Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Cold Regions Science and Technology 37 (2003) 277–288

cold regions science and technology

www.elsevier.com/locate/coldregions

Verification of regional snowpack stability and avalanche danger

Jürg Schweizer*, Kalle Kronholm, Thomas Wiesinger

Swiss Federal Institute for Snow and Avalanche Research SLF, Flüelastraase 11, Davos Dorf CH-7260, Switzerland

Received 2 September 2002; accepted 2 July 2003

Abstract

Verification of snowpack stability and avalanche danger is a prerequisite for the improvement of avalanche forecasting and the development of models. Although avalanche danger is based on snowpack stability, little is known about the variation of regional snowpack stability at a given danger level. Verification can be done by observation of avalanche occurrence and/or stability tests. To verify avalanche forecasts, and to get a more detailed picture of regional snowpack stability patterns at different danger levels, a large-scale field study has been performed. On four occasions during the winter of 2002 stability data were collected in the region of Davos. During each 1- to 3-day sampling period between 50 and 70 full snow profiles with rutschblock tests were recorded, primarily on shady slopes. At the same time the avalanche danger was estimated based on observations in the field. For analysis the profiles were assigned to one of five stability classes: Very Poor, Poor, Fair, Good, Very Good. Relating the stability to the prevailing (verified) danger level showed distinct patterns of stability. At the danger level Low, 90% of the profiles were rated as Good, or Very Good, whereas at the danger level Considerable, more than 50% showed Poor or Very Poor stability. The coefficient of variation was about 20% independent of the danger level. Significant differences in aspect and elevation existed. Some of the variation could be explained by differences in snow depth and snowpack consolidation (ram resistance). A preliminary analysis of failure layers showed that a persistent weak layer of large faceted crystals above a crust could be found in the majority of the profiles during certain periods. Despite a generally large variation in stability this weak layer was very widespread, and strongly influenced snow stability during the course of the winter, even 2 months after its formation. Due to the stability variation found in this study, verification of avalanche forecasts based on single stability tests cannot be recommended.

Keywords: Avalanche forecasting; Snow stability; Stability test; Snow stability evaluation; Avalanche danger; Spatial variability

---

1. Introduction

Verification of snowpack stability and avalanche danger is a prerequisite for both the development and operational application of models as well as for the improvement of conventional and computer-based avalanche forecasting. Avalanche observation is the best indicator of snowpack instability. Accordingly, various avalanche activity indices were proposed and compared to danger ratings (e.g. Elder and Armstrong, 1987). However, avalanche observations as means of verification is not applicable at all levels and scales of snow stability. At fair and good stability, or alternatively at the danger levels 1: Low, 2: Moderate, and partly at 3: Considerable (as described in the European

---

* Corresponding author. Tel.: +41-81-417-0111; fax: +41-81-417-0110.

E-mail address: schweizer@slf.ch (J. Schweizer).

doi:10.1016/S0165-232X(03)00070-3

avalanche danger scale) (Meister, 1995), other methods than avalanche observations must be applied to verify snowpack stability. Snowpack stability tests are best suited, combined with other observations, to verify the avalanche danger at the lower levels (Föhn and Schweizer, 1995).

In Switzerland, an avalanche forecast is issued daily in the evening, at 5 p.m., valid for the next day. The bulletin is text-based and the forecast covers the whole area of the Swiss Alps, approx. 20,000 km^2. Usually, the whole forecasting area is split into one to three regions where about the same danger level prevails. For each of these large regions the danger level, elevation and aspect of the most critical slopes is given: e.g., Moderate, above 2300 m a.s.l., north–west through north to east aspects (2,>2300, NW–N–E). In the slopes below 2300 m, and in general in the western and southern aspects, the danger is not specified, but is known to be somewhat lower. In addition to the so-called national forecast, a regional forecast is issued daily in the morning at 7 a.m. These bulletins are mainly map-based and are presently prepared for seven regions, each about 2000 km^2 in size. At this regional scale more detail about where the danger prevails can be given. However, even at this relatively fine scale the danger level can vary spatially due to the complex topography and weather patterns in the Alps. It is this regional scale that is of interest, and improvements to the forecast are mainly sought at this scale.

Soratori (1996) and Cagnati et al. (1998) proposed a preliminary scheme to operationally verify avalanche danger. In their method, the rutschblock score was directly related to a certain level of avalanche danger. Due to the limited reliability of single rutschblock test results, mainly as a consequence of the variable nature of the mountain snowpack, this direct link of rutschblock score to avalanche danger is likely inappropriate (Föhn, 1989; Kronholm and Schweizer, 2003; Jamieson, 1995; Landry et al., 2002; Stewart and Jamieson, 2002).

To our knowledge, Munter (1997) has performed the only study on the stability distribution at a given danger level for an area covering several slopes. During numerous avalanche courses he collected stability data based on about 12 rutschkeil (a wedge shaped variation of the rutschblock) tests, on a single day, evenly distributed in the four principal aspects, and related the mean and standard deviation of the rutschkeil scores to the verified avalanche danger. As one of the results of the study Munter (1997) suggested that the number of weak spots should increase exponentially with increasing avalanche danger. Birkeland (2001) investigated snow stability (as measured by stability tests) over a mountain range on two given days in order to better understand its spatial distribution and the implications for predicting dry snow slab avalanches. Spatial stability patterns could only be partly explained by variation of terrain, snowpack and snow strength properties. Wind effects in general and small-scale variability in the snowpack in particular were likely the cause for the partial lack of correlation. However, it represents the first study that quantifies snow stability in terms of terrain. Birkeland (2001) found a pattern of lower stability on high-elevation, northerly aspects.

The aim of the present study is to verify the forecasted avalanche danger level, and in particular to quantitatively describe snow stability patterns at the regional scale. This will provide a basis for a more detailed description of snow stability at a given danger level, a prerequisite for operational verification in view of quality control for avalanche forecasting. At the same time, the data collected will enable the verification of snow profile interpretation and snowpack simulations models.

2. Methods

At four occasions during January to March 2002 the snowpack stability was assessed at the regional scale by a large number of full snow profiles, each supplemented with a rutschblock test. It was assumed that a dataset of about 60 profiles and stability tests would be sufficient to get statistically reliable results. The area tested (about 400 km^2) was the region surrounding Davos (Fig. 1). However, stability tests were actually performed in only four sub-regions of about 30 km^2 in area. The main part of the remaining terrain is not suited for field studies, since it is either not steep enough for stability tests, part of a ski area, or otherwise disturbed by skiing activity. In addition, the sub-regions were chosen such that an automatic weather station was centrally located within each sub-region to enable the verification of the SNOW-

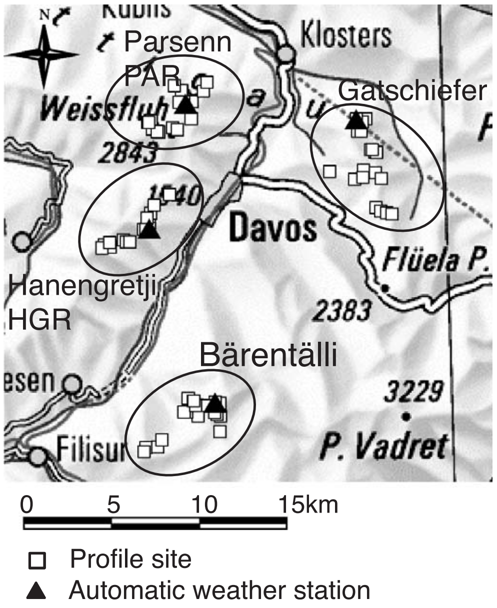

Fig. 1. Map of study area: region of Davos. The profile sites as sampled in the second period (12–13 February 2002) are shown indicating the four sub-regions: Parsenn, Hanengretji, Bärentälli and Gatschiefer. Also shown are the four automatic weather stations that are located each in one of the sub-regions.

PACK model (results will be presented elsewhere) (Lehning et al., 2002). Two sub-regions (Hanengretji: HGR and Parsenn: PAR) were located northwest of the main valley that runs south–west to north–east, and two were south–east of it (Bärentälli: BRT and Gatschiefer: GAT) (Fig. 1). The study areas are above tree-line (which is at an elevation of about 2000 m) and with peaks up to 3000 m. Most data were collected at elevations between 2200 and 2700 m above sea level.

Between 8 and 12 two-person sampling teams per day collected the snowpack and stability data. Access to the study areas was usually by helicopter or by lifts at a nearby ski area. Occasionally, during bad weather and/or critical avalanche conditions, all teams climbed on skins to the study areas or approached from a nearby ski area. Based on the prevailing avalanche conditions for each sampling day, the teams were assigned an area with an aspect and elevation range that should be covered. Otherwise, the teams were free to decide where to sample, based first on safety considerations, and second on the requirements for a representative site. A good test location should, among other things, be steep enough (about 35°), not too close to a crest or cornice, and with uniform, and particularly important, below average snow depth (Schweizer, 2002).

Sampling usually took two to three days to reach a sufficient number of profiles and stability tests per sub-region and aspect or elevation. Typically, on Day 1 of a sampling period, the northerly aspects were tested. On Day 2 the westerly and easterly slopes, and on Day 3 the southerly slopes were included. Temporal evolution was expected not to affect the results in January and February on the mainly shaded slopes. This was considered to be the case when the instability was due to persistent weak layers (that did not change strength quickly), and if slab properties did not substantially change. In addition to the profiles and stability tests on slopes during each period a snow profile, including layer density and two stuffblock tests, was taken near each of the four automatic weather stations within the study area.

Each team visited two to five sites per day. At each site, each team observed a full snow profile, including ram hardness and a rutschblock test. In addition, they recorded observations on avalanche activity, snow surface properties, occurrence of "whumpf"-sounds or any other relevant stability information. Finally, they estimated the prevailing avalanche danger level for that day and the sub-region they had traveled. This is comparable with estimating the class of snowpack stability as done in Canada. In contrast to Canada, observers in Switzerland are much more used to assigning a given avalanche situation to a level of avalanche danger than to a stability class.

For analysis, the profiles with corresponding rutschblock test results (RB score, release type and fracture type) (Schweizer, 2002) were assigned to a certain stability class (1: Very Poor, 2: Poor, 3: Fair, 4: Good, 5: Very Good) according to the scheme proposed by Schweizer and Wiesinger (2001). This stability rating system is exclusively based on snowpack properties (as measured by profiles supplemented with stability tests). The stability class scale runs in the opposite direction of the danger rating numeric scale since low stability implies high danger. Additional data on instability was recorded separately, and integrated in the subsequent assessment of avalanche danger. This stability rating system is different from other systems, e.g. the Canadian, where the stability is described primarily by the probability of avalanche release (McClung and Schaerer, 1993). Our approach follows the well-established procedures of the Swiss Avalanche Warning Service: The snowpack stability derived from snow profiles with stability tests, is combined with other observations on snowpack instability, e.g. avalanche occurrence, and with potential slab thickness, avalanche size and frequency to assess the danger level.

Verifying the avalanche danger with stability tests will likely only work when the danger level is Low, Moderate or Considerable. At higher levels of instability (High, Very High) it is expected that access to study slopes would be too limited to collect representative data. Even at Considerable, only experienced observers will be able to sample critical slopes. Therefore, a certain bias is likely unavoidable at high danger levels and avalanche activity becomes more suitable for verification.

These stability data were then analyzed for each period with emphasis on stability patterns between sub-regions, aspects or elevations. Data were also compared to verified and forecasted avalanche danger levels. Nonparametric statistics were used, primarily the Mann–Whitney *U*-test, to decide whether two stability distributions were different based on a level of significance of $p = 0.05$. The Kruskal-Wallis *H*-test was applied to compare more than two independent samples. Relations of snowpack parameters with stability are described based on Spearman rank-order correlation coefficients.

3. Results

As planned, during four 1- to 3-day sampling periods between January and March 2002 the regional avalanche danger was assessed and snow stability data were collected. In each period 50–70 full snow profiles with rutschblock tests were performed. For each sub-region about a similar number of profiles was collected. During the last period there was a snowfall event that changed conditions substantially. Consequently, the third day during the last period was analyzed separately and called Period 5 (Table 1).

3.1. Weather and avalanche activity

The weather and snowpack development is summarized in Fig. 2. Snow cover formation above 2000 m started in November, slightly later than normal. At the beginning of December, rain was recorded. During this precipitation period the snowfall limit increased several times up to at least 2800 m. This storm formed one to three rain crusts depending on elevation. Toward the end of the precipitation period, temperatures dropped and finally a few centimeters of new snow accumulated on the wet snow. Consequently, a weak layer formed above the upper crust. By mid December cold continental air masses caused a significant drop in temperature and caused significant faceting within the snowpack, in particular around the crusts. Consequently, the first major snowfall after the cold period around New Year caused very high avalanche activity and one fatality in Davos. January

Table 1
Summary of verification periods with corresponding avalanche danger

Period	Date	Days	Profiles	Regional avalanche danger		
				forecasted	verified	analyzed
1	21–23 Jan 2002	2 1/2	62	1, extreme	2, >2300, W–N–E	2, >2300, W–N–E
2	12–13 Feb 2002	2	73	3, >2400, W–N–E	3, >2300, NW–N–NE	3, >2300, NW–N–NE
3	26–27 Feb 2002	1 1/2	50	3, >1800, all	3, >2300, W–N–SE	3, >2300, W–N–NE
4	18–19 Mar 2002	1 1/2	62	2, >2500, NW–N–NE	1–2, >2600, NW–N–NE	1–2, >2500, W–N–E
5	20 Mar 2002	1/2	8	3, >2200, W–N–S	3, >2300, W–N–E	3, >2300, W–N–E

The regional avalanche danger, predicted (forecasted), verified, and analyzed, is given as danger level (1 to 5, Low to Very High), elevation above which the level prevails, and sector of aspects (clockwise): part of the compass with the highest danger. The predicted danger is the same as given in the regional avalanche forecast on the morning of the first sampling day. The verified level is the danger as observed and reported by the sampling teams. The analyzed danger level is derived from the stability distribution based on the profiles and stability tests collected by the sampling teams. If, for the sector of aspects, "extreme" is given, this means that the danger only prevails on a few extremely steep, shady and rocky slopes independent of elevation; whereas "all" means in all aspects.

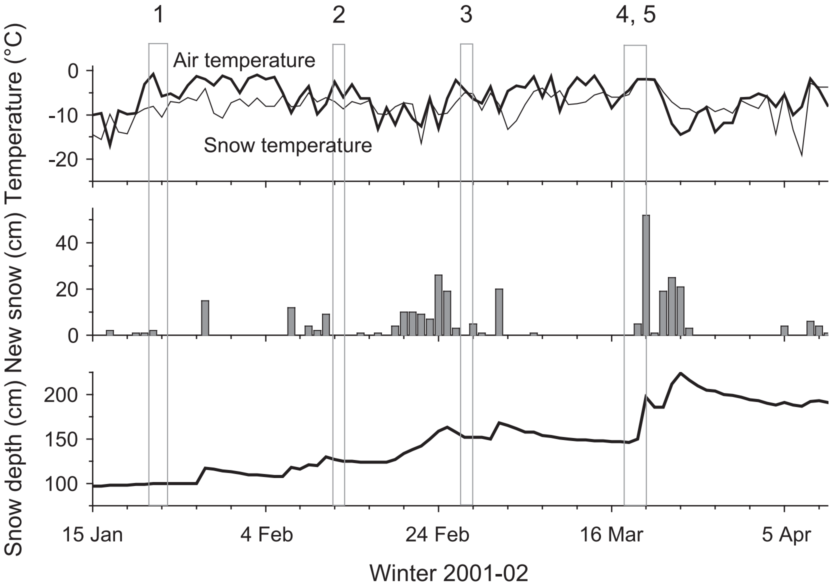

Fig. 2. Snow depth, new snow depth, air and snow temperature (10 cm below snow surface) at Weissfluhjoch 2540 m during January–March 2002 (measurements from 7 a.m.). Vertical boxes indicate verification periods 1–5.

was dry and cool. Snow depth was below average, as during most of the winter. Accordingly, poor snowpack stability, but rather relatively low release probability was observed during the first verification period. Prior to verification period 2, there were 40 cm of new snow. Natural avalanches were rare, but skier triggering was frequent, since by then just the right slab was sitting on top of the weak old snowpack. Thereafter temperatures dropped again and a longer snowfall period at the end of February preceded Period 3. Rising temperatures caused the next avalanche cycle in the first days of March. During March snow stability improved slowly. The last two periods (4 and 5) were shortly before and after a major snowfall on 19–20 March. Relatively high avalanche activity was observed on 20 March, and then again after a subsequent snowfall when artificial triggering was particularly successful. Then the snowpack stabilized quickly and spring conditions started to prevail.

Based on the observations of avalanche occurrence in the region of Davos an avalanche activity index was calculated (Schweizer et al., 1998), and compared to the avalanche danger rating (Fig. 3). The avalanche activity index (AAI) is the sum of all observed avalanches by assigning weights for size and type of triggering. It includes natural avalanches, skier-triggered avalanches and avalanches triggered by explosives. The latter two categories were weighted with weights of 0.5 and 0.2, respectively, according to Föhn and Schweizer (1995). Size of avalanches was considered with weights according to the Canadian avalanche size classes. The weights are 0.01, 0.1, 1 and 10 for the sizes 1–4 respectively. The rather poor correlation between avalanche activity index and danger ratings is obvious. Reasons might be, among others, that the avalanche observations were not consistent enough in frequency and observation area (mainly due to limited visibility), that the danger rating is occasionally inaccurate, and that, in principal, avalanche activity alone is not appropriate to verify at the lower danger levels.

3.2. Regional avalanche danger and snowpack stability

Avalanche danger ratings for the periods of verification are given in Table 1. In the following we will for each period first describe the avalanche danger and discuss deviations between forecasted, verified and analyzed danger, and then report on the snow stability patterns.

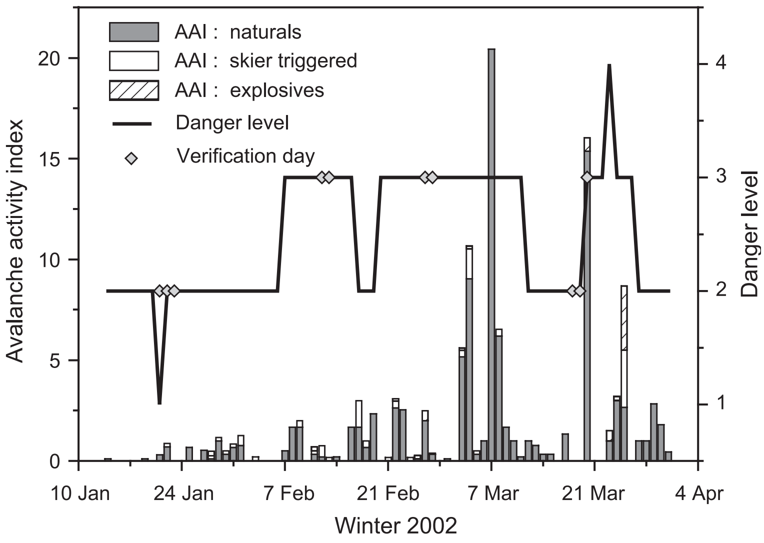

Fig. 3. Avalanche activity index compared to regional avalanche danger rating during January to March 2002 for the region of Davos. Avalanche activity index (AAI) is the sum of all observed avalanches by assigning weights for size and type of triggering. It includes natural avalanches, skier-triggered avalanches, and avalanches triggered by explosives. The danger rating is given as a number from 1 to 4, corresponding to Low, Moderate, Considerable and High avalanche danger. Verification periods are indicated as well.

3.2.1. 21–23 January 2002

At the beginning of the first period (21–23 January 2002) the regional avalanche danger was forecasted as Low. However, observers consistently rated the instability as higher. Likewise many snow profiles that were rated Poor or Very Poor were collected (see below). Accordingly, it follows that the forecast was not correct on that particular day. In fact, based on the stability information collected, the danger rating was revised for the next day.

There was a significant difference in stability between the sub-regions Bärentälli (BRT, $N=14$) and Gatschiefer (GAT, $N=19$). Bärentälli was the most unstable, and Gatschiefer the most stable sub-region. As the stability in the other two sub-regions Hanengretji (HGR, $N=14$) and Parsenn (PAR, $N=15$) were closer to the stability in Bärentälli, these three sub-regions were analyzed together. The difference in stability between BRT-HGR-PAR (Fig. 4a) and GAT (Fig. 4b) was significant ($p=0.009$). Whether evaluated for all cases ($p=0.20$), or grouped in BRT-HGR-PAR ($p=0.10$) and GAT ($p=0.70$), stability on westerly and easterly slopes was not significantly different from stability on northerly slopes. Accordingly, the stability distribution found for the group BRT-HGR-PAR ($N=43$, median: 3, mean ± S.D.: 2.9 ± 0.7) (Fig. 4a) should be representative of Moderate avalanche danger, since this danger level had been verified (Table 1). The stability distribution for GAT (Fig. 4b) represents a danger level between Low and Moderate ($N=19$, median: 3, mean ± S.D.: 3.4 ± 0.7). The relatively large number of profiles rated as Very Poor and Poor during this period confirms the verification of the regional danger as Moderate, in contrast to the forecast. There was no correlation between elevation and stability. There were only seven profiles from below 2300 m, and discriminating at 2300 m showed no significant difference ($p=0.83$).

3.2.2. 12–13 February 2002

At the beginning of the second period the situation was quite critical. There were many clear signs of instability including numerous “whumpf”-sounds. Two teams even triggered a slab avalanche, one remotely, the other at the top of a slope near the crest (so nobody got caught). Accordingly, all teams consistently rated the danger during the second period as Considerable, and occasionally somewhat higher. The forecasted regional danger level Considerable was clearly confirmed by the verification. A slightly narrower sector of aspects than forecasted was found by

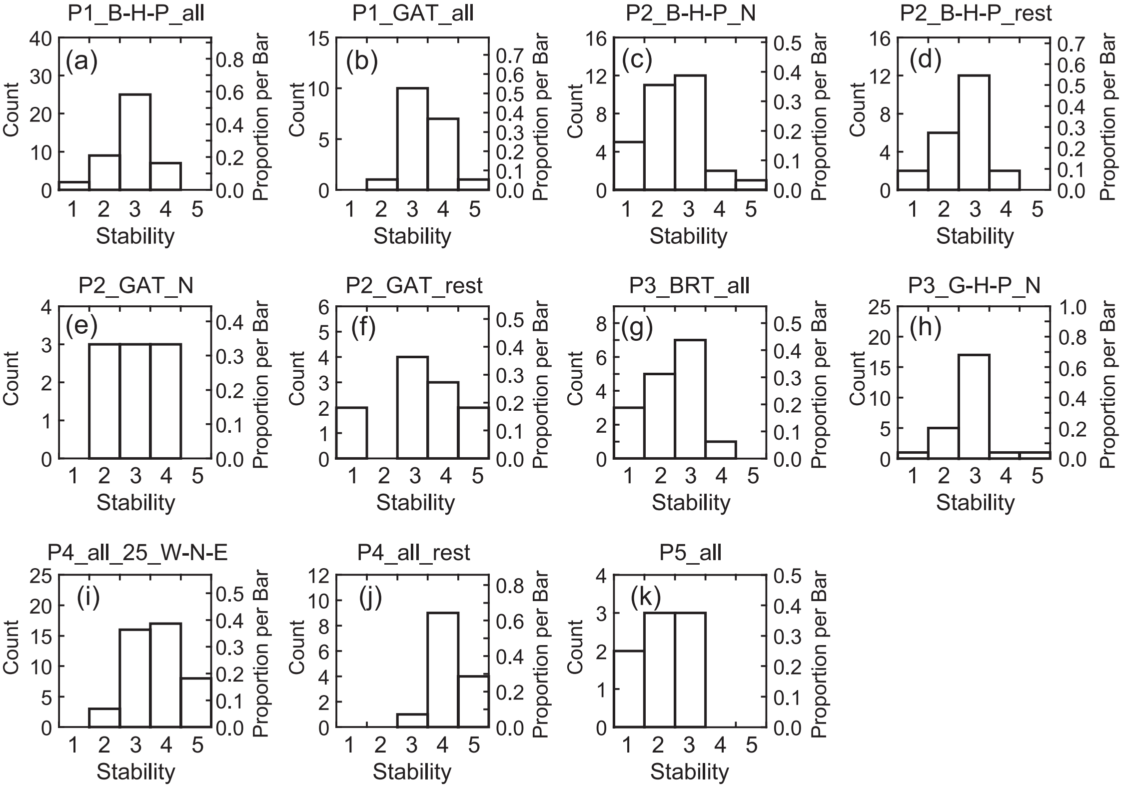

Fig. 4. Stability distributions as found during the five verification periods in different aspects and elevations (for details see text). The headings above the stability distributions shortly describe the period, the regions and aspects, e.g. in (c) "P2_B-H-P_N" means period 2, regions BRT-HGR-PAR, northerly aspects. The stability is given as 1–5 (1 = Very Poor, 2 = Poor, 3 = Fair, 4 = Good, 5 = Very Good).

the observers and supported by the snowpack stability data (see below).

Again, BRT ($N = 18$) proved to be the most unstable, and GAT ($N = 20$) the most stable of the four sub-regions. Elevation was not considered as only five profiles were taken below 2300 m a.s.l. The northerly aspects had a weaker snowpack than the westerly and easterly aspects. However, analyzing all four sub-regions jointly did not show a significant difference in stability ($p = 0.21$). Nevertheless, the stability distributions were quite different. About 50% of the Poor or Very Poor profiles came from slopes of northerly aspect, and only about 30% came from slopes of westerly/easterly aspects. The stability distribution from the slopes of westerly/easterly aspect was quite similar to the stability distribution found in the first period ($p = 0.27$), which was definitely assigned to Moderate danger. Similar results emerged when aspect was analyzed for the regions grouping BRT-HGR-PAR vs. GAT. Finally, the stability distribution found in the northerly aspects of the sub-regions BRT-HGR-PAR ($N = 31$, median: 2, mean ± S.D.: 2.45 ± 0.96) (Fig. 4c) was assigned to Considerable danger. The stability distribution for the westerly/easterly slopes in the sub-regions BRT-HGR-PAR ($N = 22$, median: 3, mean ± S.D.: 2.64 ± 0.79) are shown in Fig. 3d, and correspondingly for the sub-region GAT: northerly slopes ($N = 9$, median: 3, mean ± S.D.: 3.0 ± 0.87) (Fig. 4e), and in the westerly/easterly slopes: ($N = 11$, median: 3, mean ± S.D.: 3.3 ± 1.35) (Fig. 4f). There were only five profiles from below 2300 m. Discriminating at a certain elevation was not possible. However, interestingly, there was a slight trend towards higher stability at higher elevation.

3.2.3. 26–27 February 2002

At the beginning of the third period the avalanche warning service rated the situation as Considerable.

Despite the fact that all steep slopes above 1800 m a.s.l. were considered critical, the forecaster in charge indicated that the danger would likely be somewhat less critical than in the second period, i.e. that he expected the release probability to be somewhat lower than in the previous verification period. In fact, there were less obvious signs of instability, but the observers rated the danger as Considerable. However, they felt that the critical slopes could be more precisely described, i.e. were located primarily in the aspects west through north to south–east, above 2300 m. The stability distribution found (see below) supports this finding.

The sub-region BRT ($N = 16$) was the most unstable of the four sub-regions again, suggesting a grouping of BRT vs. GAT-HGR-PAR ($N = 34$), which revealed a significant difference ($p = 0.024$). During this period profiles with stability tests were taken on avalanche slopes from all four aspects. The grouping of aspects was not straightforward (Fig. 5). The westerly slopes were the most unstable. South facing and flat slopes were more stable than the rest. Analyzing the westerly and northerly slopes jointly vs. the rest revealed a statistically significant difference ($p = 0.045$). If the sub-regions were analyzed separately to find differences in aspect, it showed that for the sub-region BRT profiles from different aspects were not different, and thus are jointly shown in (Fig. 4g). For the grouping GAT-HGR-PAR the profiles from the slopes in the sector of aspects W–N–SE were more unstable than from slopes in the remaining sector (SW–S) ($p = 0.048$). In summary, the stability distribution from the sub-region BRT ($N = 16$, median: 2–3, mean ± S.D.: 2.38 ± 0.89) (Fig. 4g) should be typical for Considerable avalanche danger, whereas the stability distribution for the slopes of aspect W–N–SE from the sub-regions GAT-BRT-PAR (Fig. 4h) indicates a somewhat lower degree of avalanche danger ($N = 25$, median: 3, mean ± S.D.: 2.84 ± 0.75). There was a slight trend of lower stability at higher elevation. Discriminating at 2300 m as proposed by the observers, did not show a significant difference ($p = 0.9$) in stability between the two groups. The best discrimination in elevation was found at 2400 m, but still the difference was not significant ($p = 0.18$).

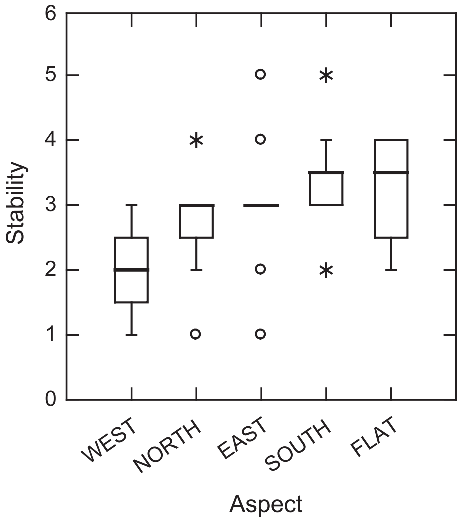

Fig. 5. Stability for different aspects as found during the third period (26–27 February 2002), all sub-regions considered jointly ($N = 50$).

3.2.4. 18–19 March 2002

During the fourth period (18–19 March 2002) the regional avalanche danger was significantly lower than in the previous periods. The avalanche warning service forecasted Moderate danger, above 2500 m, on northerly slopes. The observers rated the danger as Moderate or Low, i.e. they did not agree on the rating: six teams proposed Low, five teams proposed Moderate and another six teams estimated the danger as between Low and Moderate. There was agreement that if Moderate danger was present, it only existed in the northerly slopes and at higher elevation. The analysis of the stability essentially confirmed this estimate. In particular, it could be shown that the danger in the northerly slopes above 2500 m was less than Moderate, but more than Low. Therefore, in Table 1, the intermediate danger level of 1–2 is given, although operationally in the forecast, no intermediate danger levels are given. Our analysis revealed that above 2500 m, the northerly slopes were not significantly different from the easterly/westerly slopes (see below), and the critical aspects were described as W–N–E.

There was no significant difference in stability between the four sub-regions (significance levels p between 0.25 and 0.86). For all regions the median stability was 4: (Good). Profiles were taken at elevations between 2100 and 2900 m. Observers suggested that stability should be better below 2600 m, and accordingly poorer above 2600 m. In fact, Fig. 6 suggests a dependence of stability on elevation, and a differentiation at 2500 m. However, the profiles taken on glaciers at about 2900 m were more stable than those taken at lower elevations and so were not considered for the linear regression, which was significant ($p = 0.002$). Comparing snowpack stability above and below 2500 m shows a significant difference, even if the profiles on glaciers at about 2900 m are included ($p = 0.011$). Above 2500 m, the westerly and easterly slopes were not significantly different from the northerly slopes ($p = 0.10$). Comparing the slopes from the sector W–N–E above 2500 m ($N = 44$, median: 4, mean ± S.D.: 3.68 ± 0.86) (Fig. 4i) with the rest of the profiles showed a significant difference ($p = 0.032$). So the rest (Fig. 4j), i.e. all the profiles below 2500 m, and all the profiles above 2500 m, but not from aspects in the sector W–N–E ($N = 14$, median: 4, mean ± S.D.: 4.21 ± 0.58), should correspond to Low avalanche danger. On the other hand, the stability distribution from the northerly slopes above 2500 m was significantly different (< 0.001) from the distribution that is typical for Moderate. Accordingly, on these slopes Low to Moderate danger prevailed.

3.2.5. 20 March 2002

On the second day of the fourth period it started snowing. By the morning of the next day (20 March 2002) about 50 cm of new snow had fallen at relatively high temperatures (−2 °C at 2500 m). Consequently, the forecasted regional danger level increased to Considerable. This was fully confirmed by the verification. Although the avalanche activity index was relatively high as well (Fig. 3), there was not sufficient evidence for the danger level to increase to High.

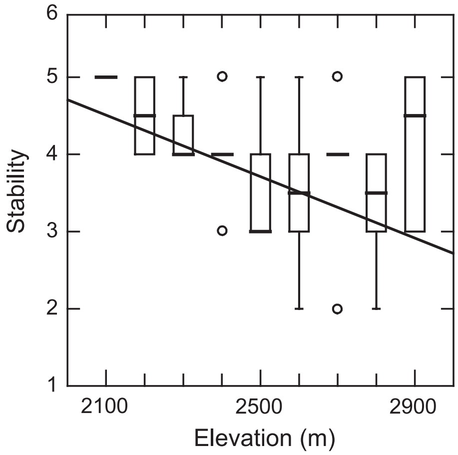

Fig. 6. Dependence of stability on elevation during the fourth period (18–19 March 2002) ($N = 62$).

Sampling possibilities were limited that day, in particular since it only cleared up in the early afternoon, but a few teams collected data nevertheless. All teams consistently rated the avalanche situation as Considerable. Due to the small number of observations ($N = 8$) no grouping by elevation or aspect was possible. The stability distribution found (median: 2, mean ± S.D.: 2.1 ± 0.83) (Fig. 4k) confirms the assessment of Considerable avalanche danger. It was the lowest average stability found during all verification periods.

3.3. Relationships with profile parameters

Stability mainly depends on snowpack properties such as weak layer strength and on slab properties such as thickness and hardness. Terrain parameters elevation and aspect also obviously affect stability, as shown above, not directly, but in a complex way. Other snowpack parameters likely to exhibit a certain relation to stability are, among others, snow depth, relative snow depth (compared to regional average), average ram resistance (snow hardness) and type of hardness profile. More directly related are rutschblock score, rutschblock release type (portion of the block that did slide), rutschblock fracture type (roughness of failure surface) (Schweizer, 2002) and fracture depth which all include stability information. Spearman rank-order correlation coefficients between snow stability and RB score, and some other snow profile parameters are given in Table 2. Highly correlated with stability are snow depth, relative snow depth, ram resistance, RB score and RB release type. However, the latter two parameters are not actually independent variables. Fracture depth is highly correlated with RB score, but not with stability. This is consistent with the fact that the RB score includes the effect of fracture depth, and with the well-established procedure to not consider fracture depth when assessing stability, but when estimating the danger level. Rutschblock fracture type (smooth, rough, irregular)

Table 2
Spearman rank–order correlations between RB score, stability class, and some other snowpack parameters

Variable	RB score	Stability
Snow depth	**0.159**	**0.382**
Relative snow depth	0.113	**0.253**
Ram resistance	**0.221**	**0.407**
RB score	–	**0.730**
Fracture depth	**0.208**	0.018
RB release type	**0.348**	**0.340**
RB fracture type	0.095	0.080

Correlations in bold are significant at $p < 0.05$.

is not significantly correlated. This is somewhat surprising, since irregular fractures are clearly associated with high values of stability (median: 4). The poor correlation seems to be due to the fact that the majority (85%) of fractures was rated as smooth. Probably two classes of fracture type would be sufficient. In fact, in that case (smooth/rough vs. irregular) correlation with stability would be significant ($p = 0.026$). Also, assessing fracture type is much more difficult than release type, and observers might not have reported fracture type consistently.

3.4. Observation bias

As indicated above, prevailing avalanche danger may affect the choice of sampling locations by the observers. Accordingly, for the Periods 1–4 we compared the profiles collected by very experienced forecasters to the ones by less experienced forecasters. In three periods the average stability found by the very experienced observers was lower, in one case equal to the stability found by the less experienced observers. Also, the spread of stability variation was in two cases substantially smaller, in one case about equal and in one case slightly larger. The results include all profiles regardless of differences in aspect and elevation.

4. Discussion and conclusions

The verification of avalanche danger, by observation in the field, or by analyzing the stability patterns, did not show substantial deviations from the avalanche forecast (Table 1). The main differences result from different sectors of aspect or elevations. Avalanche activity correlated poorly with danger rating, suggesting that verification with avalanche observation alone is not feasible at danger levels Low, Moderate and Considerable.

The stability distributions found at the different occasions cover the danger levels of Low, Moderate and Considerable. However, for Low there is only one situation with rather few profiles ($N = 14$). In total, 11 different stability distributions resulted (Fig. 4). Combining similar distributions reveals characteristic distributions for Low, Moderate and Considerable (Fig. 7). For Low danger about 90% of the profiles sampled were rated as Good or Very Good (median: 4, mean ± S.D.: 4.2 ± 0.6). For Moderate danger about 20–25% of the profiles were each rated Poor or Very Poor, or Good and Very Good (median: 3, mean ± S.D.: 2.9 ± 0.9). For Considerable danger about 50% of the profiles were rated Poor or Very

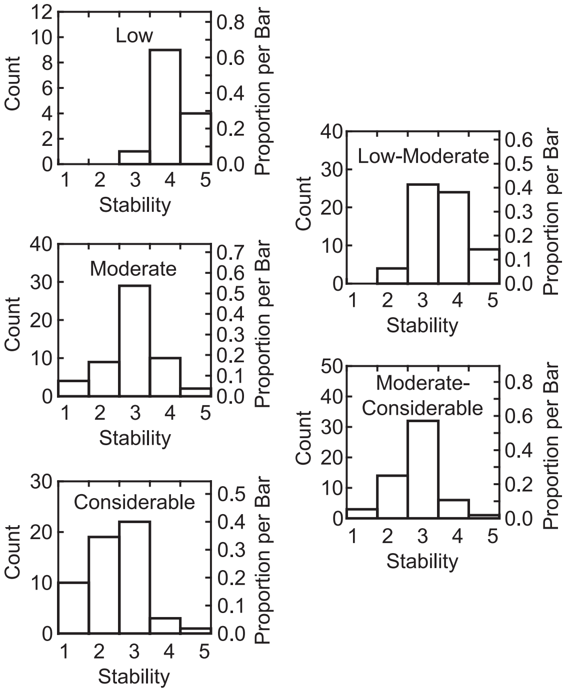

Fig. 7. Characteristic stability distributions for the danger levels of Low, Moderate and Considerable (left) compiled from the 11 stability distributions (Fig. 4) found during the five verification periods. On the right the two distributions are given that fall in between the full danger levels. These intermediate danger levels are called Low–Moderate and Moderate–Considerable.

Poor (median: 2, mean ± S.D.: 2.4 ± 0.9). Between Low and Moderate there was a stability distribution that was significantly different both from Low and Moderate (median: 4, mean ± S.D.: 3.6 ± 0.8). This distribution was assigned to an intermediate danger level of Low–Moderate. Between Moderate and Considerable there was another intermediate stability distribution, called Moderate–Considerable (median: 3, mean ± S.D.: 2.8 ± 0.8). This distribution is quite close to Moderate, but significantly different from Considerable. However, when this type of stability distribution was found observers rated the situation consistently higher than Moderate. Relative stability variation expressed as the median of the quartile coefficients of variation (Spiegel and Stephens, 1999), was about 20%, in accordance with coefficients of variation found in other studies and at different scales (Kronholm and Schweizer, 2003). No relation between median stability and stability variation was found.

The field study has shown that significant variations in aspect and elevation typically exist. Aspect clearly had more influence, and level of danger, or stability, should not be averaged over aspect, unless field checked for similarity. This is fully taken into account by the Swiss avalanche warning service that forecasts elevation and sector of aspects (part of the compass) for the *highest* prevailing danger level. However, nothing is usually said about the avalanche danger in the adjacent aspects or elevations. Based on our analysis of a limited number of situations, we now have a clear indication that, typically, the avalanche danger is at least half a danger level less in the immediately adjacent aspects and elevations. Occasionally, it is a full danger level less than in the critical slopes specified in the forecast. Specifying the elevation is more difficult and includes more uncertainty than specifying the aspects.

There were also significant differences between sub-regions. The stability was usually poorest in the sub-region Bärentälli, whereas it was usually best in the sub-region Gatschiefer. This is most likely due to their slightly different snow climates. In fact, comparing average snow depth of the profiles in the four sub-regions shows that it is lowest (115 cm) in the sub-region Bärentälli, and highest (138 cm) in Gatschiefer.

Regarding the operational verification of avalanche forecasts by snowpack stability tests, it has been shown that verification based on single stability tests is clearly not possible due to the stability variation found, even on slopes of the same aspect. However, experienced observers will likely find the appropriate spots for representative stability tests (targeted data sampling) more easily, and will therefore need fewer tests to arrive at a reliable stability result than is suggested by the present study, where the sampling was less specifically focused on seeking instability (McClung, 2002). Variability at the slope-scale might also increase the stability variation at the regional scale. To check this influence and hence the representativeness of the stability test locations, during each of the periods, for a limited number of test sites, additional SnowMicroPen measurements (Schneebeli et al., 1999) in the surroundings (12 × 12 m) of the stability test sites were performed. Those results will be presented elsewhere.

In addition, preliminary analysis showed that in addition to terrain parameters, snow depth and average ram resistance, representing simple snowpack properties, are well correlated with snow stability. Stability tests at sites with lower than average snow depth clearly give more indicative results than those from sites with a deep snow cover. Considering the rutschblock test result, besides the RB score, release type and partly fracture type are also correlated with stability, in line with the results of Johnson and Birkeland (2002).

The preliminary data analysis showed that further observations are needed to conclusively relate a given stability distribution to a certain danger level, a prerequisite for improving and refining the descriptions of the five danger levels. Furthermore, the effect on snow stability of the properties of the layers at the fracture interface, as found by the stability test, will be established. Finally, we will seek to model stability based on terrain parameters and snowpack properties to quantify snow stability variation at the regional scale.

Acknowledgements

For data collection during the field campaigns we would like to thank the following members of SLF: B. Aeberhardt, M. Aebi, M. Bründl, C. Cunningham, H.-J. Etter, T. Exner, M. Fauve, P. Federolf, A. Felber,

C. Fierz, A. Heilig, J. Hendrikx, D. Kasbauer, T. Kocher, M. Laternser, M. Lehning, T. Lichtenheld, M. Lütschg, M. Maggioni, S. Margreth, O. Martius, M. Matzl, I. Meirold-Mautner, R. Meister, M. Phillips, C. Pielmeier, N. Raderschall, HU. Rhyner, M. Schaer, G. Spreitzhofer and B. Zweifel. During the fourth and fifth period the Avalanche Service of the Swiss Army contributed to the data collection, including: F. Conceprio, O. Duvanel, J.-R. Escher, S. Felix, M. Fuchs, F. Huber, M. Lustenberger, M. Nagel, C. Pfammatter, U. Rieder, T. Schneidt, M. Schnellmann, N. Vonarburg, S. Zürcher. Their support is greatly acknowledged. We would also like to thank C. Fierz, J.B. Jamieson and I. McCammon, and two anonymous reviewers for valuable comments that helped to improve the paper.

References

Birkeland, K.W., 2001. Spatial patterns of snow stability throughout a small mountain range. J. Glaciol. 47 (157), 176–186.

Cagnati, A., Valt, M., Soratori, G., Gavaldà, J., Sellés, C.G., 1998. A field method for avalanche danger-level verification. Ann. Glaciol. 26, 343–346.

Elder, K., Armstrong, B., 1987. A quantitative approach for verifying avalanche hazard ratings. IAHS Publ. 162, 593–601.

Föhn, P.M.B., 1989. Snow cover stability tests and the areal variability of snow strength. Proceedings International Snow Science Workshop, Whistler, British Columbia, Canada, 12–15 October 1988, pp. 262–273.

Föhn, P.M.B., Schweizer, J., 1995. Verification of avalanche danger with respect to avalanche forecasting. In: Sivardière, F. (Ed.), Les apports de la recherche scientifique à la sécurite neige, glace et avalanche. Actes de Colloque, Chamonix, 30 mai–3 juin 1995. ANENA, Grenoble, France, pp. 151–156.

Jamieson, J.B., 1995. Avalanche prediction for persistent snow slabs. PhD Thesis. University of Calgary, Calgary, AB, Canada. 258 pp.

Johnson, R.F., Birkeland, K.W., 2002. Integrating shear quality into stability test results. In: Stevens, J.R. (Ed.), Proceedings ISSW 2002, International Snow Science Workshop, Penticton BC, Canada, 29 September–4 October 2002, pp. 508–513.

Kronholm, K., Schweizer, J., 2003. Snow stability variation on small slopes. Cold Reg. Sci. Technol. 37, 453–465 (in press, see paper in this issue).

Landry, C.C., Birkeland, K., Hansen, K., Borkowski, J., Brown, R.L., Aspinall, R., 2002. Snow stability on uniform slopes: implications for avalanche forecasting. In: Stevens, J.R. (Ed.), Proceedings ISSW 2002, International Snow Science Workshop, Penticton BC, Canada, 29 September–4 October 2002, pp. 532–539.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., 2002. A physical SNOWPACK model for the Swiss avalanche warning: Part III. meteorological forcing, thin layer formation and evaluation. Cold Reg. Sci. Technol. 35 (3), 169–184.

McClung, D.M., 2002. The elements of applied avalanche forecasting: Part I. The human issues. Nat. Hazards 26 (2), 111–129.

McClung, D.M., Schaerer, P., 1993. The Avalanche Handbook. The Mountaineers, Seattle, WA, USA. 271 pp.

Meister, R., 1995. Country-wide avalanche warning in Switzerland. Proceedings International Snow Science Workshop, Snowbird, Utah, USA, 30 October–3 November 1994, pp. 58–71.

Munter, W., 1997. 3 × 3 Lawinen-Entscheiden in kritischen Situationen. Agentur Pohl and Schellhammer, Garmisch Partenkirchen, Germany. 220 pp.

Schneebeli, M., Pielmeier, C., Johnson, J.B., 1999. Measuring snow micro structure and hardness using a high resolution penetrometer. Cold Reg. Sci. Technol. 30 (1–3), 101–114.

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

Schweizer, J., Wiesinger, T., 2001. Snow profile interpretation for stability evaluation. Cold Reg. Sci. Technol. 33 (2–3), 188–189.

Schweizer, J., Jamieson, J.B., Skjonsberg, D., 1998. Avalanche forecasting for transportation corridor and backcountry in Glacier National Park (BC, Canada). In: Hestnes, E. (Ed.), 25 Years of Snow Avalanche Research, Voss, Norway, 12–16 May 1998. NGI Publication, Norwegian Geotechnical Institute, Oslo, Norway, pp. 238–243.

Soratori, G., 1996. La verifica del grado di pericolo di valanghe. Neve e Valanghe, AINEVA 29, 18–27.

Spiegel, M.R., Stephens, L.J., 1999. Schaum's Outline of Theory and Problems of Statistics. Schaum's Outline Series. McGraw-Hill, New York. 538 pp.

Stewart, K., Jamieson, J.B., 2002. Spatial variability of slab stability in avalanche start zones. In: Stevens, J.R. (Ed.), Proceedings ISSW 2002, International Snow Science Workshop, Penticton BC, Canada, 29 September–4 October 2002, pp. 544–548.