tests: Extended column test, rutschblock test and compression test, Cold Reg. Sci. Technol., 59, 217–226, 2009.

Wright, C.: Introduction to structural impact, in: PDHonline Course S164, available at: http://www.pdhcenter.com/courses/s164/s164content.pdf, PDH Online, Fairfax VA, USA, 2012.