TEMPORAL CHANGES IN THE SPATIAL VARIABILITY OF SHEAR STRENGTH AND STABILITY

Spencer Logan[1], Karl Birkeland[1,2], Kalle Kronholm[1], Kathy Hansen[1]
[1]Department of Earth Sciences, Montana State University
Bozeman, MT 59717-3480 USA
spencer_logan@hotmail.com
[2]U. S. Forest Service National Avalanche Center
kbirkeland@fs.fed.us

ABSTRACT: Avalanche forecasting involves the prediction of spatial and temporal variability of the snowpack. To predict avalanches with more accuracy it is important to determine whether the snowpack is becoming more spatially variable or more spatially uniform. Greater variability increases uncertainty in extrapolation and prediction. Our results offer a look at the evolution of the spatial variability of shear strength and stability of a buried surface hoar layer in southwestern Montana, USA, from shortly after burial until it was no longer the weakest layer in the snowpack. We selected the study site for its 27-degree planar slope, uniform ground cover, and wind-sheltered location. This simplified the comparison of the plots by minimizing initial spatial differences so we could focus on temporal change. Within the site, we sampled four 14 m x 14 m arrays of more than 70 shear frame tests in a layout optimized for spatial analysis. Over a three-week period, the sampling of the four adjacent arrays showed temporal change. The variability of the shear strength of this layer initially decreased then became increasingly variable through time. This suggests that extrapolating test results to other locations becomes increasingly unreliable as layers age, a result that matches practical experience. The data also provide indications that shear strength has a correlation length, the distance at which test results are related, of just a few meters. This short correlation length demonstrates quantitatively why stability tests that are relatively close together can be quite different.

Keywords: spatial variability, avalanches, avalanche forecasting, temporal change

1. INTRODUCTION

Forecasting avalanche hazard takes information collected over limited spatial extent and extrapolates to areas of interest or to larger geographic areas. The amount of spatial variability affects the reliability of extrapolating snowpack data by introducing uncertainty (LaChapelle, 1980; McClung, 2002). Without knowing the degree of the variability, it is difficult to assess the full range of possible variations of stability

Two properties of interest in avalanche forecasting are the shear strength of weak layers and snow stability, which are related. Experience and previous research suggests that shear strength and stability can vary dramatically over a slope and that weak layers gain strength as they age (Conway and Abrahamson, 1984; Jamieson and Johnston, 1999; Kronholm and Schweizer, 2003; Landry et al., in press; Stewart and Jamieson, 2002). Other research suggests that spatial variability may change through time, with less stable slopes having less variability (Birkeland and Landry, 2002) and Kronholm and Schweizer (2003) propose a scheme to relate variability and slope stability.

Quantifying trends in the temporal changes of spatial variability will improve avalanche forecasting. Knowing that stability is becoming more variable would allow an avalanche forecaster to seek data at a greater spatial density, or confine extrapolation to shorter distances. Conversely, if stability becomes less spatially variable, reliable extrapolation would be possible over greater distances.

This project examined the temporal change in shear strength and stability over uniform slopes, and addressed three primary questions:

What are the spatial characteristics of strength and stability on a given day?

How does the spatial variability of strength and stability change through time?

Can the changes in spatial structure be related to snowfall or other changes?

This paper addressed these questions by concentrating on one of three sites sampled during the 2003-2004 winter.

2. METHODS

The study site, known as Spanky's, is 3.5 km north of Big Sky, Montana. The site is in the intermountain avalanche climate zone, which exhibits a variety of avalanche and snowpack conditions, including persistent weak layers (Mock and Birkeland, 2000; Tremper, 2001). Safe access to the site allowed for sampling while avalanche conditions were Considerable or High.

The site is a wind-sheltered grassy opening with an ENE aspect at an elevation of 2640 m. The average slope angle is 27° and the entire site varied less than 5° in angle or aspect. Vegetation ranged from grass and forbs to shrubs 0.4 m high. When sampled, the weak layer was more than 1 m above the ground, so the vegetation should have had little influence on the weak layer. We signed the site at the beginning of the winter to keep it undisturbed by skiers.

To enable comparison of the pits within the plot, we selected the site to minimize sources of variability across the plot, such as wind drifting, changes in slope angle, or rocks and shrubs.

2.1. Site Layout and Sampling

The site consisted of four 14 m x 14 m plots in two rows, separated by 3 m wide 'alleys,' for a total of 961 m^2 (Figure 1). The alleys, sampled first, allowed us to investigate the consistency between the plots by characterizing any site-scale trends. The alleys consisted of 48 shear frame tests in groups of four tests at 12 locations (Figure 1).

We sampled the first plot within a few days of the alleys, and the remaining plots at approximately weekly intervals. We selected sample days to follow periods of snowfall and to allow for sufficient changes in shear strength and stability.

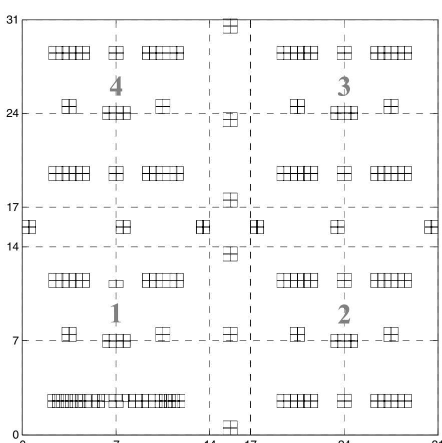

Figure 1. Layout of the Spanky's site, with shear frame locations indicated by the squares and the plot numbers in gray.

In each plot, we sampled five main pits and four smaller pits (Figure 1). The five main pits allowed for pit-to-pit and pit-to-plot comparisons (Birkeland and Landry, 2002; Landry, 2002). Including the four smaller pits improved the calculation of the spatial statistics. Within each pit, we placed tests 0.5 m apart.

We utilized shear frame tests to quantify weak layer shear strength at each test location (Jamieson and Johnston, 2001; McCammon, 2003). Shear frames allowed the targeting of a specific weak layer, even when it was no longer the weakest layer in the snowpack. Measurements of slab thickness, density, and slope angle at each pit allowed the calculation of stability ratios (Jamieson and Johnston, 2001; McCammon, 2003).

2.2. Analysis

Two statistics described the plot-wide characteristics. These are not spatial measures, and only described the test results grouped over the entire plot. Robust statistics were used because the distributions of shear strength were skewed for Plots 2, 3, and 4. The median and quartile coefficient of variation (QCV) characterized the central tendency and relative spread of the plots. Medians are less sensitive to outlying values than are means, so they provided a better measure of the central tendency. The QCV was calculated as,

$$QCV = \frac{1/2(Q_3 - Q_1)}{1/2(Q_3 + Q_1)} \qquad (1)$$

where Q_1 and Q_3 are the first and third quartiles respectively. The numerator in Eq. 1 is a robust measure of the spread of the data while the denominator is a robust measure of central tendency similar to the median. The QCV is similar to, but not equivalent to, a parametric coefficient of variation (Spiegel and Stephens, 1999).

We used two methods to describe spatial variability. Linear trend surfaces described any plot-wide trends, and variograms measured spatial auto-correlation among the test. Trend surfaces use regression on the measurement locations as predictor variables. The trend was removed if it was significant ($p < 0.05$) and explained more than 10% of the variability in the data ($R^2 > 0.10$). The residuals were used for the subsequent spatial analysis.

Using linear trends does not imply that trends present in the snowpack were actually linear. We chose linear trends over methods that are more complicated because they are relatively easy to interpret and better explained plot-wide trends.

Variograms quantify the spatial correlation within a data set by calculating the average variance between data points over distance (Webster and Oliver, 2001). The location and result of each measurement is explicitly included in the analysis. Only recently have variograms been applied to shear strength and stability data (Kronholm, 2004) though variograms have been applied to other snowpack properties (Bloschl, 1999).

A variogram is half the sample variance, the semivariance, plotted as a function of distance (Figure 2). The sill occurs where the semivariance "levels off," typically corresponding to the overall variance of the dataset. The range is the distance to the sill, and is indicative of the spatial scale of the process (Bloschl, 1999). At distances beyond the range there is no longer correlation of results; measurement results are independent and not related. The nugget is the difference between zero and the semivariance at the shortest lag distance. The nugget is variance that cannot be resolved by the measured data (Myers, 1997). If there was no correlation in the data, the variogram would be a horizontal line, termed "pure nugget."

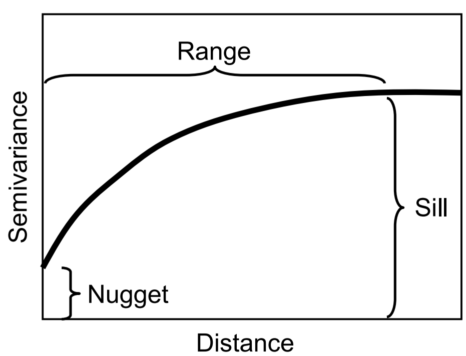

Figure 2. Variogram, with parts important to interpretation labeled

A "semi-spatial" method to characterize the spatial variability of a plot is the pit-to-plot ratio (Birkeland and Landry, 2002; Landry, 2002). Pit-to-plot ratios characterize the ability of a single pit to represent the results of the entire plot. A pit is "representative" of the plot if there was no statistically significant difference between the results in a pit and the results of all the tests for a plot. The Wilcoxon Test was used to compare the individual pits to the pooled results. The Wilcoxon Test assumes only that the data distributions are identical. We pooled all results to increase the conservativeness of the test.

If more pits represent the plot, the plot is less spatially variable. Pit-to-plot ratios are not spatial measures, because the locations of the tests are not explicitly considered in the analysis. The location of the pits within the plot, or the location of the test within the pit, does not matter. However, pit-to-plot ratios do provide a measure of the spatial variability within the data, and the representativeness of a pit for a plot. It provides practitioners with an indication of whether the results from one pit could be reliably extrapolated across the distance of a plot.

3. RESULTS AND DISCUSSION

4. SPANKY'S

4.1. Conditions prior to sampling

A layer of near surface facets topped with surface hoar developed during a period of high pressure prior to January 22 2004. As early as January 10 the local avalanche center received reports of "nice surface hoar" (Chabot, 2004). Snowfall began on January 23, with approximately 25 cm of snow by January 24 (GNFAC, 2004).

We sampled the alleys on January 26 2004 as the weather cleared and Plot 1 three days later. Sampling of the remaining plots occurred at near weekly intervals, with Plot 2 sampled on February 5, Plot 3 a week later on February 12, and Plot 4 sampled on February 20. Slab properties, weak layer strength and stability, and per day changes are summarized for the five sample days (Tables 1 and 2, and Figure 3).

4.2. The Alley, January 26

No evidence from the alley samples suggested that any site-wide trends existed across the four plots. We found no significant linear trend in either the shear strength or stability data of the Alley sample. Therefore, we assumed similar initial conditions in all four plots.

4.3. Plot 1, January 29

Warm temperatures rapidly settled and consolidated the slab in the three days between the Alley and Plot 1 samples (Table 2). The surface hoar layer remained an avalanche hazard and a few avalanches were released in steeper, wind loaded terrain by the nearby Big Sky Mountain Resort ski patrol (GNFAC, 2004). At the study site it appeared that as the slab settled, the layer of facets above the surface hoar layer interpenetrated the surface hoar grains.

The increases in both Plot 1 strength and stability were significant ($p < 0.001$) when compared to the Alleys. The QCV increased slightly, indicating an increase in the variability of shear strength. The shear frame test with the lowest shear strength was within 0.5 m of one of the tests with the highest shear strengths, suggesting spatial correlation only at short distances. Such close proximity of very high and very low test results has been noted in other studies (Landry, 2002).

On a per day basis, the change in median stability was an order of magnitude greater than observed between other sample dates (Table 2). This rapid increase occurred as the weak layer adjusted to the load of new snow. The increase in strength could have been accelerated as small facets penetrated the surface hoar layer, increasing the bonding and hence strength within the layer.

4.4. Plot 2, February 5

Light snowfall occurred throughout the week between sampling of Plots 1 and 2, increasing the median slab thickness to 47 cm. When we sampled Plot 2, grains in the oldest slab layers had begun to round. Again, small facets occurred between the surface hoar grains.

The difference in strength between Plot 1 and 2 was significant ($p < 0.001$). The QCV of shear strength decreased when compared to Plot 1, indicating a slight decrease in the relative variability of the test results. The slight decrease in median stability between Plot 1 and 2 was not significant ($p = 0.211$). Although shear strength increased rapidly between Plot 1 and 2 (Table 2), the additional snowfall increased the load on the weak layer by a proportional amount, and little net change in stability resulted.

4.5. Plot 3, February 12

Snowfall occurred for several days following the sampling of Plot 2, with the heaviest snowfall occurring on February 8 and 9. Observers noted several natural avalanches in the vicinity of the study site, and surrounding ski patrols released several avalanches, but the reports did not mention if the avalanches failed on the surface hoar layer we were sampling (GNFAC, 2004).

Strengthening continued between Plots 2 and 3 at a decreased rate compared to rates measured earlier (Table 2). The QCV of strength increased between Plots 2 and 3, and was higher than the earlier samples, indicating increased variability of shear strength as the layer aged.

Median stability continued to decrease between the plots as stress from the slab continued to increase (Table 2), but the field crew did not feel that the avalanche danger had increased. Other, easier but poorer quality failures occurred on interfaces above the surface hoar layer in stuffblock and rutschblock tests. The failures above the surface hoar layer

Table 1. Summary of slab properties, shear strength, and stability ratios for the five days of sampling.

	ALLEY	PLOT 1	PLOT 2	PLOT 3	PLOT 4
Median Slab Thickness (cm)	45	34	47	66	72
QCV	0.018	0.002	0.011	0.023	0.014
Median Slab Density (kg m^{-3})	98	129	157	148	183
QCV	0.041	0.034	0.015	0.024	0.011
Size of surface hoar grains (mm)	6-8	4-6	6	6	4
Median Shear Strength (Pa)	764	1049	1648	2118	2256
QCV	0.091	0.093	0.060	0.096	0.107
Median Stability Ratio	2.3	3.5	3.4	3.1	2.5
QCV	0.092	0.099	0.067	0.108	0.111

indicated that, while the surface hoar layer was still a critical layer, it was no longer the *most* critical layer in the snowpack.

4.6. <u>Plot 4, February 20</u>

Snowfall resumed between sample days, with 15-20 cm of dense snow falling on February 18 (GNFAC, 2004).

Strength continued to increase significantly between Plots 3 and 4 ($p < 0.001$). The rate of change in shear strength between Plots 3 and 4 continued to decrease (Table 2). The decrease in the rate of change is consistent with previous experience (Chalmers and Jamieson, 2003; Jamieson and Johnston, 1999; Jamieson and Schweizer, 2000; McClung and Schaerer, 1993; Tremper, 2001). The QCV continued to increase, suggesting that shear strength continued to become more variable.

Table 2. Per-day change in median strength, stability, and slab properties.

Between	Alley and Plot 1	Plots 1 and 2	Plots 2 and 3	Plots 3 and 4
Number of days	3	7	7	8
Change in median shear strength (Pa day^{-1})	95	86	67	17
Change in median stability (day^{-1})	0.42	-0.01	-0.05	-0.07
Change in median slab thickness (cm day^{-1})	-3.50	1.86	2.71	0.75
Change in mean slab density (kg cm^{-3} day^{-1})	10.07	3.99	-1.15	4.36
p value, change in strength	< 0.001	< 0.001	< 0.001	< 0.001
p value, change in stability	< 0.001	0.211	< 0.001	< 0.001

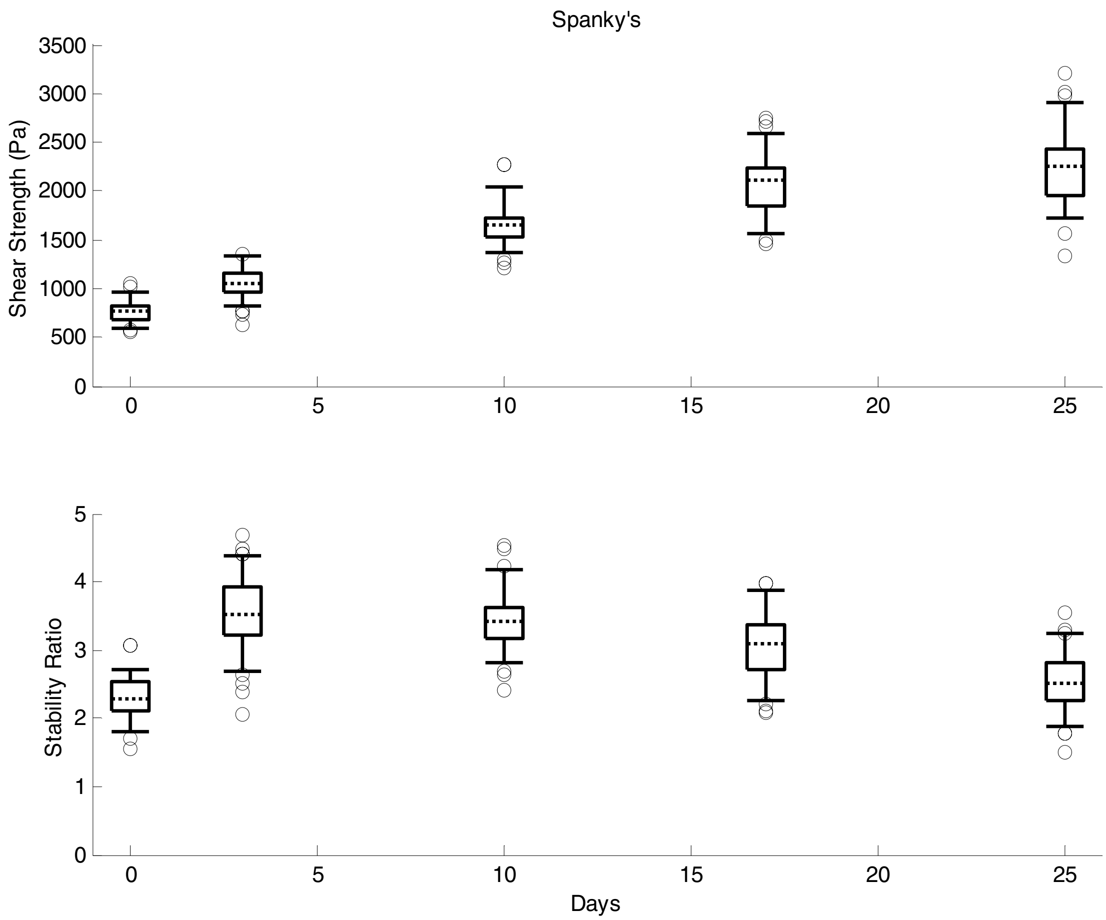

Figure 3. Box plots of shear strength (top) and stability ratios (bottom) through time. Dotted lines indicate plot medians, boxes indicate the 0.25 and 0.75 quantiles, whiskers the 0.05 and 0.95 quantiles, and circles the extreme values.

Median stability decreased significantly between Plots 3 and 4 ($p < 0.001$). The decrease in median stability of this layer did not seem to influence the avalanche hazard, and again the field crew felt that the surface hoar layer was no longer the most critical layer. The weaker interfaces above posed a greater avalanche problem.

4.7. Variogram Analysis

Only shear strength was analyzed using variograms because of problems with stability calculations; we only measured one set of slab properties for each pit, which effectively reduced the variability of the stability results. In future studies, we hope to combine additional measures of slab properties with structural data from the SnowMicroPen.

We found two significant linear trends within the stability data. Neither explained much of the variance, and we did not have enough measurements of slab properties to see if the slab had any spatial trends. Additional structural data, such as from the SnowMicroPen (Schneebeli et al., 1999), could be utilized to see if structural trends within the slab existed (Kronholm et al., in press), or if the linear trend was an artifact of the data collection and analysis.

The four variograms (Figure 4) are "noisy," with variations rather than increasing smoothly. None of the variograms indicated strong spatial correlation, and with the exception of Plot 1, all had very large nuggets. Several factors may cause large nuggets. One cause could be that the correlation distance was short relative to the spacing of the measurements. The lack of data at shorter distances would mean that the correlation length is not well resolved. The additional, closely spaced tests from Plot 1 supported this possibility.

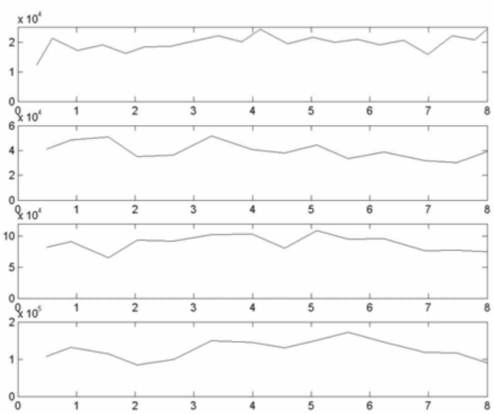

Figure 4. Variograms of shear strength for the four plots, from the top down, Plots 1, 2, 3, and 4. The horizontal scale, distance in meters, is common to all four variograms. The y axis, semivariance, differs between variograms as the semivariance increased through time.

Semivariance at the shortest distances, 0.4 m, on the Plot 1 variogram indicated a spatial pattern with correlation shorter than what could be resolved with the test array used on the other three plots. Additional support for correlation at short distances came from experience through the field season. If the operator felt a test was faulty, a second test was often placed as closely as possible to the first test. Unless the initial fault was due to an improperly prepared test, the second result tended to be more similar to the "faulty" test than two tests at the standard distance of 0.5 m. The test layout was designed to capture spatial correlation at distances of several meters. From this, it appears that there was significant variability and that there might have been correlation at distances less than 1 m over this slope. We hope to address correlation at short distances in future research.

Large nuggets could also indicate noise in the measurements, or variability that the measurements cannot resolve (Bloschl, 1999; Myers, 1997). Our coefficients of variation ranged from 12 to 24%, which compared favorably with the coefficients of variation ranging from 3% to 66% (with a mean of 15%) reported by Jamieson and Johnston (2001) for 809 sets of shear frame measurements. This suggests that the large nuggets could be due to the test itself, and that there is considerable variation in shear strength that the shear frame test cannot resolve.

4.8. Temporal Changes in Spatial Variability

There were no indications of site-wide trends in the Alley samples. We therefore assumed that the plots were initially similar, and attributed differences between the Plots to temporal change.

Though we remain cautious about our spatial analysis given the quality of the variograms, there are a few points worth discussing. First, the nugget increased from Plot 1 to Plot 2 to Plot 3, indicating a larger amount of "noise" in the results and reflecting the lack of tests at close distances in Plot 2 and 3. There was little change in the nugget or range from Plot

3 to Plot 4, suggesting little increase in the noise. Second, in both Plots 2 and 4, the semivarince increased markedly between 2.5 m and 3.5 m. At those distances, tests were no longer within the same pit, so the increase indicated that there was more difference between pits than within pits.

4.9. Pit-to-Plot Ratios

The only pits not representative of the plots were two pits of Plot 1. By the pit-to-plot measure, only Plot 1 demonstrated spatial variability. On all other plots, any individual pit statistically represented the shear strength or stability of the plot. By this measure, all other plots were spatially uniform.

Of the four plots, the Plot 1 variogram indicated the least difference between tests at the inter-pits distances (all the variability occurred at distances within the pits), but Plot 1 had the only pits not representative of the plot. A pure nugget variogram or one that indicated spatial correlation at distances less than 2.5 m (the intra-pit distance) would be more consistent with the pit-to-plot results.

We anticipated that more pits would not be representative of the plot based on Landry's (2002; in press) study. In that work, over a third of the pits were not representative of the plots, while this research had only 2 of 16 pits (13%) that were statistically not representative. Several differences between the studies might explain the differences. One difference was the size of the plots. Our plots were about a fourth of the area of Landry's, with distances between our pits about a fourth of the distance between Landry's pits. Spatial trends at scales that Landry's tests could pick up might be undetectable with our layouts.

The type of test used was another difference between the current study and Landry's work. Landry used the Quantified Loaded Column Test (Landry et al., 2001), which integrates slab characteristics into the test result. The shear frame test removes the slab from the test, and tests only the weak layer. This could account for some of the differences, if trends were present in the slab.

5. LIONHEAD

5.1. Conditions Prior to Sampling

An extensive layer of surface hoar formed during the middle of January 2004 throughout the intermountain west. Reports of "trophy sized" surface hoar came from Utah; in the mountains around the Lionhead study site observers found grains measuring 50 mm (Chabot, 2004; UAC, 2004). Observations made near the study site on January 22 recorded two layers of surface hoar, separated by a thin layer of precipitation particles, and buried under a thin layer of recent snow (Figure 2).

A storm system on January 24-26 buried and preserved the surface hoar layer. Storm totals of 20 cm were reported for the Lionhead area (Chabot, 2004). On the 26th, natural avalanche activity was reported on the avalanche paths near the study site (Chabot, 2004). Stormy weather continued, with daily snowfall of 2 to 10 cm. A widespread avalanche cycle occurred at the end of the month, with an Avalanche Warning issued by the GNFAC January 31 through February 2 (GNFAC, 2004).

Small snowstorms continued to deposit snow through the week prior to the sampling of the alleys. Avalanche conditions remained interesting: many slopes had slid on the surface hoar layer, effectively removing it as a hazard on those slopes. Slopes that had not avalanched remained quite sensitive to human triggering, and posed a hazard to backcountry travelers. On February 3, an observer reported that the surface hoar layer was so sensitive they "did not want to get on anything over 30°" (Chabot, 2004). Stability tests conducted by the observer on slopes near the study site failed easily on the surface hoar layer (SB 30 Q1, CT 16 Q1).

5.2. The Alleys

The alleys were sampled February 7th, 2004. Avalanches adjacent to the study site allowed for safe access to the site. Avalanche debris had come within 3 m of the eastern edge of the site, along the outside margins of Plots 2 and 3. The avalanches probably ran during or after the January 26 avalanche cycle. Recent snow made for good skiing on the bed surface of the old avalanches, and composed about half of the sampled slab. The slab was 53 cm thick, with a median density of 151 kg m^{-3} (Table 1).

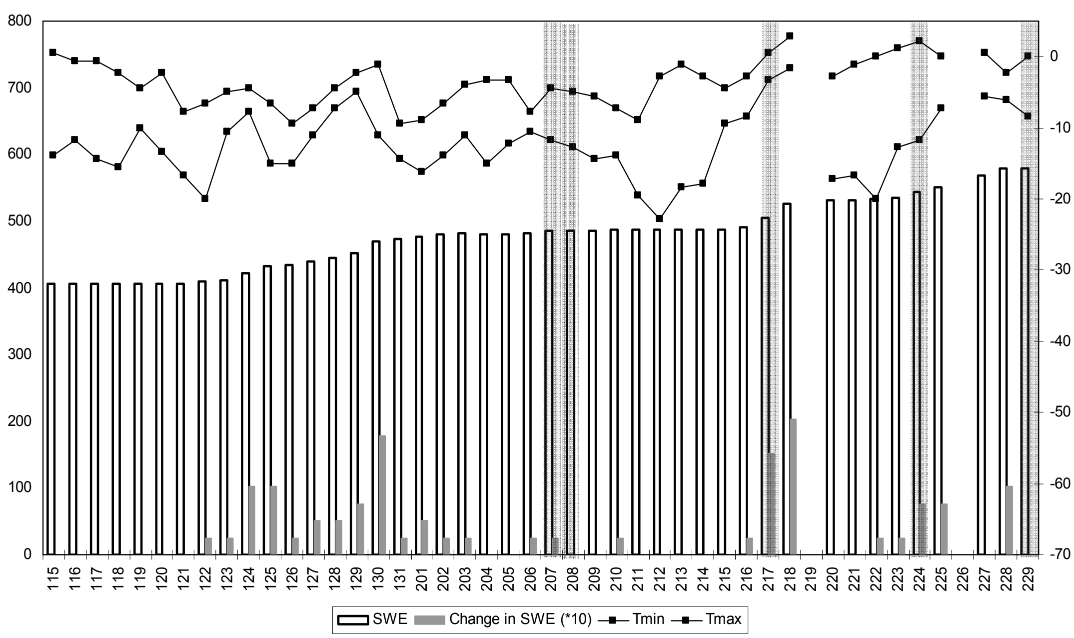

Figure 1. Plot of weather from the Madison Plateau Snotel site, located 20 km SE of the study site at an elevation of 2360 m. Snow depth is not available from the site, so snow water equivalent (SWE) is shown instead.

Figure 2. Close up photo of the double surface hoar layers, taken after the collapse while samplingthe Alleys. Ruler marked in centimeters.

Table 1. Slab and weak layer properties of the Lionhead plots.

	ALLEY	PLOT 1	PLOT 2	PLOT 3	PLOT 4
Median Thickness (cm)	53.25	53.00	61.00	61.00	86.00
QCV	0.012	0.009	0.008	0.016	0.006
Median Density (kg m^{-3})	151.40	158.49	181.67	213.11	209.30
QCV	0.051	0.021	0.018	0.030	0.011
Median Shear Strength (Pa)	774.72	1363.12	2000.55	2550.81	2927.27
QCV	0.440	0.125	0.130	0.108	0.093
Normality of Strength	< 0.001	0.5	0.09	< 0.001	0.092
Median Stability	1.37	2.29	2.38	2.65	2.45
QCV	0.440	0.125	0.130	0.124	0.104
Normality of Stability	< 0.001	0.5	0.5	<0.001	0.086

The first shear frame pits and SMP measurements indicated that the surface hoar layer had not collapsed, eliminating the field crew's concern that the surface hoar layer had collapsed across the site when the adjacent avalanches ran. About one quarter of the way into the sampling, the field crew collapsed one of the surface hoar layers. The collapsed occurred between shear frame measurement 2 and 3 in Pit 4. It collapsed with a thunderous whumph and sounded as if it propagated a considerable distance. The propagation distance turned out to be relatively short; possibly the thickness of the slab and the thickness of the surface hoar layer contributed to the loud sound (because I was down in the pit sampling, I heard the collapse it from within the snowpack, which made it even

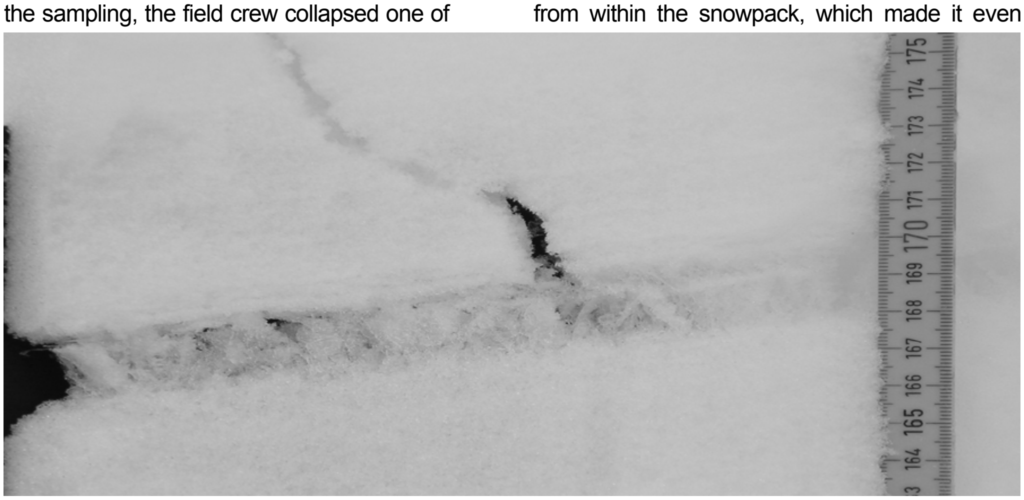

Figure 3. Close up photo of the tensile crack near the lower right corner post of Plot 1.

more dramatic). A tensile crack opened up across the site, arcing from the corner posts of the cross-slope alley, across and up to the upper corner posts of the up-slope alley. The corner posts, excavated earlier that morning, had become areas of stress concentration.

The collapse occurred within the upper of the two surface hoar layers, leaving the lower surface hoar layer upright and intact (Figure 2). Prior to the collapse, it was hard to determine which of the surface hoar layers failed in the shear frame tests, because both layers were disrupted as the test failed. After the collapse, the shear frames tended to fail on the upper, collapsed layer.

The shear strength measured by the shear frame tests changed dramatically after the collapse ($p < 0.001$). The median shear strength prior to the collapse was 1697 Pa, with median stability of 3.08 (Table 1). Post collapse, the median shear strength was 657 Pa, and median stability 1.16. I made two measurements above the fracture at the upslope end of the alley, extending pit 6. These two tests, 25 and 26, were more than 1000 Pa stronger than the adjacent tests below the fracture.

The post collapse median is somewhat misleading, because shear strength (and therefore stability) increased with time. Using test order as a proxy for time, increase in shear strength was significant ($p < 0.0143$; $R^2 = 0.174$). The collapse, and the increasing shear strength as the day continued, is apparent in Figure 4 . Sometimes, observers note that shear strength increases dramatically after the collapse of a weak layer. With this layer, the presence of the second, un-collapsed surface hoar layer may have confounded the trend. The layers did gain strength rapidly, as can be seen from the increase through time and the much smaller difference between collapsed and un-collapsed tests in Plot 1.

It was not possible to determine from the Alleys if any site-wide trends existed. The number of pre-collapse tests was insufficient, and concentrated within the lower limb of the up-slope alley, and the post collapse trends changed through the sample period.

Table 2. Significance and correlation of linear trends of Lionhead plots. Indicated in bold are trends that were significant, explained more than 10% of the variance, and were removed from the variogram analysis.

		Plot 1	Plot 2	Plot 3	Plot 4
Strength	P value	0.002	0.3117	< 0.001	0.003
	R^2	0.005	0.239	0.128	0.063
Stability	P value	0.005	0.017	< 0.000	0.0906
	R^2	0.116	0.085	0.289	0.04

5.3. *Plot 1 (CO)*

We sampled Plot 1 (CO) February 8, the day following the sampling of the alleys. The presence of the tensile crack dictated our choice of the plot, because we wanted to compare collapsed and un-collapsed tests. The slab changed little overnight, with a median thickness of 53 cm, and median density of 159 kg m^{-3} (Table 1).

The location of the tensile crack was still apparent, so tests could be classified as un-collapsed if above the crack, and collapsed if below. Although the collapsed and un-collapsed still differed ($p = 0.061$), the difference was much less than the day before. The median shear

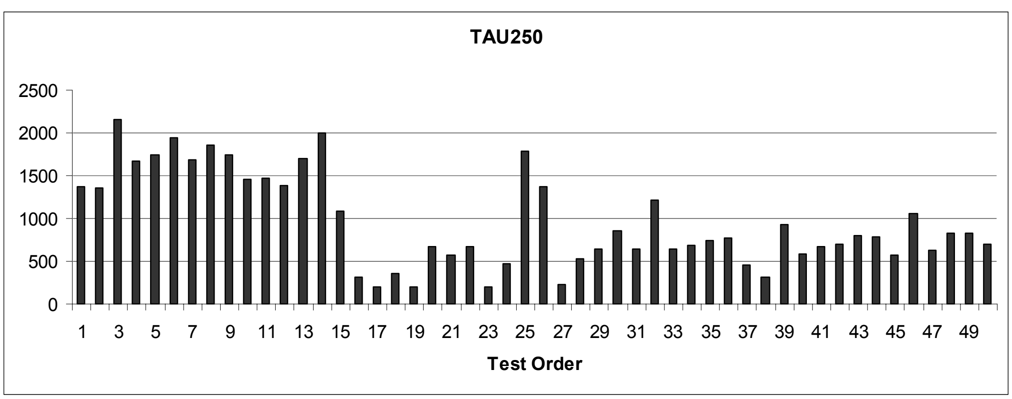

Figure 4. Plot of test number and shear strength. The collapse occurred between tests 15 and 16. Tests 25 and 26 were made above the tensile crack.

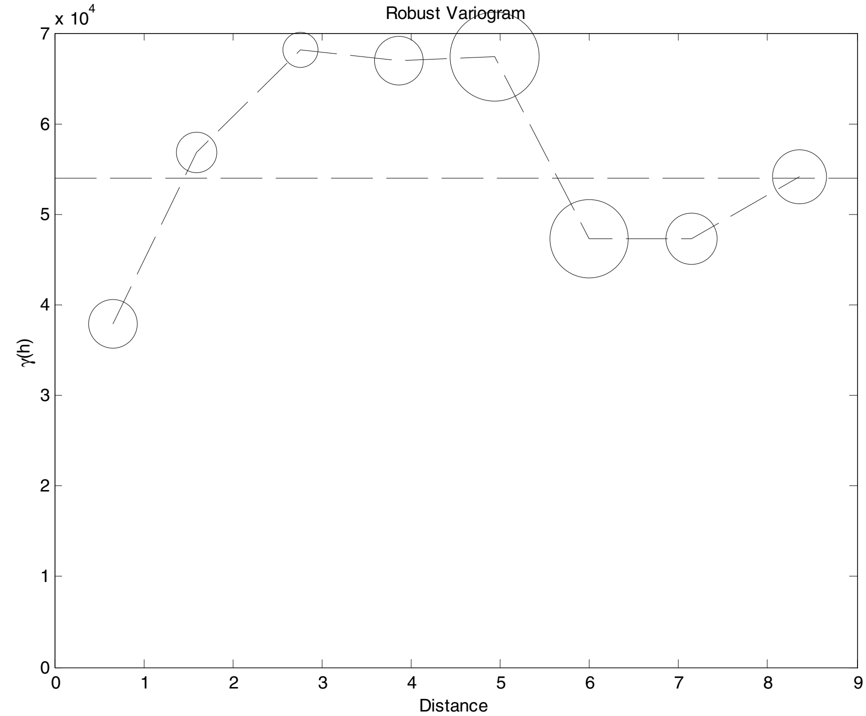

Figure 5. Variogram of Lionhead Plot 1, plotted with a lag distance of 1.1 m.

strength of the un-collapsed tests was 1432 Pa, with median stability of 2.32 (Table 3). For collapsed tests, the median shear strength was 1314 Pa and median stability of 2.26 (Table 3). The post-collapse strengthening observed in the alley sample continued through the night. It appeared that the un-collapsed areas weakened, although the presence of the double surface hoar layer complicates interpretation because it was had to tell in which layer the shear frames fractured.

Table 3. Shear strength (Pa) and stability for collapsed and un-collapsed tests.

	ALLEY	PLOT 1
Median shear strength, un-collapsed	1697	1432
Median shear strength, collapsed	657	1314
Median stability, un-collapsed	3.08	2.32
Median stability, collapsed	1.16	2.26

A significant linear trend was present in the shear strength data, but it explained only 0.5% of the variance so it was not removed for analysis (Table 2). The trend indicated that tests at the top of the plot tended to be stronger. The un-collapsed tests occurred in the upper portion of the plot, and potentially explained the trend.

The variogram (Figure 5. Variogram of Lionhead Plot 1, plotted with a lag distance of 1.1 m.), plotted with a lag distance of 1.1 m, had a nugget estimate of 0.46. Semivariance decreased between 5 and 6 m. The decrease could be interpreted as a waveform, but could also represent comparisons between tests across the tensile fracture.

The range was just less than 2 m, within the intra-pit distance. This relatively short correlation distance meant that the entire spatial pattern occurred within the individual pits. This was reflected in the pit-to-plot ratios, with all pits representative of the plot. Because there was no significant spatial structure at the inter-pit distances, any pit should represent the plot.

Table 4. Pit-to-plot ratios of Lionhead plots for strength and stability. Pits that were not representative of the plots are indicated in bold.

STRENGTH	Plot 1	Plot 2	Plot 3	Plot 4
Pit 1	0.3182	0.8461	0.1987	0.4778
Pit 2	0.4301	0.1247	0.3694	0.8719
Pit 3	0.3000	0.0693	0.6606	0.6397
Pit 4	0.1866	0.0008	0.1650	0.0185
Pit 5	0.7291	0.8780	0.2518	0.2762
STAB	Plot 1	Plot 2	Plot 3	Plot 4
Pit 1	0.1791	0.9039	0.9807	0.5195
Pit 2	0.0808	0.4541	0.0022	0.9679
Pit 3	0.5628	0.0027	0.5626	0.5765
Pit 4	0.0248	0.0956	0.0148	0.0329
Pit 5	0.7053	0.1841	0.2633	0.2807

5.4. <u>Plot 2 (UT)</u>

Light snowfall occurred on February 9-12th with storm totals estimated at 15 cm, and again on the 15th and 16th with estimated storm totals of 10 cm (GNFAC, 2004). On the 17th, when Plot 2 (UT) was sampled, the slab had a median thickness of 61 cm, with median density of 182 kg m^{-3}. The median shear strength increased to 2000 Pa, with a QCV of 0.130 and median stability was 2.38, with a QCV of 0.130 (Table 1).

Collapsed and un-collapsed tests could not be definitively separated by shear strength because snowfall in the intervening week obscured all traces of the tensile crack. The crack was somewhat symmetrical around the up-slope alley, so Pit 3 was probably above the crack and un-collapsed, while Pits 1.5, 2, 3, and 3.5 were probably below the crack and collapsed. The location of the other pits relative to the crack could not be completely determined. There were not obvious differences in values within Pits 1 and 5 to differentiate where the crack may have been located.

A significant linear trend existed within the shear strength across the plot, with values near Pit 4 having a greater probability of lower

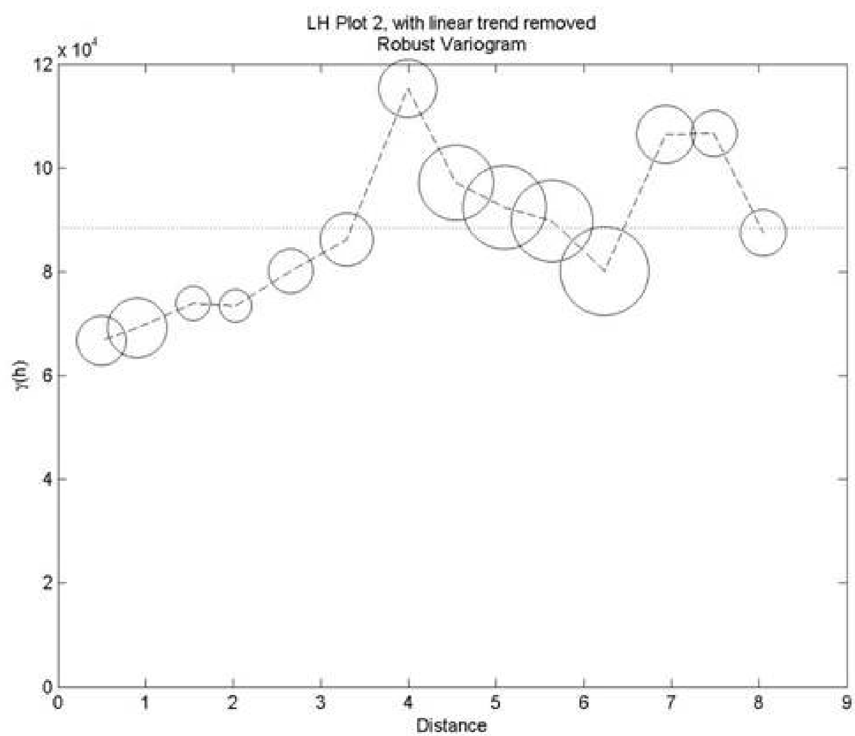

Figure 6. Variogram of Lionhead Plot 1, plotted with a lag distance of 0.6 m.

strength than areas near Pit 2. The trend probably reflected the location of the tensile crack because tests below the tensile crack tended to be stronger than tests above. The collapsed surface hoar layer may have contributed to the strengthening by acting in a manner similar to an ice lens. The force of the collapse may have compacted the lower surface hoar layer, and then increased the rate and strength of bonds between the surface hoar grains. Unfortunately, reliable SMP measurements were not available from this Plot, so detailed structural differences between collapsed and un-collapsed areas can only be hypothesized.

The variogram of the de-trended data, Figure 6, was plotted with a lag distance of 0.6 m, had a range of 4 m and nugget estimate of 0.7. Within pits, shear strength exhibited more correlation than results between pits.

The pit-to-plot ratios suggested some inter-pit differences. Pit 4 was not representative of the plot strength ($p < 0.0001$) and there was some evidence that Pit 3 was not representative ($p = 0.069$; Table 4). Pit 3 was not representative of plot stability ($p = 0.0027$; Table 4). As noted above, the tensile crack likely ran below or across the corner of Pit 4, and tests above the crack tended to be weaker than tests below. Some evidence existed that Pit 3 might not be representative of the plot ($p = 0.07$; Table 4). With one exception, shear strength in pit 3 was quite high.

5.5. Plot 3 (AZ)

A warm storm occurred on February 18th and 19th , with rain at the lower elevations and 30 to 45 cm of dense snow at the upper elevations (GNFAC, 2004). The minimum temperature recorded at the Madison Plateau Snotel site was -1.6° C for the 18th. The station then experienced problems, so no data was recorded for the 19th. A period of high pressure, with clear skies and colder temperatures dropping to -20° C on the 22nd, followed the storms. The weather began to change on February 24th with snowfall when Plot 3 was sampled. On February 24, the median slab thickness was 61 cm, with a median density of 213 kg m^{-3}. Although slightly more variable, the slab thickness did not differ markedly from the Plot 2. Density did increase between Plot 2 and 3.

Median stability of Plot 3 was 2.65, with a CV of 0.124, and median shear strength was 2551 Pa, with a QCV of 0.108 (Table 1). The shear strength in Plot 3 was quite skewed, with several very strong measurements. The very strong shear frame measurements occurred at the top of the plot. There was a significant linear trend across the site ($p < 0.001$; $R^2 = 0.123$). Shear strength tended to be lower at the bottom of the site than at the top.

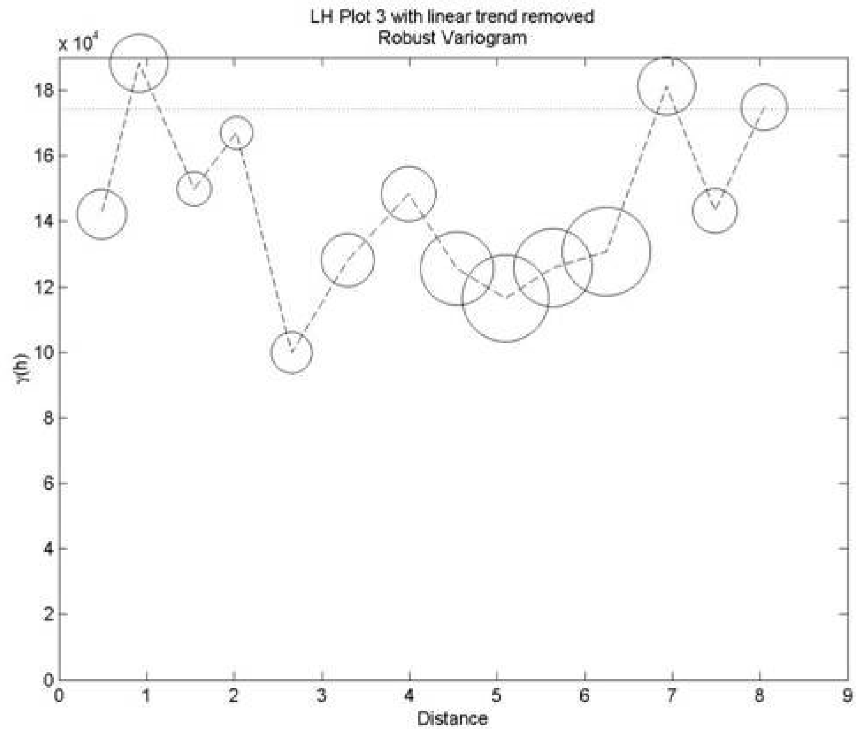

Figure 7. Variogram of Lionhead Plot 3, plotted with a lag distance of 0.6 m.

The trend was removed from the variogram plotted in Figure 7 with a lag distance of 0.6 m. The variogram indicated negative spatial correlation at all but the two shortest lag distances. The slight increase between the 0.6 and 1.1 m bins could be an indication of some spatial correlation at shorter distances.

The pit-to-plot ratios for shear strength indicated that all pits were representative of the plot. Results for stability, however, differed. Pits 2 and 4 were not representative of the plot ($p = 0.002$, $p = 0.0145$ respectively; Table 4). Because stability data were calculated from a single set of slab properties for each pit, spatial variations in slab properties may not have been reflected in the stability data.

Structural data from SMP measurements provide some interesting additional information. Slab thickness was well correlated, and the thickness was kriged to coordinates of the shear frame tests (Lutz, 2004). There was a significant linear trend within the kriged slab thickness ($p < 0.001$, $R^2 = 0.764$; Table 2), with the slab at the top of the plot tending to be thicker than the slab at the bottom by 1 or 2 cm. Other studies have found slab structure more correlated than weak layer structure, and trends in slab characteristics and strength in similar directions (Birkeland et al., in press; Kronholm, 2004; Kronholm et al., in press).

I had hoped to be able to relate slab thickness to density, and recalculate stability for each test using the kriged thicknesses. Unfortunately, there was not enough correlation within the nine slab thickness and water content measurements to build a relationship between the two ($p = 0.38$, $R^2 = 0.107$). Recalculating the stability ratios using kriged slab thickness and the measured water content changes the calculated stability ratio by no more than 0.25.

5.6. Plot 4 (NM)

Another storm moved through southwestern Montana on February 25-29. Reports of storm totals of 35 cm came from the Lionhead area (GNFAC, 2004). Avalanches occurred in the new snow—probably running on near-surface facets formed February 20-24th—but the GNFAC did not report any avalanches fracturing on the January surface hoar layer. Stability tests conducted on February 27, on a slope similar to the study site, produced very hard or no results on the surface hoar layer (Chabot, 2004). Only trace amounts of snowfall were reported between February 29 and March 2, when Plot 4 was sampled (GNFAC, 2004). Slab thickness increased to 86 cm, with a QCV of 0.006, and median density was 209 kg m^{-3}, with a QCV of 0.011 (Table 1). The slab was the most uniform of the five sample days.

Median shear strength increased to 2927 Pa, with a QCV of 0.093. A significant linear trend existed within the shear strength data, but explained too little of the variance (3%) to be removed form the data (Table 2). Median stability of the plot was 2.45, with a QCV of 0.104. The non-spatial QCVs indicated that Plot 4 was the most uniform of the plots.

The variogram plotted with a lag distance of 1.1 m (Figure 8), supported that conclusion to some extent. Plot 4 had the largest range of the four plots, near 4.5 m. Considerable variation occurred at distances less than 4.5 m, indicated by the nugget estimate near 1, and high semivariance at distances less than 4.5 m. Even thought the range was 4.5 m, the variogram was similar to a pure nugget variogram, indicating that what little correlation existed was not strong.

With such little spatial structure indicated by the variogram, all pits could be expected to be representative of the plot in the pit-to-plot ratios. However, Pit 4, with the highest median pit strength, was not representative of either plot strength ($p = 0.019$) or stability ($p = 0.033$).

Slab thickness kriged from SMP data (Lutz, 2004) again indicated a significant slope-scale trend ($p = 0.012$, $R^2 = 0.121$). The slab tended to thicken slightly towards the top of the plot. Unfortunately, as in the case of Plot 3, there was not a strong relationship between the measured slab properties ($p = 0.162$, $R^2 = 0.260$) so stability ratios could not be recalculated using kriged values.

5.7. Temporal change

The collapse of the upper surface hoar layer allowed the comparison of both a weak layer and a collapsed weak layer over a ten-day period, although the doubled surface hoar layer added some complication. The results of the collapsed layer while sampling the alley increased significantly through the day. We were able to capture the rapid strengthening of a broken weak layer, or at least the layer adjacent. A day later, there was little difference in test results between collapsed and un-collapsed tests.

The change in median shear strength

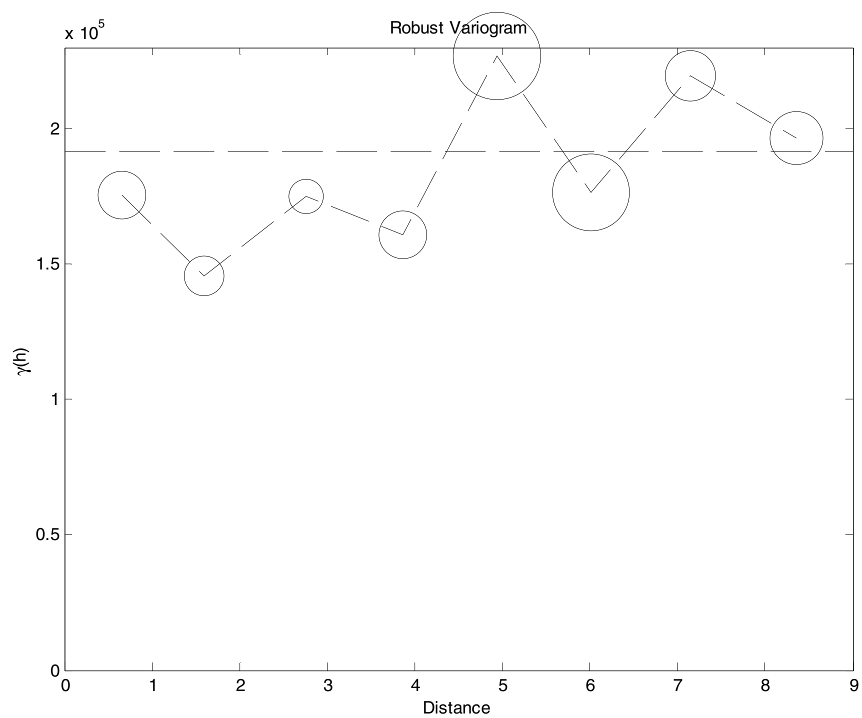

Figure 8. Variogram of Lionhead Plot 4, plotted with a lag distance of 1.1 m.

should not be directly compared between the Alleys and Plot 1. Instead, I compared the differences between the collapsed and un-collapsed tests on each day. The difference in shear strength between un-collapsed tests of the Alley to Plot 1 was significant ($p = 0.005$), as was the difference between collapsed tests ($p < 0.001$). The lack of significant difference in shear strength of collapsed and un-collapsed tests in Plot 1 ($p = 0.061$) could have been due to the presence of the second, un-collapsed surface hoar layer. It was quite hard to determine in which of the two surface hoar layers the shear frames failed. Failure in the lower layer that had not collapsed could explain the smaller difference between tests on Plot 1.

The differences in both shear strength and stability between Plot 1 and 2 were significant ($p < 0.001$ fro strength, $p = 0.037$ for stability; Table 5). The surface hoar layers continued to strengthen through time and tests above and below the tensile fracture could not be distinguished by eye or in the field. The linear trend fit did reflect the presence of the crack, but explained little of the variance across the plot. Although the change in stability was slight, the median increased enough to be significant.

The increase in range between Plot 1 and 2 suggested divergence. So did the decrease in the pit-to-plot ratio and slight increase in the QCV of shear strength.

The increases in shear strength between and stability between Plots 2 and 3 were significant ($p < 0.001$). The increase in shear strength was greater than the proportional increase in shear stress. Although the slab thickness did not change, the density increased. Divergence would be expected, based on my hypothesis, as the layer gained strength without significant change in the slab. Unfortunately, the Plot 3 variogram did not allow much interpretation, although the range decreased between the two. The pit-to-plot ratio for strength increased, and the ratio for stability decreased. The QCVs of both strength and stability deceased. The changes between the plots indicated both divergence and convergence, with no one indicator providing strong evidence of the temporal trend.

The increase in shear strength between

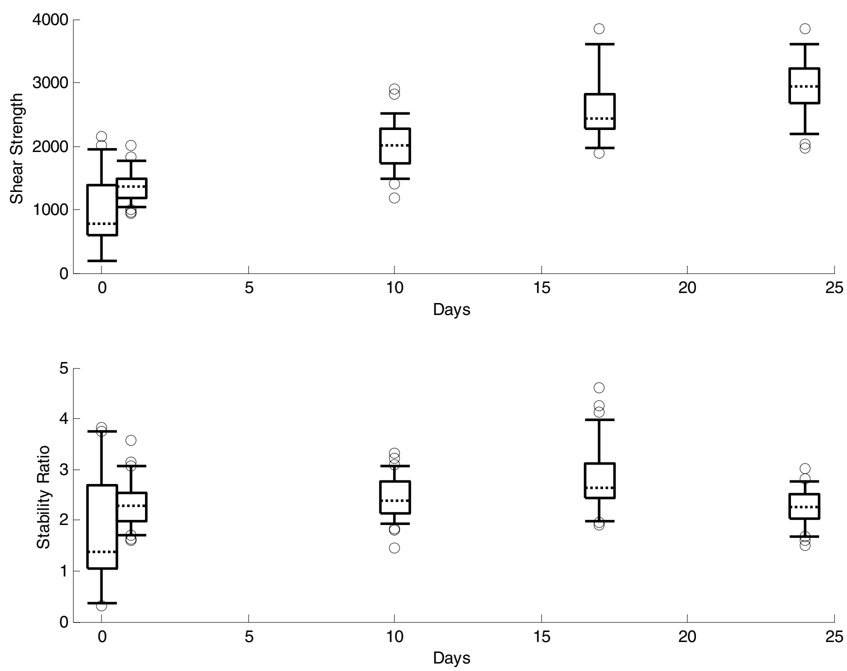

Figure 9. Timeline plot of Lionhead shear strength data

Plots 3 and 4 was significant ($p < 0.001$). The decrease in stability between Plots 3 and 4 was also significant ($p < 0.001$). The QCVs of both strength and stability decreased, indicating convergence. The range increased between the variograms, but both variograms contained considerable noise and provided little evidence of change. Again, the pit-to-plot ratios indicated opposite trends, divergence indicated by the pit-to-plot ratios of strength, and convergence indicated by the pit-to-plot ratios of stability.

Table 5. Per day change in weak layer and slab properties between Lionhead plots.

	X and Plot 1	Plots 1 and 2	Plots 2 and 3	Plots 3 and 4
Number of days	1	9	7	9
Change in median shear strength (Pa d^{-1})	588	71	79	42
Change in median stability	0.92	0.01	0.04	-0.02
Change in median slab thickness (cm d^{-1})	-0.25	0.89	0.00	2.78
Change in mean slab density (kg m^{-3} d^{-1})	7.09	2.58	4.49	-0.42
P value, change in strength	<0.001	<0.001	<0.001	<0.001
P value, change in stability	<0.001	0.037	<0.001	<0.001

6. CONCLUSIONS AND IMPLICATIONS

This study focused on three questions about spatial structure, temporal change, and causes of the change. Spatial patterns in the data proved elusive. Our variograms did not indicate strong spatial correlation. There may be several reasons; primary among them that the spatial correlation of shear results occurs at distances less than 1 m. Both Plot 1 and our field experience support such short correlation distances. In the coming winters, we plan to sample short distances more thoroughly.

Another possibility is that we chose such uniform sites that we removed all spatial trends, leaving a nearly uniform slope that had no correlation or trends. If this were the case, the expected variograms would be pure nugget. Our variograms were quite similar to pure nugget variograms, but they did exhibit some spatial structure.

Our sampling array could be another reason we found little spatial correlation in the data. The array was designed to balance pit-to-plot ratios, geostatistical analysis, and the number of tests we could conduct on one day. Because geostatistical analysis proved more promising than pit-to-plot ratios at the scale of our study sites, arrays in the coming winters will better optimize the variograms, and ignore pit-to-plot ratios.

Since our spatial analyses at this site did not show conclusive trends or spatial structure, we can make few conclusions about temporal changes in spatial variability. However, the significant changes in non-spatial measures of variability could be easily related to weak layer aging and changes in the slab-induced shear stress.

Shear strength and the relative spread of the results increased through time as the weak layer aged and strengthened. This increase in spread across what our analyses show are relatively spatially uniform sites demonstrates how evaluating stability becomes more difficult and less reliable as the weak layer ages. We were able to follow the surface hoar layer from burial through strengthening to the point that it was no longer a critical weak layer for avalanching. Changes in stability were related to changes in the slab thickness, density, and weak layer aging.

One issue that may be applicable to more than just this layer and slope is the correlation of adjacent tests. Commonly, stability tests are conducted adjacent to each other within a pit. If the correlation length of shear strength is, as indicated by this study, shorter that 1 m, results of these adjacent tests would be related, and under-represent the potential variability of stability or shear strength. To represent the potential variability adequately, tests would need to be spaced at distances greater than the correlation length, and a sufficient number of tests conducted to statistically represent the variability.

The analysis allowed us to examine the ability of a forecaster to extrapolate stability tests. On this layer, over this uniform slope, stability results could be statistically extrapolated at least 17 m—provided sufficient tests were conducted to properly characterize the distribution of test results.

How many tests are sufficient to characterize the plots? Sample size can be calculated many ways (there are easy to use sample size calculators at www.stat.uiowa.edu/~rlenth/Power/). A minimum of three uncorrelated tests would be necessary to insure, at a 95% confidence level, that the minimum stability ratio of the Alleys was not less than 1.5. By the time Plot 4 was sampled, ten tests would be necessary. The margin of error is quite large with so few samples. To reduce the margin of error to within one standard deviation, 15 tests would be necessary under conditions similar to the Alley, and 18 under conditions similar to Plot 4. As the weak layer ages, more tests are required to adequately characterize the distribution of results. The test results must not be correlated, so, according to this study on this single slope, they should be spaced as much as a meter or more apart.

Would this help to evaluate the avalanche hazard near the study area? Yes, but not directly. Because the surrounding, avalanche prone slopes are steeper and more wind affected than the study plot, results from the study area would need to be interpreted, just as most forecasting requires.

7. ACKNOWLEDGEMENTS

Many thanks to Eric Lutz for all his help in the field in the summer and winter, for his assistance in organizing field days, and for his discussions on the ideas in this paper. We also thank Shannon Moore for his assistance in the field. Linda Bishop and Grant Estey facilitated site access. This research was financially supported by National Science Foundation (Grant #BCS-0240310). Spencer Logan was partially supported by the Barry C. Bishop Scholarship for Mountain Research and the Department of Earth Sciences at Montana State University.

8. REFERENCES

Birkeland, K.W. and Landry, C.C., 2002. Changes in spatial patterns of snow stability through time, International Snow Science Workshop, Penticton, Canada.

Bloschl, G., 1999. Scaling issues in snow hydrology. Hydrological Processes, 13(14-15): 2149-2175.

Chabot, D., 2004. Gallatin National Forest Avalanche Center Backcountry Observations.

Chalmers, T.S. and Jamieson, J.B., 2003. A snow-profile-based forecasting model for skier-triggered avalanches on surface hoar layers in the Columbia Mountains of Canada. Cold Regions Science and Technology, 37(3): 373-383.

Conway, H. and Abrahamson, J., 1984. Snow Stability Index. Journal of Glaciology, 30(106): 321-327.

GNFAC, 2004. Gallatin National Forest Avalanche Center Backcountry Avalanche Advisories. Gallatin National Forest Avalanche Center.

Jamieson, B. and Johnston, C.D., 1999. Snowpack factors associated with strength changes of buried surface hoar layers. Cold Regions Science and Technology, 30(1-3): 19-34.

Jamieson, B. and Johnston, C.D., 2001. Evaluation of the shear frame test for weak snowpack layers. Annals of Glaciology, 32: 59-69.

Jamieson, J.B. and Schweizer, J., 2000. Texture and strength changes of buried surface-hoar layers with implications for dry snow-slab avalanche release. Journal of Glaciology, 46(152): 151-160.

Kronholm, K., 2004. Spatial Variability of Snow Mechanical Properties with regard to Avalanche Formation, Universitat Zurich, Zurich, 177 pp.

Kronholm, K., Schneebeli, M. and Schweizer, J., in press. Spatial variability of penetration resistance in snow layers on a small slope. Annals of Glaciology, 38.

Kronholm, K. and Schweizer, J., 2003. Snow stability variation on small slopes. Cold Regions Science and Technology, 37(3): 453-465.

LaChapelle, E.R., 1980. The Fundamental Processes in Conventional Avalanche Forecasting. Journal of Glaciology, 26(94): 75-84.

Landry, C.C., 2002. Spatial variations in snow stability on uniform slopes: Implications for extrapolation to surrounding terrain. MS Thesis, Montana State University, Bozeman, 248 pp.

Landry, C.C. et al., in press. Variations in snow strength and stability on uniform slopes. Cold Regions Science and Technology.

Landry, C.C., Borkowski, J.J. and Brown, R.L., 2001. Quantified loaded column stability test: mechanics, procedure, sample-size selection, and trials. Cold Regions Science and Technology, 33(2-3): 103-121.

McCammon, I., 2003. Using the shear frames, Snowpit Technologies, Salt Lake City.

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

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

Mock, C.J. and Birkeland, K.W., 2000. Snow avalanche climatology of the western United States mountain ranges. Bulletin of the American Meteorological Society, 81(10): 2367-2392.

Myers, J.C., 1997. Geostatistical Error Management: quantifying uncertainty for environmental sampling and mapping. Van Nostrand Reinhold, New York, 571 pp.

Schneebeli, M., Pielmeier, C. and Johnson, J.B., 1999. Measuring snow microstructure and hardness using a high resolution penetrometer. Cold Regions Science and Technology, 30(1-3): 101-114.

Spiegel, M.R. and 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. and Jamieson, B., 2002. Spatial variability of slab stability in avalanche start zones, 2002 International Snow Science Workshop, Penticton Canada.

Tremper, B., 2001. Staying Alive in Avalanche Terrain. The Mountaineers, Seattle, 284 pp.

Webster, R. and Oliver, M.A., 2001. Geostatistics for Environmental Scientists. Statistics in Practice. John Wiley and Sons, Chichester, 271 pp.