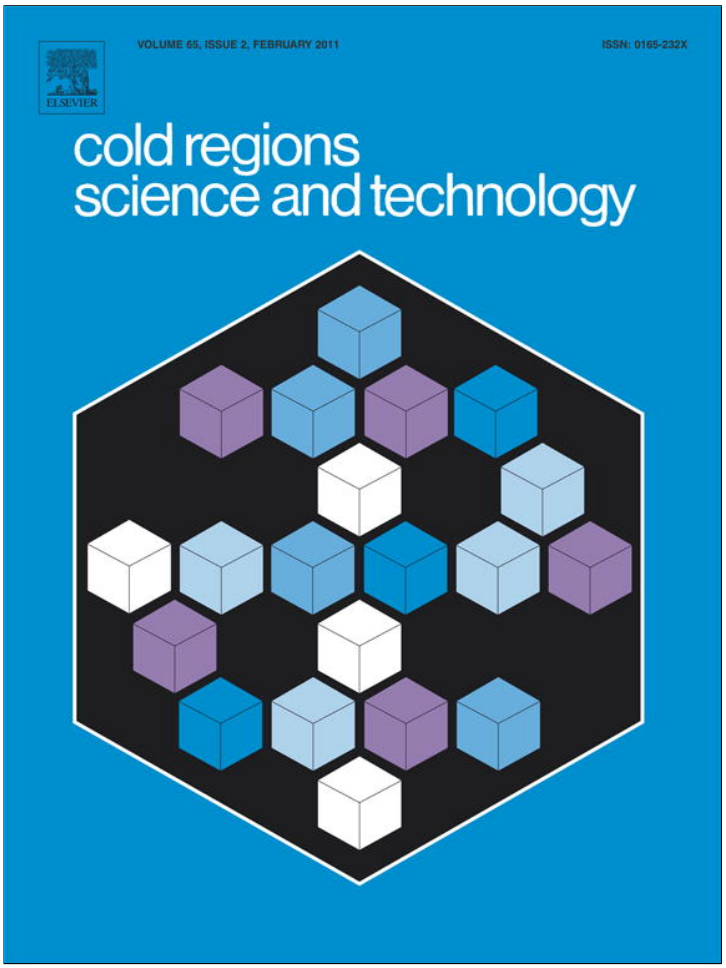

Cold Regions Science and Technology 65 (2011) 234–241

Contents lists available at ScienceDirect

Cold Regions Science and Technology

journal homepage: www.elsevier.com/locate/coldregions

Measuring spatial variations of weak layer and slab properties with regard to snow slope stability

Sascha Bellaire *, Jürg Schweizer

WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland

ARTICLE INFO

Article history:
Received 5 November 2008
Accepted 25 August 2010

Keywords:
Snow mechanical properties
Snow stability
Avalanche formation
Spatial variability
Geostatistics

ABSTRACT

Spatial variations of weak layer and slab properties are believed to affect snow slope stability. To quantify spatial variability at the slope scale, penetration resistance was measured with a high-resolution snow micro-penetrometer (SMP) in a partly randomized grid pattern. The grid design consisted of 46 SMP measurement locations. In addition, a full snow profile and 20 compression tests as well as a Rutschblock test at the snow profile location were performed within the grid. Fifteen slopes of different aspects were sampled of which 11 could be analysed. Weak layer and slab properties were characterised using non-spatial as well as spatial statistics and results were related to slope stability. The geostatistical analysis revealed that in more than half of the cases a range could be determined. Slab layers tended to have more spatial structure than the weak layer. Though some trends are apparent, firm conclusions on the dependence of slope stability on spatial variations were not possible due to the limited range of snow conditions in the dataset, and the fact that the definition of slope stability remains elusive. Based on our limited data set, we can therefore not specify the conditions when spatial variations of weak layer and slab properties are most relevant for snow slab release.

1. Introduction

Spatial variations of weak layer and slab properties are considered a key factor for snow slope stability. Spatial variability is presumably caused by meteorological conditions such as wind during snow deposition, as well as metamorphic processes after deposition (Schweizer et al., 2008a). Kronholm and Schweizer (2003) suggested slope stability to be related to the degree of stability variation, the length scale of variation and the mean point stability.

Previous studies analysed weak layer properties at the slope scale (Kronholm and Schweizer, 2003; Birkeland et al., 2004). Kronholm et al. (2004) determined the spatial structure of seven weak layers and the corresponding slab layers. They found the spatial structure of slab layers to be more variable than the spatial structure of weak layers. Schweizer et al. (2008a) concluded that weak layers were often spatially continuous and showed less variability than the stability of small column tests.

The snow micro-penetrometer (SMP) developed by Schneebeli and Johnson (1998) measures penetration resistance at high-resolution and allows one to derive micro-structural snow parameters (Johnson and Schneebeli, 1999). Pielmeier and Schweizer (2007) made a first attempt to classify a priori known weak layers with regard to stability based on these micro-structural parameters. Recently, Pielmeier and Marshall (2009) have improved this procedure based on a series of rutschblock tests in conjunction with SMP measurements. Kronholm and Schweizer (2003) were unable to relate spatial SMP measurements to the manual stability observations partly due to (a) the sampling design that did not allow a spatial analysis for the manual stability observations and (b) the lack of an SMP derived stability parameter. Bellaire et al. (2009) developed an algorithm for analysing the SMP signal with the aim to detect potential weak layers and estimate the degree of instability with regard to skier-triggering.

For this study, we analysed a series of spatial measurements at the slope scale of penetration resistance (SMP) and point stability (compression test, Rutschblock test) performed on slopes near Davos, Switzerland. The aim was to detect a relation between weak layer and slab properties (and their spatial variations) and slope stability.

2. Methods

2.1. Snow micro-penetrometer (SMP)

Schneebeli and Johnson (1998) developed a snow micro-penetrometer for rapid snow cover investigations at high vertical resolution. The SMP consists of a cone-shaped probe which is driven into the snow cover by a motor at a constant speed of 20 mm s^{-1}. Penetration resistance is measured every 4 μm by a piezo-electric force sensor.

* Corresponding author. Present address: ASARC — Applied Snow and Avalanche Research, Dept. of Civil Engineering, University of Calgary, Calgary AB, Canada. Tel.: +1 403 220 5830.
E-mail address: sascha.bellaire@ucalgary.ca (S. Bellaire).

0165-232X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.coldregions.2010.08.013

To derive micro-mechanical properties of snow from the SMP signal, Johnson and Schneebeli (1999) developed a micro-structural model. A micro-structural element will rupture within a typical length of dimension L_n, and will induce a force peak in the SMP signal F_{max}. The rupture force f_r is the force needed to rupture the crystal structure. The rupture is assumed to occur between two crystals, i.e. at a bond. The number of ruptures per unit length n_{peaks} is defined as the number of peaks (local maxima). Fig. 1 illustrates these three parameters for two different snow types.

The peak force and the number of ruptures are typically large for well-bonded snow and small for poorly bonded snow structures. Johnson and Schneebeli (1999) (Fig. 4) showed that, for example, a wind slab has approximately a ten times larger rupture force than snow consisting of depth hoar.

2.2. Sampling design

To quantify the effect of spatial variations on slope stability, spatially distributed measurements of snow cover properties need to be performed on potential avalanche slopes. For spatial analysis geostatistics have been used to identify the spatial structure, in particular the correlation length of snow properties, which is believed to be related to avalanche formation (e.g. Kronholm, 2004). An optimized sampling design enables one to identify the spatial structure with a minimum of measurements. If the correlation length is unknown, then the sampling design should contain some randomness (Kronholm and Birkeland, 2007). However, in practice, randomly distributed measurement points are difficult to locate on grids where the extent (maximum distance between sampling locations) is small (≤20 m), i.e. locating points, for example, by GPS requires differential GPS.

Skøin and Blöschl (2006) demonstrated that the estimated correlation length can be biased depending on the choice of extent, spacing (distance between sampling points) and support (area of sampling). This means that the sampling design is of particular importance for estimating the correlation length. Furthermore, they pointed out that the sampling design should be adapted to the expected correlation length. As a rule of thumb, Skøin and Blöschl (2006) suggested an optimized sampling design with an extent larger and a spacing smaller than the expected correlation length.

Correlation length of snow cover properties can range from a few decimeters to several meters, or more. Therefore, the correlation length is a priori unknown. Also, the number of processes acting on the snow cover, and the typical length scale of these processes are not sufficiently known. Many processes, such as wind, cause variations at different scales (e.g. Schweizer et al., 2008b). Since we focus on avalanche formation, in particular on the effect of spatial variations on avalanche formation, we can make an assumption on the length scale that is most relevant. This scale is related to the avalanche release process which can be described in terms of fracture mechanics. In order for an initial failure to propagate so that eventually the slab becomes detached, the failure has to reach a critical size. Independent estimates suggest that the critical size is <10 m but larger than the slab thickness (e.g. Schweizer et al., 2003). Therefore the sampling design should be such that a correlation length of a few meters can be determined — at least approximately. In other words, we are not primarily seeking a robust estimate, but are interested whether a correlation length of a few meters exists or not.

The sampling design used in this study was developed for a grid of 18 m×18 m (Fig. 2). This area was divided into nine sub-grids of 6 m×6 m. Each sub-grid contained five SMP measurements in an L-shaped design as suggested by Cline et al. (2001). The distance between each SMP measurement within a sub-grid differed (0.25 m, 0.5 m, and 1 m). This sampling design has a well distributed spacing h, with a mean spacing of 9 m and an extent of 19 m; for lag distances h = 4–17 m the number of point pairs is >30 (compare Fig. 3). Our sampling design is partly randomized and has therefore no fixed spacing. The mean spacing is on the order of the maximum expected correlation length and might therefore be somewhat too large. However, our random sampling design also includes distances smaller than the expected correlation length, which allows one to cover the small-scale variability.

2.3. SMP signal analysis

For the present study we only focused on the micro-structural parameter Ψ introduced by Bellaire et al. (2009), which we used to describe the structure of a layer. The parameter Ψ (Pa), which is

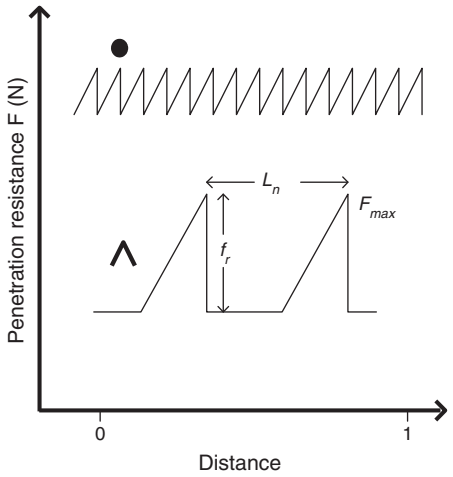

Fig. 1. Schematic SMP signals for small rounded grains (top) and depth hoar (bottom), and the definition of the micro-structural parameters, rupture force f_r, element length L_n and peak force F_{max}. The number of peaks n_{peaks} corresponds to the number of ruptures per unit length.

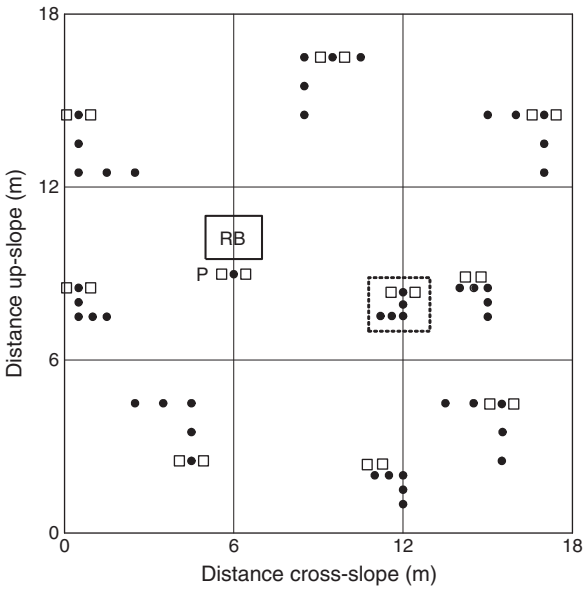

Fig. 2. Sampling design of the present study. Dots indicate locations of SMP measurements, squares location of compression tests adjacent to the SMP measurement. The position of the manual profile is indicated by "P", where two additional compression tests and a SMP measurement were performed. "RB" locates the position of the Rutschblock test. The SMP measurements within the dashed square are 0.25 m apart; the other ones are up to scale (0.5 and 1 m apart).

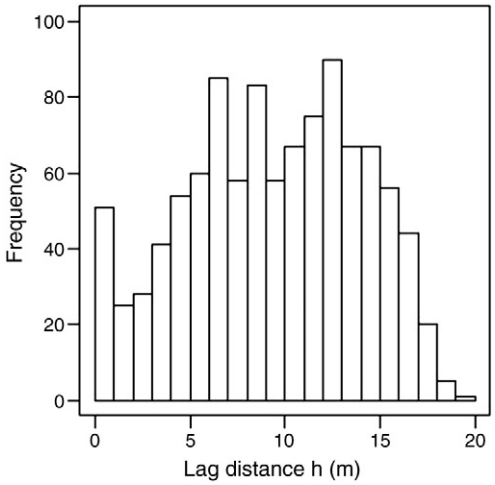

Fig. 3. Frequency distribution of available point pairs per lag distance for the sampling design shown in Fig. 2.

similar to the micro-scale strength used by Pielmeier and Marshall (2009), is defined as:

$$\Psi = \frac{\bar{f}_r \; n_{\text{peaks}}}{A} \tag{1}$$

with $\bar{f}_r$ (N) the mean rupture force averaged over one millimetre, n_{peaks} the number of ruptures within 1 mm of the signal, and A the lateral surface area of the conical sensor tip ($A = 39\ \text{mm}^2$). The parameter Ψ is small for poorly bonded layers (depth hoar, surface hoar) and is rather large for well-bonded layers (small rounded crystals, crusts).

Bellaire et al. (2009) suggested various parameters and split values for the stability estimation. The parameter Ψ with a threshold value of about 5 kPa allowed classifying SMP profiles into rather "poor" (smaller than threshold) or rather "fair-to-good" snowpack stability (larger than threshold) with a an accuracy of 78% (not cross-validated).

2.4. Measurements and manual observations

Each grid consisted of 46 SMP measurements and 10 pairs of compression tests (9 in sub-grids and one pair near the manual profile site) as well as a manual profile and a rutschblock test (Föhn, 1987). All compression tests (Jamieson, 1999) within one grid were performed by either one of two experienced observers. Early season training sessions ensured comparable results. The failure layer identified by the Rutschblock test was defined as the predominant weak layer. For further analysis, we considered only compression test scores for this weak layer. We used the minimum score of two side by side compression tests and assigned this score to the stability of the sub-grid or the profile location. The median difference between the two compression tests was one score (3rd quartile: 2 scores). In most cases, the RB failure layer also failed in both of the two adjacent compression tests. In five out of 110 compression test pairs, only one compression test of the pair failed on the weak layer identified by the Rutschblock test.

The 46 SMP measurements in each grid were visually inspected and signals of questionable quality were excluded. Therefore the number of SMP measurements available for analysis differed per grid (from 41 to 46). For all SMP signals we manually located the weak layer using the manual profile and the compression tests as a reference. The weak layer was present in all SMP signals and was analysed as described above using the parameter Ψ. All layers above the weak layer we considered as part of the slab layer.

The manual profiles were classified into five stability classes (1: "very poor" to 5: "very good") based on the stability classification introduced by Schweizer and Wiesinger (2001).

2.5. Data analysis

To assess whether the study slopes showed spatial structure we used spatial statistics. First, we estimated the spatial structure with the Moran's I coefficient, a measure of spatial autocorrelation (Moran, 1948) that has been recently applied in a snow study by Hendrikx et al. (2009). The coefficient ranges from -1 (dispersion) to $+1$ (clustered). A value of zero indicates a random pattern. The null hypothesis chosen was "no spatial autocorrelation exists". For this study, we choose a level of significance $p<0.05$ to reject the null hypothesis. The Moran's I statistics was calculated with the Moran I-function implemented in the ape-package (Paradis et al., 2004) for R (R Development Core Team, 2009).

Second, geostatistics were used to determine the correlation length. Therefore an experimental semivariogram was calculated using a robust method to remove outliers (Cressie and Hawkins, 1980):

$$\gamma(h) = \frac{\left[\frac{1}{N(h)} \sum_{i=1}^{N(h)} |Z(x_i + h) - Z(x_i)|^{\frac{1}{2}}\right]^4}{\left(0.914 + \frac{0.988}{N(h)}\right)} \tag{2}$$

with Z the variable of interest at sampling location $x_i = (X_i, Y_i)$ and $N(h)$ the number of point pairs separated by the lag distance h.

Slope scale trends have been quantified and removed prior to the geostatistical analysis. Trends have been quantified by applying a first order polynomial to the data, with coordinates X and Y as independent variables. The correlation coefficient r^2 of this regression gives the accuracy of the fit. In addition, it describes how much of the spatial variation can be explained by slope scale trends.

By fitting a spherical model to the experimental semivariogram we determined the sill σ^2, range R, and the nugget τ (Webster and Oliver, 2007). Only lag distances to half the extent were considered. The weighted sum of squares (WSS) determined the accuracy of the fit.

2.6. Testing the sampling design

Geostatistical analysis is influenced by the sampling design (Skøin and Blöschl, 2006). To assess the applicability of the sampling design introduced above for our specific purpose (presence or absence of a range of a few meters), Gaussian random fields with defined initial covariance parameters ($\sigma^2 = 1 \pm 0.2$; $R = 2$ m, 5 m and 8 m ± 0.2 m, $\tau = 0$; 100 simulations per range) were generated using the grf-function implemented in the RandomFields-package (Schlather, 2001) for R (R Development Core Team, 2009). The fields were generated on a rectangular grid with a regular spacing of 0.25 m (5328 points). Each point of the sampling location was assigned to the corresponding value of the generated field and a sample variogram was calculated.

A spherical model λ was fitted to each sample variogram and the parameter range R, sill σ^2 and the nugget τ were determined ($\lambda = R, \sigma^2, \tau$). A method suggested by Cressie (1993) was used to estimate the best fitting variogram. This method used the weighted least-squares approach and is given by the function J:

$$J(\lambda) = \sum_{i=1}^{K} N(h(i)) \left\{ \frac{\hat{\gamma}(h(i) - \gamma(h(i); \lambda)}{\gamma(h(i); \lambda)} \right\}^2 \tag{3}$$

where $N(h(i))$ is the number of point pairs separated by the lag distance $h(i)$ ($i = 1,\ldots,K$), and $\hat{\gamma}$ is the sample variogram and γ the theoretical variogram. The function J gives more weight to smaller

lags and to lags with more point pairs $N(h(i))$. The initial covariance parameters of the theoretical variogram were the same as used to generate the random fields.

3. Data

During the winters 2006–2007 and 2007–2008 fifteen grids were observed on different slopes above timberline (at about 2400 m a.s.l.) in the region of Davos, Switzerland. The slopes had four different aspects (N, NE, S, and SW) and the slope angle varied between 18° to 34°. Four grids had to be discarded since the majority of the SMP measurements were erroneous. In most of the remaining eleven grids (eight out of eleven) the predominant weak layer was a layer of persistent grains consisting of depth hoar or large faceted crystals. The other three weak layers or interfaces were (1) a layer of small graupel, (2) a mixed layer of decomposed fragmented particles and small rounded grains, and (3) a hardness change (fist to four fingers, grid 11) within a layer of small rounded grains, i.e. not a distinct weak layer.

The study slope was classified with regard to stability into the stability classes of "poor", "fair" and "good". The slope stability estimate was based on the profile stability class (Schweizer and Wiesinger, 2001), the median compression test score of the grid as well as on any signs of instability such as "whumpfs", cracking or recent avalanche activity on nearby slopes (Table 1). Following Schweizer and Jamieson (2003), we assigned a median compression test score per grid of ≤13 to the stability class of "poor", and scores >13 to "good" slope stability. Grids where the manual profile was classified as "fair" (profile stability class 3), were rated as "fair" (slope stability), if either the mean compression test score was ≤13 or signs of instability had been observed.

4. Results

We present (1) the non-spatial statistics and relate them to slope stability, and (2) the spatial statistics for the weak layer and slab layer properties.

4.1. Non-spatial analysis of slab and weak layer properties

Descriptive statistics for all eleven grids are compiled in Table 2. In general, the median and the semi-interquartile range (SIQR; Spiegel and Stephens, 1999) of the parameter Ψ were lower for the weak layer than for the slab layers. One exception was grid 11. In this grid, Ψ had low median values and little spread (i.e. small SIQR) for both the weak layer and the slab, and Ψ was larger for the weak layer than for the slab. The failure layer was not a distinct weak layer, but rather a distinct step in the hardness within the snowpack. Whereas the dispersion was almost twice as large for the slab layers than for the weak layers, the slab layers had on average a lower quartile coefficient of variation, i.e. a lower relative dispersion than the weak layers.

Bellaire et al. (2009) proposed a threshold value for Ψ of 5 kPa to discriminate between rather unstable and rather stable SMP profiles. According to this, five grids (1, 2, 5, 6 and 11) should have rather poor stability (Table 2). The first two of these grids were indeed rated as "poor" and were associated with soft slabs (i.e. low Ψ for the slab). Grids 5 and 6 had well consolidated slabs (i.e. high Ψ) and were rated as "fair". Finally, grid 11, which was classified as "good", was not associated with a distinct weak layer.

All grids with a median value of Ψ for the weak layer larger than 5 kPa were rated as "fair" or "good", with the exception of grid 3. For this grid, both the weak layer and the slab had large median Ψ values and large spread. This grid was performed on a wind affected slope where large variations in slab and weak layer thickness existed (snow height varied from 30 cm to 200 cm). Prior to sampling, the slope was triggered (whumpfed) from a thin spot.

In Fig. 4 the variation of Ψ for the weak layer and the slab layers is shown for the stability classes "poor", "fair" and "good" ($N_{poor}=136$, $N_{fair}=134$, $N_{good}=219$). Despite the limited number of observations some trends are apparent.

The variation of Ψ for the weak layer was smaller for slopes classified as "poor" than for slopes classified as "fair" or "good". The largest spread was observed for the group of "fair" slope stability. Furthermore, the median values of Ψ for the weak layer were smaller for the slopes classified as "poor" and "fair", than for the slopes classified as "good". For the slab, on the other hand, the spread in Ψ was similar for all stability classes. However, the median Ψ values for the slab were lower for the class "poor" than for the classes "fair" and "good".

4.2. Spatial analysis of weak layer and slab properties

To quantify the spatial structure of each grid we used the Moran's *I* coefficient as well as a geostatistical analysis. In order to estimate the reliability of the geostatistical analysis we also analysed the above described sampling design.

4.2.1. Spatial analysis using Moran's I

The Moran's *I* coefficients, a measure of spatial autocorrelation, for Ψ for the weak layer and the slab layer are shown for all eleven grids in Table 3. Spatial structure was identified in 8 out of 11 cases for the weak layer and in 10 out of 11 cases for the slab layers. The median Moran's *I* coefficient was slightly larger for the slab ($I=0.24$) than for the weak layer ($I=0.18$), suggesting that slab layers were slightly more clustered than weak layers.

4.2.2. Geostatistical analysis

Fig. 5 shows the variation of simulated ranges for the three generated ranges of 2 m, 5 m and 8 m sampled with the sampling

Table 1
Classification of study slopes into either "poor", "fair" or "good" slope stability. Slope stability is based on profile classification according to Schweizer and Wiesinger (2001) (1: very poor, 2: poor, 3: fair, 4: good, 5: very good) and the median point stability (MPS) (median CT score ≤13) or on the presence of any signs of instability.

Profile classification	MPS		Signs	Estimated slope stability
1, 2	≤13	and	Present	Poor
3	≤13	or	Present	Fair
3, 4, 5	>13	and	Absent	Good

Table 2
Summary statistics for the eleven grids. Rutschblock score (RB, scores given in bold indicate a whole block release), PC the profile classification into five stability classes (1: "very poor" to 5: "very good") according to Schweizer and Wiesinger (2001), the median point stability (MPS) based on CT score, presence (1) or absence (0) of signs of instability, and estimated slope stability are given. Furthermore, the median and the semi-interquartile range (SIQR) of Ψ for the weak layer and the slab layers are shown.

Grid						Weak layer Ψ		Slab layers Ψ	
No	RB	PC	MPS	Instability	Slope stability estimate	Median kPa	SIQR kPa	Median kPa	SIQR kPa
1	**2**	2	10.5	1	Poor	1.6	0.9	3.5	1.4
2	**2**	1	11	1	Poor	5.1	2.5	7.7	3.7
3	**3**	2	12	1	Poor	17.2	8.7	59.2	22.9
4	**5**	3	11.5	1	Fair	53.0	11.0	67.6	14.7
5	5	3	12	0	Fair	1.5	1.6	53.3	7.6
6	**4**	3	12.5	1	Fair	3.1	2.0	28.0	7.7
7	**5**	3	13.5	0	Good	20.7	5.1	24.1	4.2
8	**4**	3	14.5	0	Good	12.1	3.5	30.3	6.9
9	6	4	22.5	0	Good	48.9	12.8	92.4	20.1
10	3	3	14	0	Good	7.9	2.7	42.1	6.3
11	4	4	19	0	Good	2.5	0.4	2.1	0.3

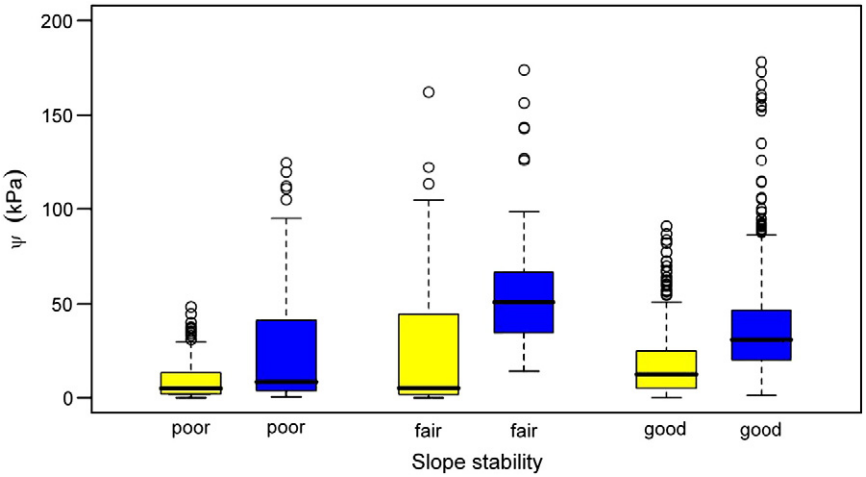

Fig. 4. Distributions of parameter Ψ for the weak layer (yellow) and the slab layers (blue) for the grids rated as either "poor", "fair" or "good". ($N_{poor} = 136$, $N_{fair} = 134$, $N_{good} = 219$). Boxes span the interquartile range. Open circles indicate outliers.

design used in this study. The median, mean and the standard errors of the simulated ranges are given in Table 4 as well as the 95% confidence interval. The differences between the mean simulated range and the generated range varied from 1.4 m to 2.4 m, while the deviation from the median of the simulated ranges was always smaller than 1 m. The largest standard error (±0.7 m) was observed for the generated 2 m range, the smallest standard error (±0.3 m) for the 5 m range. Values for the estimated range that were far beyond the extent (5 out of 300) were discarded for the calculation of the mean and standard error.

The results of the geostatistical analysis for the parameter Ψ for the weak layer and the slab layers are shown in Table 5. In 3 out of 11 grids the range for the weak layer Ψ after trend removal was larger than 10 m (unbounded variogram) (Fig. 6). For one grid a pure nugget variogram was found. The remaining seven grids showed ranges between 2 m and 8 m, with a median range of 4 m. For the slab layer Ψ four grids showed unbounded and one a pure nugget variogram. The six remaining grids showed ranges between 2 m and 6 m (median: 3 m). The median total sill was smaller for the weak layers than for the slab layers. The median correlation coefficient r^2 was larger for the slab layers than for the weak layers.

5. Discussion

Assessing snow slope stability (or avalanche release probability) – by either extrapolating from a single point measurement or from many spatially distributed measurements – is of primary interest for avalanche forecasting. Our dataset contains only eleven grids, which proved to be insufficient to draw any firm conclusions based on a statistical analysis. Furthermore, our study is limited by (a) the fact that stability cannot readily be measured (but at best extrapolated) and (b) by safety concerns for the field crew which limits sampling to days when the avalanche conditions are not very unstable. Really unstable conditions are only found when an avalanche is triggered on the study slope. Otherwise, various degrees of stability can be found, but the differences between those might be too small for reliable discrimination. Therefore, it is not surprising that in our dataset no clear patterns of spatial variations with regard to slope stability can be seen since there are only very few measurements in the unstable range.

Apart from spatial variability, measurement uncertainties also introduced variations in the measurements. To maximize the number of measurements available for analysis, we manually identified the weak layer. The manual detection of weak layers is subjective and if the weak layers or interfaces are not very distinct this introduces some uncertainty, which cannot be quantified. Other sources of uncertainty stem from the stability tests and from the stability classification. For instance, a threshold value CT ≤ 13 was used as a criterion for slope stability classification, i.e. the uncertainty of the

Table 3
Summary of the Moran's *I* statistics. Given are the estimated slope stability, the Moran's *I* coefficient with its corresponding *p*-value, as well as the presence (1) or absence (0) of clustering (C) of Ψ.

Grid		Weak layer			Slab layers		
No.	Slope stability	Moran's *I*	*p*	C	Moran's *I*	*p*	C
1	Poor	0.04	0.119	0	0.31	<0.001	1
2	Poor	0.07	0.031	1	0.24	<0.001	1
3	Poor	0.21	<0.001	1	0.3	<0.001	1
4	Fair	0.13	<0.001	1	0.35	<0.001	1
5	Fair	0.03	0.161	0	0.11	0.002	1
6	Fair	0.27	<0.001	1	0.09	0.019	1
7	Good	0.18	<0.001	1	0.13	0.001	1
8	Good	0.23	<0.001	1	0.3	<0.001	1
9	Good	0.37	<0.001	1	0.56	<0.001	1
10	Good	0.04	0.136	0	0.08	0.026	1
11	Good	0.26	<0.001	1	−0.02	0.643	0

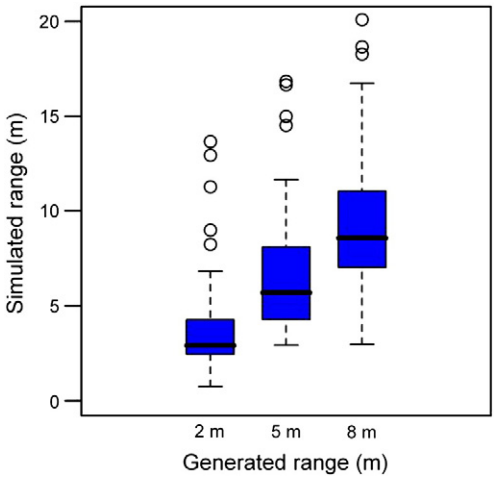

Fig. 5. Distribution of simulated ranges for the three generated ranges of 2 m, 5 m and 8 m. Boxes span the interquartile range. Open circles indicate outliers.

Table 4
Summary statistics for the sampling design test. Given are the generated range, the median of the simulated range, the mean of the simulated range as well as the standard error of the mean (SE) and the 2.5% and 97.5% quantiles.

Generated range	Median	Mean	SE	$Q_{2.5}$	$Q_{97.5}$
m	m	m	m	m	m
2	2.9	4.4	0.7	1.4	13.3
5	5.7	6.4	0.3	3.3	14.8
8	8.5	9.8	0.5	4.8	16.5

compression test results greatly affects the stability classification, in particular if the median score is close to the threshold value.

The non-spatial analysis suggests that the median values of Ψ for the weak layer and the slab layer increased with increasing slope stability (Fig. 4). The spread in Ψ for the slab layers showed no trend. A small spread would be related to rather uniform slab layers whereas a large spread would indicate variable slab layers. The median values of Ψ for the weak layer were rather larger for the stability class of "good" than for the classes of "poor" and "fair". Small values of Ψ indicate poorly bonded grains, i.e. fracture initiation should be possible. The largest spread was observed for the class of "fair" stability. However, a small median Ψ was also related to this class indicating areas of poorly bonded grains. These areas might be larger than the critical length required for fracture propagation.

Fig. 4 suggests that the stability classification with a threshold value of $\Psi = 5$ kPa reported by Bellaire et al. (2009) is not suited for classifying slopes in regard to stability. In fact, applying this threshold and grouping the slope stability classes "fair" and "good", an unweighted average accuracy of 58% results. As the datasets (spatial vs. point measurements) and stability classification (slope estimate vs. Rutschblock) are different, a direct comparison with the results of Bellaire et al. (2009) is not meaningful. In the present study we attempted to discriminate between three stability classes ("poor", "fair" and "good") instead of two. This is obviously a more difficult classification problem; in particular the middle class ("fair") is not distinct enough. In addition, when classifying slopes the spatial stability measurements were considered in the present study. For the future, we intend to develop an improved classification scheme which should include a combination of weak layer and slab layer properties and possibly some information on their spatial structure.

The Moran's I statistics suggest that the slab layer Ψ was more clustered than the weak layer Ψ. For most of the grids, correlation length between 2 m and 8 m were identified for the weak layer and the slab layer Ψ. With our sampling design ranges below 10 m can be identified with an accuracy of about 2 m (Fig. 5), but the exact correlation length cannot be determined. Slope scale trends were observed for the weak and slab layer Ψ; trends were larger for the slab layer Ψ (see r^2 in Table 4). This supports the results of the Moran's I statistics and might be related to the fact that different processes at different scales can cause slope scale variations. For example, wind might affect the slab and cause slope scale variations, whereas the weak layer will be more influenced by metamorphic processes. Further slopes need to be investigated to relate possible causes of variability to the type and amount of variability.

The non-spatial statistics showed that slab layers had lower relative spread, though the absolute spread was higher. The mean Ψ tended to be larger for the slab layers and the slab layers were more clustered than the weak layer. These results contradict the findings by Kronholm et al. (2004) who concluded that slabs may be more variable than weak layers. The reasons for this discrepancy are presently unclear.

The observed lack of correlation between spatial variations and slope stability might be explained by the hypothesis that a spatial structure is only relevant when there are variations between unstable and stable conditions, i.e. stable and unstable areas exist on the same slope. Otherwise, the presence of a spatial structure with only variations within the unstable or stable range still results in an overall instability or stability, respectively. Variations and their length scale seem irrelevant if they occur fully within the stable or unstable range. Spatial variability, and especially the correlation length, would only become crucial if variations occur across stability classes. Considering the threshold of 5 kPa for the weak layer introduced by Bellaire et al. (2009) the investigated weak layers showed primarily variations in either the stable or the unstable range. However, only one of the slopes (grid 3) with a potentially unstable weak layer was triggered during investigation. This suggests that besides the weak layer characteristics, the slab layer properties need to be taken into account as well for slope stability classification.

In summary, combining the results of the spatial and non-spatial analysis suggests that some amount of spatial variability always exists. We hypothesize that variations and their length scales are only relevant when the variations are around the threshold between rather unstable and rather stable conditions. Otherwise, it seems that the slope is either stable or unstable, independent of spatial variations, and point stability measurements might well be indicative — as exemplified by the average classification accuracy of stability tests (70–90%) (e.g. Schweizer and Jamieson, 2010; Schweizer and Bellaire, 2010). In other words, spatial variations might be most critical for slope stability evaluation when stability is "fair". This conjecture is in agreement with the view that incidental skier-triggering is most critical in situations of intermediate or fair stability with substantial variability in snowpack properties (Schweizer, 2004).

We also analysed (not shown) the spatial structure of the compression tests (initialization) and point stability derived from the algorithm introduced by Bellaire et al. (2009). No structures were

Table 5
Results of the geostatistical analysis. Given are the range R, the sill σ^2, the nugget τ as well as the weighted sum of squares (WSS) as a measure of fit quality for all 11 grids. Also shown is the correlation coefficient r^2 of the trend analysis. The second column contains the estimated slope stability. No range is given (–) if a pure nugget variogram was found.

Grid		Weak layer					Slab layers				
No.	Slope stability	R	σ^2	τ	WSS	r^2	R	σ^2	τ	WSS	r^2
		m	kPa	kPa	kPa		m	kPa	kPa	kPa	
1	Poor	>10	2.8×10^3	1.8	6.2×10^2	0.19	2	1.4	0	6.7×10^1	0.67
2	Poor	–	0	17	9.5×10^5	0.04	3	30	0.9	4.5×10^4	0.04
3	Poor	2	63	29	3.8×10^5	0.33	6	3.2×10^2	56	5.3×10^6	0.44
4	Fair	>10	3.0×10^5	1.8×10^2	1.3×10^7	0.18	>10	2.3×10^5	0	1.6×10^7	0.63
5	Fair	4	5.5	4.9	2.3×10^3	0.09	6	1.7×10^2	0	5.2×10^5	0.29
6	Fair	5	7.6	2.2	1.4×10^4	0.21	3	92	14	5.3×10^5	0.20
7	Good	6	85	0	3.9×10^5	0.14	–	0	34	4.8×10^3	0.25
8	Good	>10	4.8×10^4	0	6.3×10^4	0.02	>10	6.7×10^4	4.6	1.7×10^5	0.03
9	Good	8	2.7×10^2	50	2.9×10^6	0.35	>10	3.0×10^3	0.0	4.4×10^7	0.49
10	Good	8	28	8.1	2.8×10^3	0.16	>10	7.2×10^4	35	1.3×10^6	0.17
11	Good	3	0.8	0	1.4×10^2	0.35	3	1.4	0	1.3×10^3	0.04

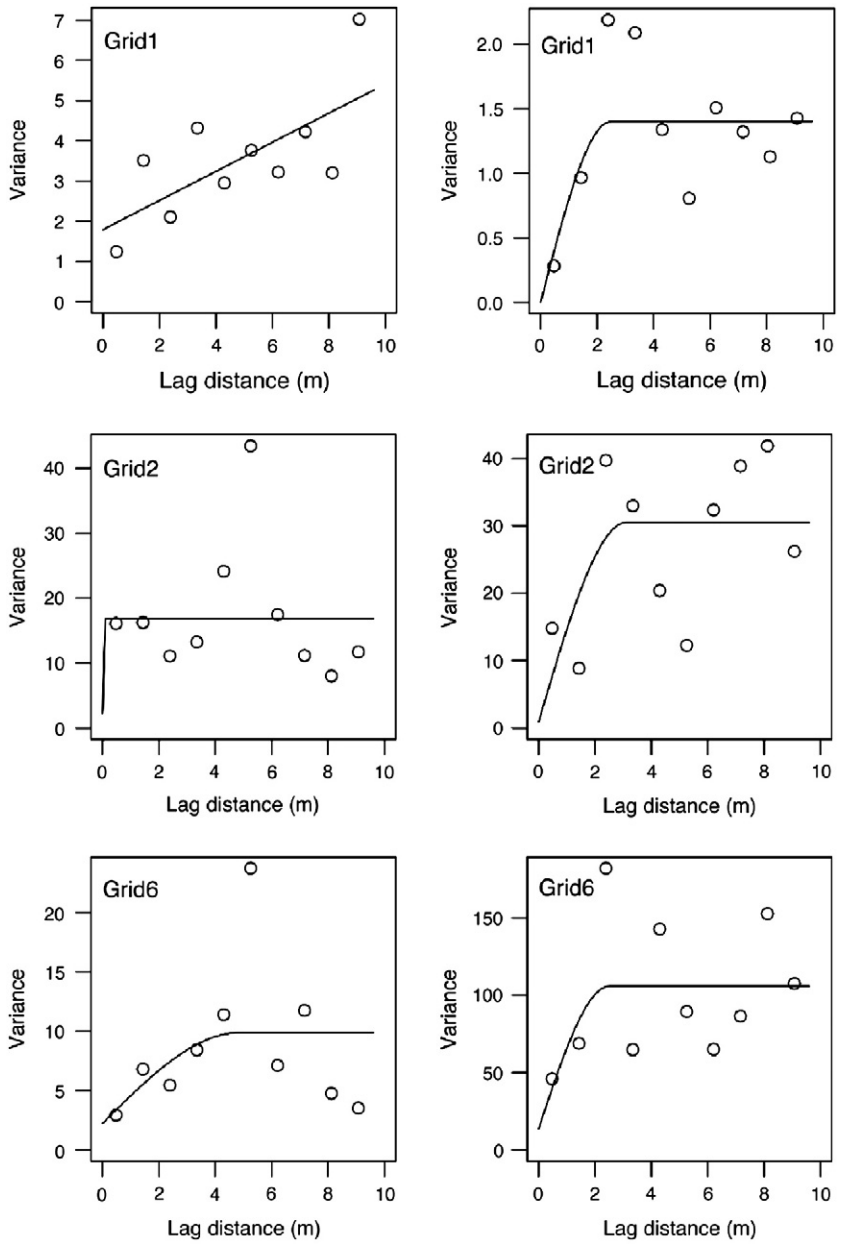

Fig. 6. Exemplary variograms for the parameter Ψ for grids 1, 2 and 6, for the weak layer (left) and the slab layers (right).

identified, which allowed one to distinguish between slopes of "poor", "fair" or "good" stability. In addition, as described above the point stability values (compression tests) are prone to errors. Point stability values should at best be considered as an index of failure initiation. In other words, it seems questionable whether spatially distributed point stability observations can be used to estimate slope stability. At least, low variability of the compression test scores within a grid was related to rather poor slope stability, which might favour fracture initiation as well as propagation. Whereas initiation is limited to a small area, the scale of propagation is much larger.

There is no doubt that the properties of the mountain snowpack are spatially variable. However, measuring, analysing and interpreting spatial distributed measurements remains challenging which is why the present study, as well as many previous studies, is partly inconclusive.

6. Conclusions

We analysed spatial measurements of penetration resistance performed using a partly randomized novel grid design on eleven slopes near Davos, Switzerland. A micro-structural parameter Ψ, which describes the structure of snow layers (related to micro-structural strength), was analysed non-spatially as well as spatially. The spatial measurements of Ψ were related to an estimate of slope stability.

The non-spatial analysis showed that the median and spread of the parameter Ψ for the weak layer and the slab layers can be very different. These differences could partly be related to the observed slope stability. For example, on rather stable slopes higher values of Ψ for the weak layer and the slab layers were observed compared to values of Ψ on slopes rated as "poor". Whereas the spread for slab layers did not show a trend, the spread for the weak layer tended to increase with increasing slope stability, but the apparent increase was statistically not significant.

The geostatistical analysis showed that in more than half of the cases a range of a few meters was identified for the weak layer and the slab layers. The Moran's *I* index suggests that slab layers were more clustered than weak layers; they also showed more often a slope scale trend. The weak layers showed generally lower values of Ψ and lower

absolute (but higher relative) dispersion than the slab layers. We were unable to relate a certain correlation length to slope stability and therefore the effect of length scale on slope stability remains unknown for the time being.

Our results should be interpreted as a preliminary attempt towards a better understanding of the complex interaction of snow cover properties – including their spatial variations – with slope stability. Whereas, some of the apparent trends seem to be plausible in comparison to previous hypotheses on how spatial variations of weak layer and slab properties might affect avalanche release probability, our study has to be considered as inconclusive with regard to establishing a relation between spatial variations and avalanche release probability. However, our measurements provide new insight into the nature of spatial variations at the slope scale and point out the difficulties that need to be tackled in order to clarify the effect of spatial variations on avalanche release probability. In particular, the definition of slope stability in the course of field studies seems problematic.

Acknowledgements

For help with the field work we would like to thank Alec van Herwijnen and Christoph Mitterer, as well as Sina Schneider, Michael Schirmer, and Charles Fierz. We are grateful for the comments by Alec van Herwijnen, Karl Birkeland and two anonymous reviewers. We gratefully acknowledge financial support by the European Commission (FP6-STReP-NEST, Project no. 043386: TRIGS).

References

Bellaire, S., Pielmeier, C., Schneebeli, M., Schweizer, J., 2009. Stability algorithm for snow micro-penetrometer measurements. J. Glaciol. 55 (193), 805–813.

Birkeland, K., Kronholm, K., Schneebeli, M., Pielmeier, C., 2004. Changes in shear strength and micropenetration hardness of a buried surface-hoar layer. Ann. Glaciol. 38, 223–228.

Cline, D., Armstrong, R., Davis R., Elder K., Liston G., 2001. NASA Cold Land Processes Field Experiment Plan. [Available online at http://www.nohrsc.nws.gov/~cline/clpx.html].

Cressie, N.A.C., 1993. Statistics for Spatial Data. Wiley Series in Probability and Mathematical Statistics. Wiley, New York. 900 pp.

Cressie, N.A.C., Hawkins, D.M., 1980. Robust estimation of the variogram: I. Math. Geol. 12 (2), 115–125.

Development Core Team, R., 2009. R: A Language and Environment for Statistical Computing. R Foundation Statistical Computing, Vienna, Austria. Available online at http://www.R-project.org.

Föhn, P.M.B., 1987. The Rutschblock as a practical tool for slope stability evaluation. In: Salm, B., Gubler, H. (Eds.), Symposium at Davos 1986 — Avalanche Formation, Movement and Effects, IAHS Publ., 162. International Association of Hydrological Sciences, Wallingford, Oxfordshire, U.K, pp. 195–214.

Hendrikx, J., Birkeland, K., Clark, M., 2009. Assessing changes in the spatial variability of the snowpack fracture propagation propensity over time. Cold Reg. Sci. Technol. 56 (2–3), 152–160.

Jamieson, B., 1999. The compression test — after 25 years. Avalanche Rev. 18 (1), 10–12.

Johnson, J.B., Schneebeli, M., 1999. Characterizing the microstructural and micro-mechanical properties of snow. Cold Reg. Sci. Technol. 30 (1–3), 91–100.

Kronholm, K., 2004. Spatial variability of snow mechanical properties with regard to avalanche formation, Ph.D. Thesis, University of Zurich, Zurich Switzerland, 192 pp.

Kronholm, K., Birkeland, K.W., 2007. Reliability of sampling designs for spatial snow surveys. Comput. Geosci. 33 (9), 1097–1100.

Kronholm, K., Schweizer, J., 2003. Snow stability variation on small slopes. Cold Reg. Sci. Technol. 37 (3), 453–465.

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

Moran, P.A.P., 1948. The interpretation of statistical maps. J. R. Stat. Soc. B 10, 243–251.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.

Pielmeier, C., Marshall, H.-P., 2009. Rutschblock-scale snowpack stability derived from multiple quality-controlled SnowMicroPen measurements. Cold Reg. Sci. Technol. 59 (2–3), 178–184.

Pielmeier, C., Schweizer, J., 2007. Snowpack stability information derived from the SnowMicroPen. Cold Reg. Sci. Technol. 47 (1–2), 102–107.

Schlather, M., 2001. RandomFields: Simulation and Analysis of Random Fields. R package version 1.3.37. [Available online at http://www.stochastik.math.uni-goettingen.de/institute].

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

Schweizer, J., 2004. Skier-triggered avalanches: observations and concepts. In: Kristensen, K. (Ed.), Proceedings Snoskred og friluftsliv, Stryn, Norway, 24–27 October 2002, pp. 52–55.

Schweizer, J. and Bellaire, S., 2010. On stability sampling strategy at the slope scale. Cold Reg. Sci. Technol., doi:10.1016/j.coldregions.2010.02.013.

Schweizer, J., Jamieson, J.B., 2003. Snowpack properties for snow profile analysis. Cold Reg. Sci. Technol. 37 (3), 233–241.

Schweizer, J., Jamieson, J.B., 2010. Snowpack tests for assessing snow-slope instability. Ann. Glaciol. 51 (54), 187–194.

Schweizer, J., Wiesinger, T., 2001. Snow profile interpretation for stability evaluation. Cold Reg. Sci. Technol. 33 (2–3), 179–188.

Schweizer, J., Jamieson, J.B., Schneebeli, M., 2003. Snow avalanche formation. Rev. Geophys. 41 (4), 1016, doi:10.1029/2002RG000123.

Schweizer, J., Kronholm, K., Jamieson, J.B., Birkeland, K.W., 2008a. Review of spatial variability of snowpack properties and its importance for avalanche formation. Cold Reg. Sci. Technol. 51 (2–3), 253–272.

Schweizer, J., Heilig, A., Bellaire, S., Fierz, C., 2008b. Variations in snow surface properties at the snowpack depth, the slope and the basin scale. J. Glaciol. 54 (188), 846–856.

Skøin, J.O., Blöschl, G., 2006. Sampling scale effects in random fields and implications for environmental monitoring. Environ. Monit. Assess. 114, 521–552.

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.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists. Statistics in Practice Wiley, Chichester. 315 pp.