

Sampling and Statistical Analyses of BTS Measurements

Alexander Brenning,^{1*} Stephan Gruber² and Martin Hoelzle²

¹ Geographisches Institut, Humboldt–Universität zu Berlin, Berlin, Germany

² Department of Geography, University of Zürich, Zürich, Switzerland

ABSTRACT

Basal temperature of snow (BTS) data show characteristic spatial autocorrelations at distances typically less than 200 m, leading to non-independent regression residuals. Systematic temporal variations may also introduce model bias resulting in a shift in the predicted lower limit of permafrost. Both phenomena are analysed. The selection of an appropriate sampling design for a BTS measurement program appears critical in order to minimize problems typical of observational data in complex terrain. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: BTS; mountain permafrost; statistical analysis; spatiotemporal variability; sampling design

INTRODUCTION

The basal temperature of the snow (BTS) method predicts mountain permafrost distribution. It was introduced by Haeberli (1973) and consists of measuring the temperature at the base of the mid-winter snow cover. It is based on two assumptions: (1) BTS remains constant in mid-winter below a snow cover of at least 0.8 to 1 m thickness; and (2) the BTS value is determined by heat flux from the subsurface and serves as an indicator of subsurface thermal conditions.

Haeberli (1978) empirically established BTS thresholds of $\theta_1 = -3^\circ\text{C}$ and $\theta_2 = -2^\circ\text{C}$ for discriminating between 'likely permafrost', 'possible permafrost' and 'no permafrost'. This was done in the Swiss Alps in a statistically non-systematic way, but the mentioned thresholds have been applied in other high mountain areas without further discussion (compare e.g. Kneisel, 1999; Ødegård *et al.*, 1999; Ishikawa and Hirakawa, 2000; Fukui, 2003). Recently, the results presented by Lewkowicz and Ednie (2004)

showed that different BTS thresholds may need to be applied in Yukon Territory, Canada.

In recent years, the statistical and functional relationships between BTS, topographic parameters and remotely-sensed data have been used in order to statistically predict the regional BTS field and hence permafrost distribution (Hoelzle *et al.*, 1993; Hoelzle, 1994; Hoelzle and Haeberli, 1995; Etzelmüller *et al.*, 2001; Gruber and Hoelzle, 2001; Lewkowicz and Ednie, 2004).

Statistical methods—in contrast to deterministic approaches—not only provide a regionalized prediction of permafrost distribution, but also allow for estimating uncertainties. Also, statistical methods provide a means for minimizing the effort or cost that is necessary in order to obtain results of a desired quality.

This paper presents some methodological foundations of statistical BTS modelling for permafrost prediction. Cookbook-like recommendations to improve BTS measurement programs and data analysis, and to avoid typical problems and pitfalls, are suggested. Emphasis is put on meso-scale modelling (i.e. 25 to 200 m of spatial resolution; Etzelmüller *et al.*, 2001), but remarks on local-scale models are also included.

* Correspondence to: A. Brenning, Hartmannstr. 105, 91052 Erlangen, Germany.
E-mail: alexander.brenning@alumni.hu-berlin.de

TYPICAL PROBLEMS

Gruber and Hoelzle (2001) present a statistical analysis of BTS measurements in the surroundings of Zermatt, Switzerland. The sample consisted of 451 individual BTS measurements obtained in three separate study areas. Three hundred and seventy-one of these measurements were used for fitting linear regression models that explain BTS by means of elevation and other variables related to ground thermal regime. The remaining 80 measurements were used for model validation. The model finally selected by Gruber and Hoelzle (2001) was:

$$\text{BTS} = 7.809 - 0.0048 Z + 0.102 \text{PSWR} \\ + 4.933 \text{SAVI}$$

where the variables represent BTS temperature in °C (BTS), elevation in m ASL (Z), potential short-wave radiation (PSWR), and soil-adjusted vegetation index (SAVI). This model was applied to BTS regionalization and assessment of permafrost distribution using the mentioned thresholds of -2 and -3°C . Here, we disregard SAVI due to its limited importance as an explanatory variable, and we re-analyse the entire set of 451 measurements.

Statistical Problems

Probably the most striking characteristic of the Zermatt BTS data is related to the distribution of measurement locations with respect to terrain aspect. Forty-seven per cent of the measurement sites are exposed towards the north quadrangle, 29% towards the west quadrangle, and 10% look southward. This distribution is a consequence of the general orientation of the three field monitoring areas and has to be taken into account during model fitting in order to get an acceptable fit for all aspect classes.

As an example of the effect of the uneven distribution of orientations, we can try to fit the empirical radiation model $\text{PSWR} = \beta_0 + \beta_1 \text{NORTH} + \beta_2 \text{SLOPE}$, where NORTH denotes the 'north-exposedness' $\cos(\text{aspect})$ (north = 0° , clockwise angles), and SLOPE is the slope angle in degrees. In the European Alps, the more north-exposed a point is, the less solar radiation will it receive. Furthermore, the steeper the north-facing slope, the lower the incoming solar radiation. Theory tells us that this relation is wrong for south-facing slopes. Nevertheless, the model performs very well for the (mainly north-exposed) data; one gets $R^2 = 0.81$ with highly significant explanatory variables. We can even substitute PSWR in the model for BTS and get the fit:

$$\text{BTS} = 15.1 - 0.0061 Z - 0.65 \text{NORTH} \\ - 0.032 \text{SLOPE} \quad (R^2 = 0.454, \text{AIC} = 485.36) \quad (1)$$

with highly significant variables. The Akaike Information Criterion (AIC) is a relative measure of model fit, which penalizes models for the number of variables. Due to the strong linear relation of PSWR with NORTH and SLOPE, this model yields qualitatively the same results as

$$\text{BTS} = 11.4 - 0.0061 Z + 0.12 \text{PSWR} \\ (R^2 = 0.451, \text{AIC} = 485.99) \quad (2)$$

This is one of the models suggested by Gruber and Hoelzle (2001). But if both models are almost the same and equation (1) is not valid for southerly exposures, why should one believe that the almost numerically equivalent equation (2) is valid for south-facing slopes? Furthermore, almost two thirds of all measurements yield values below $\theta_1 = -3^\circ\text{C}$, i.e. they were made in 'likely permafrost' areas; the median is -3.7°C . As a consequence, the BTS model shows the best fit and gives best prediction results within an altitudinal belt where the likely presence of permafrost is known *a priori* e.g. due to the existence of active rock glaciers. However, modelling mountain permafrost distribution essentially means modelling its lower limit. As a consequence, model quality in this example might have benefited from measurement locations on average 200 m lower since a BTS median of $(\theta_2 + \theta_1)/2 = -2.5^\circ\text{C}$ would have been desirable.

Another issue in the Zermatt data set is that most BTS measurements either form clusters of some three to 10 individual measurements within a circle of radius 50 to 100 m, or follow the contour lines. Geostatistical analysis shows that BTS model residuals are correlated up to distances of 180 m. Similar results have been reported for other mountain areas (Figure 1, Table 1). Furthermore, when BTS measurements are made within a distance of few metres, these must be considered as different observations of the same random variable. If these observations are treated as independent, pseudoreplication occurs (Hurlbert, 1984; Crawley, 2002). This leads to incorrect distributional assumptions and invalid significance statements for estimates of model parameters.

On the other hand, measurements within small distances show considerable variability due to measurement error and micro-scale variability of snow cover, vegetation and soil properties (Keller and Gubler, 1993; Bernhard *et al.*, 1998; Ikeda, 2000). This micro-scale variability, expressed by a positive value of the

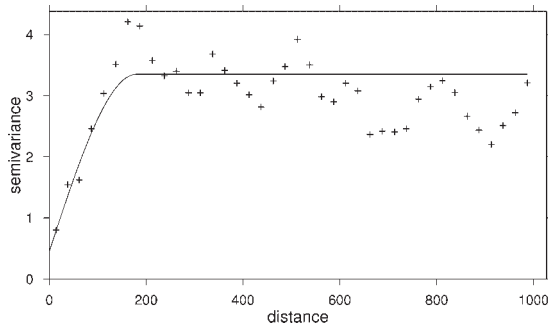


Figure 1 Empirical semivariogram of the residuals of the detrended Zermatt data using model (2). The range of spatial autocorrelation is 180 m, the nugget effect 0.45. Compare Figure 5 of Ødegård *et al.* (1999).

empirical semivariogram close to the origin, is called the *nugget effect* (compare e.g. Cressie, 1993; Goovaerts, 1997). For the Zermatt BTS data, the nugget effect is 0.45. A similar value was obtained by Ødegård *et al.* (1999) (Table 1). In other words, if a normal distribution is assumed, micro-scale variability and measurement error may produce deviations from an average local BTS of up to 1°C with 90% of confidence. This amount of uncertainty could lead to misclassifications between the classes ‘*likely permafrost*’ and ‘*no permafrost*’.

Improved Analysis

A number of methods might help solve the above-mentioned problems with ‘typical’ BTS data (Table 2).

In the case of the Zermatt data, measurement sites were predominantly north-exposed and on average 200 m too high for an optimal model fit at the lower limit of mountain permafrost. In order to compensate for this, an intuitive approach would give more ‘weight’ to those observations located around the expected lower limit of permafrost (approx. 2650–

2850 m ASL) and especially to those that show one of the ‘neglected’ terrain aspect classes. This is achieved by fitting linear regression models with *weighted least squares*, i.e. by minimizing the sum of squared errors $\sum_{i=1}^n w_i e_i^2$, where the w_i are appropriate weights assigned to each individual observation. The weights may be chosen based on the histograms of variables having unfavourable distribution.

The limitations of this method are clear: if there is just a handful of data within an underrepresented aspect or elevation class, one must trust these measurements more than those within a well-represented class. The time and money spent obtaining the numerous measurements on northerly slopes (in the case of the Zermatt data) were, to a large extent, unnecessary.

Statistical weighting accounts for heteroscedasticity, i.e. for a variable measurement variance. The weighting scheme proposed here invalidates prediction variances and renders questionable the *t*-test based significance statements.

With respect to the nugget effect (micro-scale variability) and spatially clustered data, it is recommended to compute the mean values within each cluster and base the model on these aggregated data. In the case of the Zermatt data, 101 groups of up to 20 (median: 4) measurements within diameters <200 m (median: 75 m) were formed manually. In this way, both the nugget effect and the strongest spatial autocorrelations disappear. Within these clusters, or groups of data, values of the explanatory variables are also averaged. This makes sense only if measurement locations within each cluster represent similar environmental conditions. Aggregation with non-uniform cluster sizes should be modelled with an adequate weighting scheme since the variance of estimated cluster means depends on cluster size. Alternatively, spatial autocorrelations may be represented in a regression model using generalized least squares and an appropriate autocorrelation model (Cressie, 1993).

Table 1 Geostatistical characteristics of BTS data from meso-scale studies (Norway, Zermatt) and from local-scale studies on rock glaciers.

	Norway ¹	Zermatt ² , raw data	Laurichard rock glacier ³	Ecuries de Mille rock glacier ⁴
Local variability (variance)	0.41	0.45	0.8	n.a.
—Measurement error (variance)	0.05	n.a.	n.a.	n.a.
Range of autocorrelation	200 m	180 m	80 m	20 m

Note: Parameters given for meso-scale data refer to residual semivariograms (n.a. = not available).

¹Ødegård *et al.* (1999).

²Gruber and Hoelzle (2001).

³X. Bodin, unpublished BTS data measured on the rock glacier’s surface.

⁴Delaloye (2004); range orthogonal to the transverse ridges.

Table 2 Three types of problems that occur during BTS measurements, and suggested approaches for handling them.

	Problem	Consequences	Workaround	Warning
Spatial	Unfavourable distribution of measurement sites	Model not valid for all elevations, expositions, etc.	Use weighted least squares	More weight on certain measurement implies confidence in their precision
	Small distances between measurements (e.g. <100 m)	Autocorrelation, pseudoreplication, incorrect degrees of freedom	Aggregate data that are close together. Use generalized least squares and a model of spatial autocorrelation	Average only data with similar aspects, elevations, etc.
	Nugget effect: local variability at distances <20 m	Individual BTS values do not allow inference on permafrost	Aggregate data that are close together	Average only data with similar aspects, elevations, etc.
	Early/late snow cover	Overall BTS bias	Rest/add 0.1°C	Rough bias correction
	Late snow cover	Short period of constant BTS, negative bias	Use data loggers for bias estimation and removal	Locally variable temporal bias is likely to remain
Climatic	Local early melting	Heat transport; strong positive bias, great spatial variability	Add a binary variable indicating melting	Removing data that are affected by melting may introduce a bias
	Overall early melting	Heat transport; strong positive bias, great spatial variability	None	
Operational	Different observers	Observer bias	(a) Joint measurements, aggregate data; (b) add an observer variable	Avoid confounding with instrument variable
	Use of several probes	Instrument bias	(a) Joint measurements, aggregate data; (b) add an instrument variable	Avoid confounding with observer variable

Table 3 Mean GST100 at four automatic snow observation stations of the Swiss IMIS network for the month of March. Observations for the years 1999/2003, except for STN2 (2000/03).

Code	Name	Elevation	GST100 (March)
ANV2	Anniviers Orzival	2630 m	-0.01°C
ARO2	Arolla Les Fontanesses	2850 m	-0.13°C
STN2	St Niklaus Oberer Stelligletscher	2910 m	-0.82°C
ZER2	Zermatt Triftchumme	2750 m	-0.26°C

TEMPORAL VARIATION OF BTS

Magnitude

An important source of uncertainty in permafrost modelling is the variation of BTS with time. This can be shown by examining the variability of ground surface temperature below a snow cover of at least 100 cm (GST100) as registered at automatic snow observation stations of the IMIS network (Interkanto-

nales Mess- und Informationssystem für die Lawinenwarnung) in the Swiss canton of Wallis (Table 3). These stations are close to the regional lower limit of permafrost. Ground surface temperature data for the Swiss Alps and Hokkaido, Japan, are also used (Kneisel, 1999; Ishikawa and Hirakawa, 2000; Dela-loye, 2004).

Figure 2 shows the temporal development of GST100, snow height and air temperatures between October 1998 and October 2003 at STN2 and ZER2. Both time series show parallel trends due to the short separation distance of 14 km. Two different regimes of the seasonal development of GST100 can be observed:

R1 In the winters 1998/99, 2000/01 and 2002/03, a snow cover ≥ 100 cm formed in November and kept the ground surface insulated at relatively high temperatures. GST100 fell gradually until it reached a value that remained almost constant from January to April. Depending on snow melt, GST100 started to rise in April or May.

R2 By contrast, in the winters of 1999/2000 and 2001/02, GST100 reached very low values at the beginning of February, and then rose steadily until the

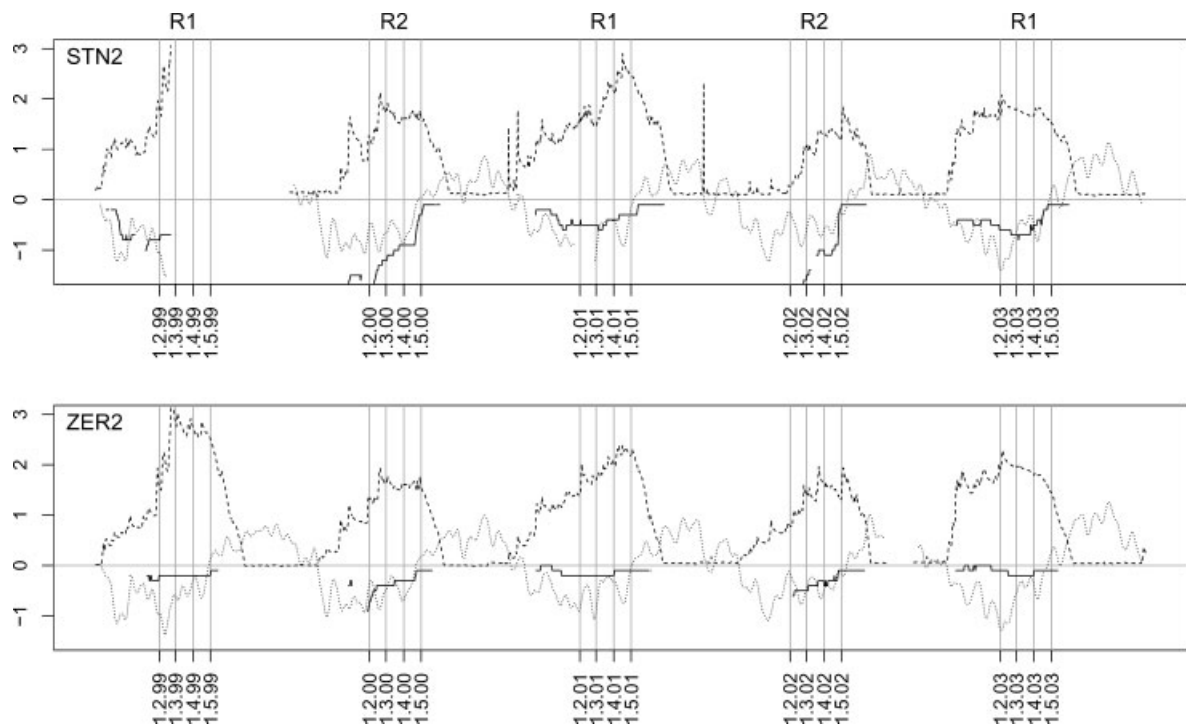


Figure 2 Snow height (dashed line, in m), daily mean ground surface temperature (°C) below ≥ 100 cm of snow (solid line), and smoothed daily mean air temperature (°C, dotted) at STN2 and ZER2. The ground surface temperature regime (R1 or R2) for each year is indicated and explained in the text. Note: breaks in solid lines represent periods when snow cover was <100 cm.

beginning of May. The low values in early winter were due to late accumulation of an insulating snow cover (especially in winter 2001/02 at STN2), and to a melt period in mid-winter (1999/2000), both of which led to lower values of GST100.

Years with R1 regime are better suited for determining BTS than those with the R2 type. March and the first half of April are the most reliable times for BTS measurements, but even then R1 type winters produce higher BTS values than R2 type winters.

More general results can be obtained by examining the deviation of GST100 from its mean value shown in Table 3. The GST100 time series were centred and then averaged between the four stations in order to get a general temporal behaviour without spatial effects. The result is shown in Figure 3 (bold black line). Attention focuses on the month of March, since earlier and later GST100 values are more likely to reflect unstable conditions, and 'true' BTS (or its mean on a temporal scale) is best approximated in this month.

Due to the small amount of years that are available in the analysis, it can only be estimated that R1 type winters (early snow cover) produce an average 0.1°C increase in GST100 in March, while R2 type winters (late snow cover or mid-winter melt) result in a reduction by the same amount with respect to the average. These values do not include local effects. As an example, ground surface temperature data from Kneisel (1999) (below an unknown but stable snow cover) show that inter-annual variation at individual loggers may reach $\pm 0.5^{\circ}\text{C}$ to $\pm 1^{\circ}\text{C}$. This is higher than the value given above but includes local effects.

Figure 3 also demonstrates the large amount of intra-seasonal variation (i.e. the scatter around each

season's mean GST100 for March). This gives a standard deviation of approximately 0.12°C , or $\pm 0.2^{\circ}\text{C}$ at the 90% confidence level. It is not known which part of this variation corresponds to systematic meso-scale trends due to climatic conditions, and to non-systematic local-scale scattering. Ground surface temperature data presented by Kneisel (1999) show intra-seasonal fluctuations of similar magnitude to those given above. Further uncertainty is added if seasonal snow melt in late winter begins before BTS measurements are made (e.g. Ishikawa and Hirakawa, 2000). The local temporal offset of the beginning of seasonal snow melt implies serious constraints for BTS measurement programs. For example, compare the mutually exclusive phases of constant ground surface temperatures at sites BTS1 and BTS3—both situated in 'likely permafrost' areas—in Figure 6 of Ishikawa and Hirakawa (2000). Such conditions hinder the measurement of stable BTS values during a single field campaign.

Diurnal variations of GST100 are in general not measurable at a resolution of 0.1°C . In March, they may vary by $\pm 0.1^{\circ}\text{C}$ around a daily average GST100 (Table 4).

The combined inter-seasonal and intra-seasonal temporal variation can be examined using the annually-repeated BTS measurements made over 8 years at 41 sites on the Ecurie de Mille rock glacier (Delaloye, 2004). The value of $\pm 1^{\circ}\text{C}$ that is obtained is higher than those presented in this paper, partly because the exceptional years 1996 and 1997 are included in Delaloye's time series. Local effects are also probable: For example, Delaloye (2004) suggests that temporal variation of BTS is greater at colder locations. Thus, the results presented above from the Swiss IMIS network may be optimistic because of

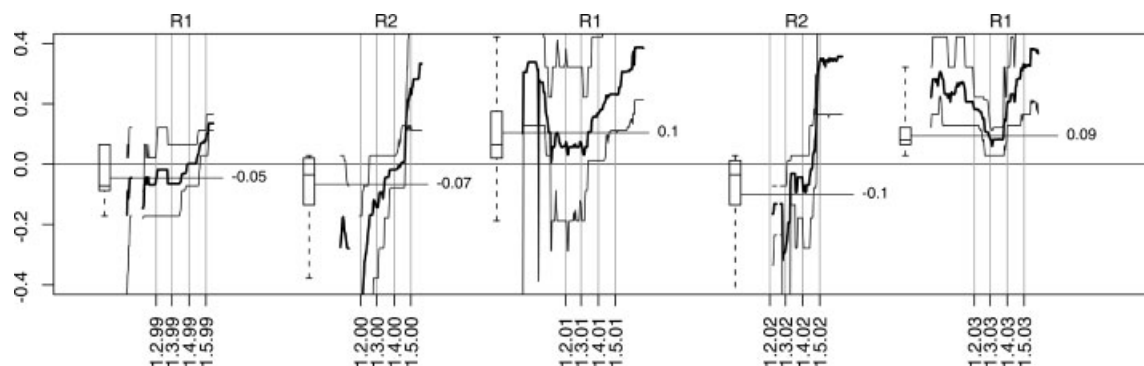


Figure 3 Deviation (in $^{\circ}\text{C}$) of GST100 from mean GST100 of March 1999/2003. The graphs represent mean (bold black line), maximum and minimum (both dark grey) deviation at AVN2, ARO2, STN2 and ZER2. Annual horizontal lines represent the mean of each month of March, and inset box-and-whisker plots summarize the deviation in the month of March.

Table 4 Variation of ground surface temperature below a thick snow cover on different time scales with respect to a multiannual mean. Values represent temporal effects on a regional scale and do not consider local effects.

Temporal scale	Variation at $p = 90\%$	Extreme values
Diurnal variation (March)	$\pm 0.0^\circ\text{C}$	$\pm 0.1^\circ\text{C}$
Intra-seasonal variation (March)	$\pm 0.2^\circ\text{C}$	$\pm 0.4^\circ\text{C}$
Intra-seasonal variation (Feb–Apr)	$\pm 0.3^\circ\text{C}$	$\pm 0.8^\circ\text{C}$
Inter-annual variation	at least $+0.1^\circ\text{C}$ for R1, -0.1°C for R2 locally	± 0.5 to $\pm 1^\circ\text{C}$

relatively high March temperatures (Table 3) and their location outside rock glaciers or other coarse debris surfaces.

Control of Temporal Variation

For permafrost modelling purposes, it is important to reduce systematic deviation of measured BTS from long-term average GST100 or BTS values. This eliminates a bias in model prediction that may be equivalent to a vertical shift of the lower limit of permafrost of several tens of metres. If available, long-term ground surface temperature measurements and snow heights from the surroundings of the study area should be analysed. If an appropriate time series is unavailable, it is suggested that 0.1°C is subtracted from the BTS data in the case of an R1 type winter, and that the same amount is added if an R2 type regime occurs.

Whether these values are appropriate for areas outside the Swiss Alps is not yet known. Likewise, systematic intra-seasonal variation can be detected by using miniature data loggers during one single winter season. However, this bias is difficult to differentiate from spatially heterogeneous non-systematic variation. The intra-seasonal regional average systematic error may always remain in the order of 0.1 to 0.2°C .

Finally, efforts to reduce bias will be beneficial only if the BTS–permafrost relation, as expressed by BTS thresholds θ_1 and θ_2 , is calibrated under well-documented reference conditions.

BTS-BASED MESO-SCALE PERMAFROST MODELLING

General Description of the Problem

The statistical characteristics of BTS data can be described as a random field with external drift and spatially-autocorrelated residuals. It can be written as:

$$\text{BTS}(\mathbf{x}) = \beta^T \mathbf{f}(\mathbf{x}) + e(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^2$$

where $\beta^T \mathbf{f}(\mathbf{x})$ is a deterministic linear trend term, and $e(\mathbf{x})$ is a residual random field assumed to be second-order stationary with mean 0. The range of spatial autocorrelation of the residuals, i.e. the distance up to which the random component of BTS is autocorrelated, is small in the context of meso-scale models, and the nugget effect is considerable when compared to the total sill variance, i.e. the variance reached at distances greater than the range.

In order to model and predict BTS, the most adequate methods will be linear regression and universal kriging (Goovaerts, 1997). The latter will be beneficial only in areas with sufficiently dense data, i.e. where the ‘gaps’ between the measurement locations are smaller than the range.

The final goal of BTS modelling, namely, the meso-scalar prediction of mountain permafrost, has three implications: (1) discrimination between the presence and absence of permafrost is of primary interest, while the actual BTS values are of secondary interest, (2) local mean BTS values should be used in order to reduce micro-scale effects and measurement error, and (3) resolution for statistical prediction should be no finer than ~ 20 – 30 m because variability at shorter distances is dominated by local effects.

Sampling Design

It seems reasonable to cope with the nugget effect of raw BTS data by measuring BTS at each measurement site \mathbf{x} k times and then averaging the measured values $\text{BTS}_i(\mathbf{x})$, $i = 1, \dots, k$, in order to obtain $\text{mBTS}_k(\mathbf{x})$:

$$\text{mBTS}_k(\mathbf{x}) := \frac{1}{k} \sum_{i=1}^k \text{BTS}_i(\mathbf{x})$$

For example, if the nugget variance σ_{nug}^2 is 0.45, as for the Zermatt data, averaging $k = 4$ individual measurements will reduce the nugget standard deviation by the factor $1/\sqrt{k}$ to 0.34°C . In other words, precision is increased by a factor of 2, and more reasonable confidence intervals are obtained for mBTS_4 (90%

confidence interval $\text{mBTS}_4 \pm 0.5^\circ\text{C}$, no misclassification between ‘likely’ and ‘no permafrost’). The k individual measurements should be made within a 10 m diameter characterized by homogeneous environmental conditions. Only the aggregated mBTS_k values should be used for statistical modelling and permafrost indication. Nevertheless, the estimated variances:

$$\widehat{\text{var}}(\text{mBTS}_k(\mathbf{x})) = \frac{1}{k-1} \sum_{i=1}^k (\text{BTS}_i(\mathbf{x}) - \text{mBTS}_k(\mathbf{x}))^2$$

at each measurement site \mathbf{x} may be used for estimating the nugget effect, and furthermore each $\text{mBTS}_k(\mathbf{x})$ may be weighted by the inverse of its estimated variance.

With respect to sample size and linear regression modelling, 50 to 70 measurement sites are sufficient. For example, the size of the aggregated Zermatt data set ($n = 101$) was randomly reduced by resampling without replacement to half or two thirds in a series of trials. Model fits (i.e. parameter estimates) and automated variable selection by AIC-based stepwise methods qualitatively coincide with the results obtained for the full sample size. A further reduction of the sample size to 30 significantly increased uncertainty. Sample sizes needed for logistic regression (see below) may be significantly higher (Hosmer and Lemeshow, 2000).

A second consideration is to achieve a spatial distribution of measurement sites that overrepresents the altitudinal belt surrounding the expected lower limit of mountain permafrost together with a uniform distribution with respect to aspect. The following procedure is suggested:

1. Determine the approximate average lower limit of mountain permafrost within the study area from secondary information (e.g. rock glacier distribution, pilot studies).
2. Select ‘comparable’ catchments (or valley slopes) that are of different orientation.
3. Choose five different altitudes centred at the average permafrost limit. For instance, in the case of Zermatt, around 2350 m ASL (‘no permafrost’ altitude), 2600, 2750, 2900 and 3150 m ASL (approximately following the quartiles of a normal distribution around the average lower limit at 2750 m ASL).
4. Within each selected catchment and at the selected altitudes, determine three measurement sites, e.g. valley floor and both valley-sides. Maintain horizontal distances of more than 150–200 m between

the measurement sites and try to represent different terrain aspects and slope angles uniformly.

5. At each measurement site, take at least $k = 4$ individual BTS measurements within about 10 m of distance.

Several measurement-related parameters may influence observed BTS values and explain part of the nugget variance. First, different probes and individuals may introduce systematic measurement error in part of the sample. Second, signs of melting or other disturbances of the snow cover (such as funnels and voids) should be registered. Different types of disturbance should be tested for statistically significant effects on BTS. This may be important at the local scale, e.g. when BTS distribution on rock glaciers is studied.

Model Design

Two model types can be used for modelling permafrost distribution based on BTS data: (1) linear regression (Gruber and Hoelzle, 2001) and (2) logistic regression (Lewkowicz and Ednie, 2004). The former models BTS values, and three somewhat ‘fuzzy’ classes of permafrost susceptibility are assigned in a second step to the predicted BTS values.

In logistic regression (Hosmer and Lemeshow, 2000), the presence of permafrost is modelled as a binary response variable:

$$\text{PF}(\mathbf{x}) := \begin{cases} 1, & \text{if } \text{BTS}(\mathbf{x}) < \theta_1 \\ 0, & \text{if } \text{BTS}(\mathbf{x}) > \theta_2 \\ \text{NA}, & \text{otherwise} \end{cases}$$

Since probabilities are confined to the interval $[0, 1]$, they cannot be adequately modelled by linear regression models. Logistic regression is therefore performed for so-called *logits* (i.e. the logarithm of the odds $p/(1-p)$, which are the ratio of the probability that something occurs to the probability that it does not occur):

$$\text{logit}(p) := \ln \frac{p}{1-p} \in] -\infty, +\infty [\quad \text{for } p \in]0, 1[$$

The logits are modelled linearly.

The advantage of logistic regression is its ability to predict probabilities (Figure 4, Table 5). On the other hand, linear regression models can more easily answer questions that typically arise when permafrost distribution is analysed. For example, how much is the lower limit of mountain permafrost depressed on north-facing slopes with respect to south-facing slopes? Model equation (2) shows that BTS values measured in north-exposed areas ($\text{NORTH} = +1$) are on average 1.3°C colder than south-exposed areas

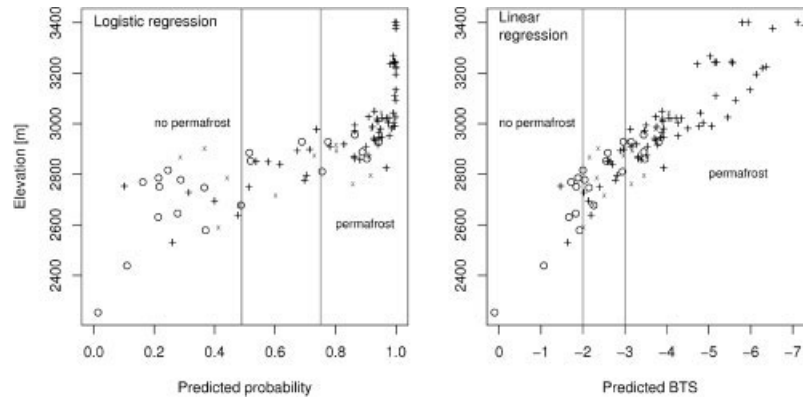


Figure 4 Comparison of predictions made by the logistic regression model (left) and the linear regression model (right) for the probability of permafrost occurrence and BTS, respectively, using the aggregated data set and the explanatory variables Z and PSWR. Based on the BTS field data, plus marks represent 'likely permafrost', grey crosses 'possible permafrost', and empty circles 'no permafrost'. In logistic regression, the BTS thresholds were used to obtain a binary response variable. In the left panel, the vertical lines represent probability thresholds that produce the same number of permafrost and no-permafrost classifications as inferred from BTS values.

Table 5 Qualitative comparison of different approaches to permafrost modelling.

	Linear regression	Logistic regression
Models	BTS values	Presence (1) vs. absence (0) of permafrost
Predicts	BTS values; indirectly: 'fuzzy' permafrost classes ('likely', 'possible', 'no')	Probabilities of permafrost occurrence. Zonation according to mapping units for lowland permafrost is possible.
Answers the question	e.g. How great is the permafrost depression on north-facing slopes?	How much is the permafrost probability increased on north-facing slopes at a given altitude compared to south-facing ones?
Other permafrost indicators	Cannot be included within the model	Can be included

(NORTH = -1; coefficient of NORTH: -0.65) This corresponds to an altitudinal permafrost depression of 200 m (vertical gradient of BTS: 0.61 K per 100 m).

Variable Selection

A fundamental principle in statistics is that of parsimony, i.e. the desire to reduce the complexity of a model without loss of model fit. Possible variables that are used within the models have been presented by Gruber and Hoelzle (2001) and in this paper. Primary topographic attributes (e.g. elevation, aspect, slope) appear to produce model fits and predictions as good as those achieved with more complex topographic attributes (potential solar radiation) or even using variables derived from remotely-sensed data (net short-wave radiation, vegetation index SAVI). This is due to correlation with simple topographic attributes and, in the case of remotely-sensed data, a consequence of the instantaneous character of these

data. However, certain variables will have to be transformed (e.g. by log, sin, cos, square root, power functions) to obtain normally-distributed data or to represent the frequently non-linear relation between explanatory and response variables.

LOCAL-SCALE MODELS

The range of the BTS semivariogram (Figure 1, Table 1) may be considered as the limit between local- and meso-scale. At greater distances, regional effects of topography (elevation, radiation) predominate, while at lesser distances, local effects of soil properties, vegetation and snow cover gain importance.

From an operational perspective, this implies the need for a dense BTS sampling network with between-measurement distances clearly below the range of autocorrelation (Thompson, 2002). In this way, local structures can be mapped, and geostatistical

prediction methods such as universal kriging may be applied. Highest sampling density is required on rock glaciers and other rugged surfaces (Table 1; Delaloye, 2004).

CONCLUSIONS

In the context of meso-scale modelling of BTS values, the implementation of an appropriate sampling design is an important first step in order to obtain well-distributed data and thereby reduce the weaknesses of observational data in complex terrain. A nested sampling design, consisting of a stratified overall distribution and small clusters of local measurement that can be averaged, may be used in order to reduce the effect of nugget variance on BTS data.

Temporal variation of BTS on different scales has to be considered as a major obstacle for the use of BTS as a proxy for permafrost occurrence. Since temporal variation must be expected to show a complex spatial behaviour, even a set of long-term BTS ground surface temperature observation sites cannot be expected to entirely solve the problem. In this case, BTS has to be employed as a relative measure of ground thermal state and not as a permafrost indicator.

ACKNOWLEDGEMENTS

Thanks to Xavier Bodin, Grenoble, for providing unpublished BTS data measured on the Laurichard rock glacier, and to him and Reynald Delaloye, Fribourg, for valuable discussion. The authors wish to thank A. G. Lewkowicz and H. M. French for helpful reviews of the manuscript.

REFERENCES

- Bernhard L, Sutter F, Haeberli W, Keller F. 1998. Processes of snow/permafrost interactions at a high-mountain site, Murtel/Corvatsch, Eastern Swiss Alps. In *Proceedings of the Seventh International Conference on Permafrost, Yellowknife, Canada, 23–27 June 1998*, Lewkowicz AG, Allard M (eds). Nordicana, **57**, Centre d'Etudes Nordiques: Quebec City; 35–41.
- Crawley MJ. 2002. *Statistical Computing: an Introduction to Data Analysis using S-Plus*. Wiley: New York.
- Cressie NAC. 1993. *Statistics for spatial data*. Wiley: New York.
- Delaloye R. 2004. Contribution à l'étude du pergélisol de montagne en zone marginale. GeoFocus 10. Unpublished doctoral thesis, University of Fribourg, Switzerland.
- Etzel Müller B, Ødegård RS, Berthling I, Sollid JL. 2001. Terrain parameters and remote sensing data in the analysis of permafrost distribution and periglacial processes: Principles and examples from Southern Norway. *Permafrost and Periglacial Processes* **12**: 79–92. DOI: 10.1002/ppp.384.
- Fukui K. 2003. Investigating mountain permafrost distribution by ground temperature measurements in the Tateyama Mountains, the northern Japanese Alps, central Japan. *Zeitschrift für Geomorphologie N. F., Suppl.-Bd.* **130**: 179–193.
- Goovaerts P. 1997. *Geostatistics for natural resources evaluation*. Oxford University Press: New York.
- Gruber S, Hoelzle M. 2001. Statistical modelling of mountain permafrost distribution: local calibration and incorporation of remotely sensed data. *Permafrost and Periglacial Processes* **12**: 69–77. DOI: 10.1002/ppp.374.
- Haeberli W. 1973. Die Basis-Temperatur der winterlichen Schneedecke als möglicher Indikator für die Verbreitung von Permafrost. *Zeitschrift für Gletscherkunde und Glazialgeologie* **9**: 221–227.
- Haeberli W. 1978. Special aspects of high mountain permafrost methodology and zonation in the Alps. In *Proceedings of the Third International Conference on Permafrost, Edmonton, Canada*, National Research Council of Canada: Ottawa; 378–384.
- Hoelzle M. 1994. Permafrost und Gletscher im Oberengadin. Grundlagen und Anwendungsbeispiele für automatisierte Schätzverfahren. Unpublished doctoral thesis, Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie der ETH Zürich, 132.
- Hoelzle M, Haeberli W. 1995. Simulating the effects of mean annual air-temperature changes on permafrost distribution and glacier size: an example from the Upper Engadin, Swiss Alps. *Annals of Glaciology* **21**: 399–405.
- Hoelzle M, Haeberli W, Keller F. 1993. Application of BTS measurements for modelling mountain permafrost distribution. In *Proceedings of the Sixth International Conference on Permafrost, Beijing*, Vol. 1. South China University of Technology: Beijing; 272–277.
- Hosmer DW, Lemeshow S. 2000. *Applied logistic regression*. Wiley: New York.
- Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* **54**: 187–211.
- Ikeda A. 2000. Rock glacier morphology, structure and processes in the Upper Engadin, Swiss Alps: the effect of rock types and thermal conditions. Unpublished doctoral thesis, University of Tsukuba, Japan.
- Ishikawa M, Hirakawa K. 2000. Mountain permafrost distribution based on BTS measurements and DC resistivity soundings in the Daisetsu Mountains, Hokkaido, Japan. *Permafrost and Periglacial Processes* **11**: 109–123.
- Keller F, Gubler HU. 1993. Interaction between snow cover and high-mountain permafrost, Murtel/

- Corvatsch, Swiss Alps. In *Proceedings of the Sixth International Conference on Permafrost, Beijing*, Vol. 1. South China University of Technology: Beijing; 332–337.
- Kneisel C. 1999. Permafrost in Gletschervorfeldern—Eine vergleichende Studie in den Ostschweizer Alpen und in Nordschweden. Trierer Geographische Studien 22. Unpublished doctoral thesis, Geographische Gesellschaft Trier in Zusammenarbeit mit dem Fachbereich IV—Geographie/Geowissenschaften der Universität Trier” [Geographical Society Trier in cooperation with the Department IV—Geography/Geosciences of the University of Trier].
- Lewkowicz AG, Ednie M. 2004. Probability mapping of mountain permafrost using the BTS method, Wolf Creek, Yukon Territory, Canada. *Permafrost and Periglacial Processes* **15**: 67–80. DOI: 10.1002/ppp.480.
- Ødegård RS, Isaksen K, Mastervik M, Billdal L, Engler M, Sollid JL. 1999. Comparison of BTS and Landsat TM data from Jotunheimen, southern Norway. *Norsk Geografisk Tidsskrift* **53**: 226–233.
- Thompson SK. 2002. *Sampling*. Wiley: New York.