DATA PROCESSING AND ANALYSIS METHODS

1. Pre-processing of raw temperature time series

At the field sites MID, NON and POP measurements were initiated in summer 2007, whereas, data acquisition started in 2008 for the other field sites. The data time series for this analysis were retrieved between summer 2011 and 2013 for the last time. Figure 2 gives an overview of the data completeness after filtering of invalid data. This filtering comprises two procedures and produces gaps of different characteristics: a) automated filtering of invalid/corrupted values (not numeric or out of realistic range) cause short gaps (gap type I: single values); b) manual filtering of values from broken sensors (e.g. water damage or cable disruption) are applied over long time periods and cause long gaps (gap type II: weeks to months). Maintenance effects on logger operation lead to missing data with the following characteristics: Data read out lead to type I gaps while low power supply lead to type II gaps. The majority of the gaps observed in our data is of type II mainly caused by sensor damage and low battery except for the air temperature on Pink Mt. (Figure 2). Because of these gaps it is not possible to directly compare all time series and simply calculate annual means for the same years. To account for this data characteristic we applied the processing described in section 3.2.

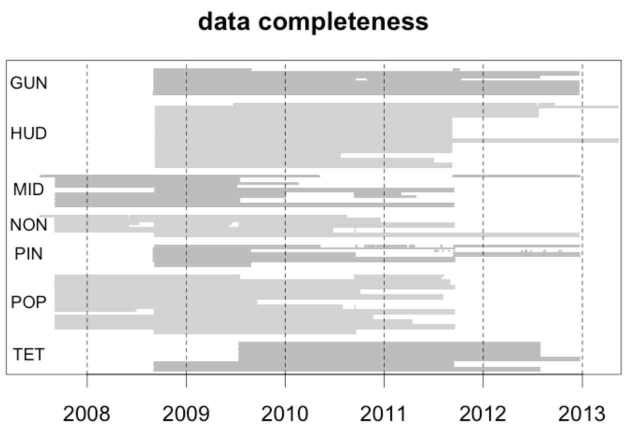

Figure 2: Data compleetness for the seven field sites. Grey bars indicate the presence of valid data.

For all the data analysed in this study there is at least one continuous year of valid data. One exception is the air temperature measurement of the weather station at Mt. Gunnel. The very good correlation of 11 months existing data with the surface temperature recorded in a near-by rock cleft allow a reliable estimation of the mean annual air temperatures. The related data and linear regression model (LRM) is presented in Figure 3. The standard error for a 365 days mean is 0.015°C (residual std. error / sqrt(n) with n= 24*365).

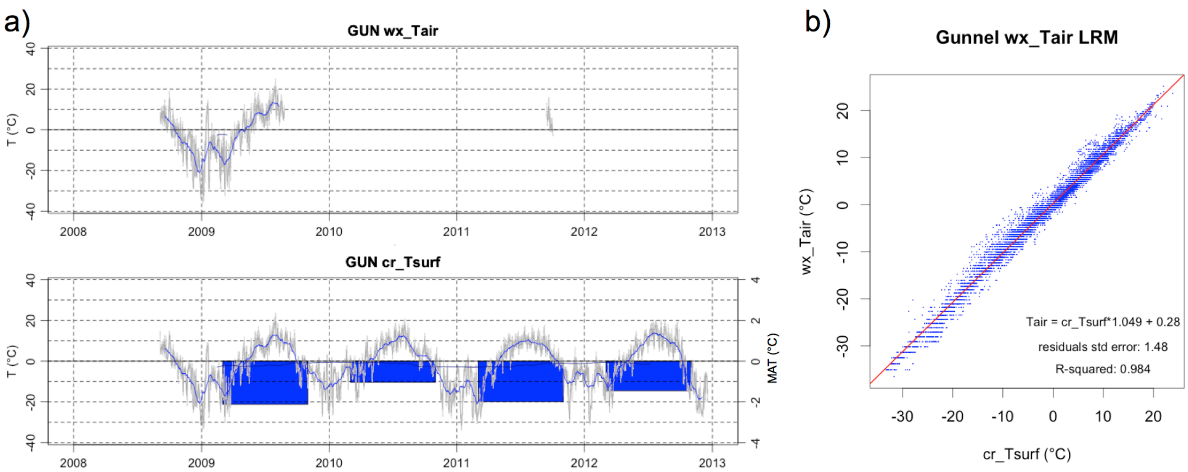

Figure 3: a): Time series of weather station air temperature and surface temperature in rock cleft (grey=raw and blue line = running monthly means) at Mt. Gunnel. Blue bars are mean annual temperatures (MAT) on right scale. b): Linear regression model for wx_Tair with scatter data (blue dots) and regression function (red line).

2. Calculation of mean annual temperatures and their inter-comparability

Annual means of temperature time series (MAT) depend on the averaging period and the completeness of the raw data (Figure 3). Surface and thermal offsets, the differences between such annual means, are highly error sensitive to this mean calculation. As described above, our data is characterized by inconsistencies regarding the available measurement periods and data completeness. To minimize errors introduced by the data aggregation and to avoid misinterpretations of the resulting offsets due to temporal variations, we conduct the following processing steps: 1. Calculate daily mean temperature; 2. Calculate running mean annual temperature; 3. MAT statistics; and 4. Surface and thermal offsets calculation.

The hourly data is aggregated to daily means. Gaps up to two missing values per day are interpolated, if more values are missing no value (*“NA”*) is introduced to the daily mean data. This is important to avoid effects of systematic type I gaps.

Then, *running mean annual temperatures* (RMAT) are calculated for a 365 day window with 99% of data available (Figure 4). In Figure 4 the example of Pink Mt. illustrates possible problems with the inter-comparability of annual means if time series are incomplete or if the running means are asynchronous: MATs from different points in time cannot be easily compared and offsets between RMATs vary strongly for some locations (Figure 4b). This is considered with the next two steps of the data processing, the MAT statistics calculation and the surface/thermal offset calculation.

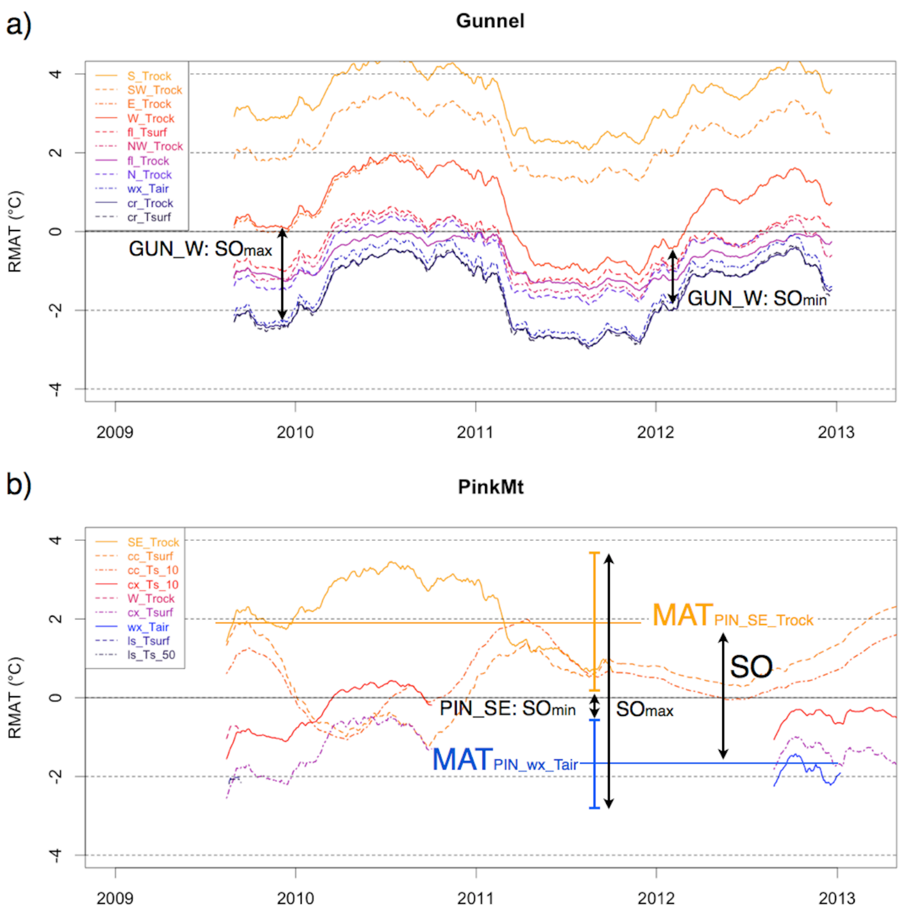

Figure 4: Two examples of running mean annual temperatures (RMAT). a): Mt. Gunnel. Mainly near-surface rock temperatures from steep rock faces without snow. All temperatures have similar variations in annual means. b): Pink Mountain. The annual means from convex and concave landforms show inverse development. Offset terms strongly depend on the point in time of the comparison of instantaneous MATs. These two examples are the best and worst case in terms of data completeness (compare Figure 2). Indications of maximum and minimum surface offset are explained in text.

For the *MAT statistics* the mean and the spread of all RMAT values are calculated (minima, maxima and standard deviation) over the entire available data. This allows us to obtain a MAT value independent of sufficient data within particular calendar years and to get measures for the temporal variation of the MATs. We use the min.- and max.-values as indicator of MAT spread (Figure 4b), because the variations are not normally distributed. For short RMAT time series (below 50% of longest time series at field site), the means and spreads are corrected by using a longer time series as a reference. As a reference the RMAT time series from the same field site with the best correlation during the overlapping time period is chosen (e.g. cx_Tsurf for wx_Tair at Pink Mt., Figure 4). The mean annual temperature of such a short temperature time series is estimated by:

$$MAT = \overline{RMAT_{overlap}} - \overline{RMAT_{ref\_overlap}} + \overline{RMAT_{ref}} \qquad (1)$$

with $RMAT_{overlap}$ being the RMAT time series of interest where it overlaps with the reference, $RMAT_{ref\_overlap}$ being the reference time series of the same period and $RMAT_{ref}$ being the mean of the entire reference time series. The uncertainties of this estimate are approximated by

$$U_{MAT} = \sigma\left(RMAT_{overlap} - RMAT_{ref_{overlap}}\right) * \frac{\sigma(RMAT_{ref})}{\sigma\left(RMAT_{ref_{overlap}}\right)} \qquad (2)$$

assuming that the variations of the difference of the two variables scales with the relative variance captured by the overlapping period. This uncertainty is then added/subtracted to max.-/min.-values of the MAT statistics of the temperature of interest, resulting in a larger spread for shorter time series (higher uncertainty; cf. Figure 4b *PIN_wx_Tair*). The uncertainty introduced by measurement errors and data pre-processing (about ±0.2°C) are small compared to the temporal variations and negligible for the MAT statistics if considered as independent errors (root of square sum). Figure 5 shows the result of the MAT statistics on the example of Pink Mountain, which is a worst case in terms of data continuity: *W_Trock* or the *ls_Tsurf* MAT values show spreads up to 4.5 °C due to the limitation of the available data.

The *surface and thermal offsets* are calculated by a simple subtraction of the MAT values. However, the spread of these offsets is smaller than spreads of the MATs in many cases because the temporal variations of RMATs are often synchronous (cf. Figure 4a). Given this, the minimum and maximum difference between the RMATs (e.g. $SO_{min/max}$ at GUN_W) is used as indication of the uncertainty of the offsets, introduced by temporal variation of the SO (Figure 4a). Where only short time series (less than 50% of longest time series) or no overlap between RMATs exists, the spreads of the MAT statistics are added to obtain a conservative (large) uncertainty estimate (cf. $SO_{min/max}$ of PIN_SE in Figure 4b). To these uncertainty values an independent measurement error of ±0.3 °C is added (square root of quadratic sum) to obtain a total uncertainty for each temperature offset (c.f. Hasler et al. 2011).

$$U_{offset} = \sqrt{\left(Offset_{\max} - Offset_{\min}\right)^2 + 0.6^2} \qquad (3)$$

For the Hudson Bay Mt. field site, where the air temperature is measured at a meteorological station at 300m to 500m lower elevation (Table 1), an air temperature lapse rate between –3.75 to –6.25 °C/km is used for the calculation of the mean annual air temperature (MAAT), the SO and its uncertainty. In Figure 5b an example of a temperature profile shows the surface and thermal offset at one location at Pink Mt. In the further analysis, offsets are treated as significant (solid lines) if they are larger than the (inner) half of the uncertainties ($U_{offset}/2$) indicated by the spreads in Figure 5b.

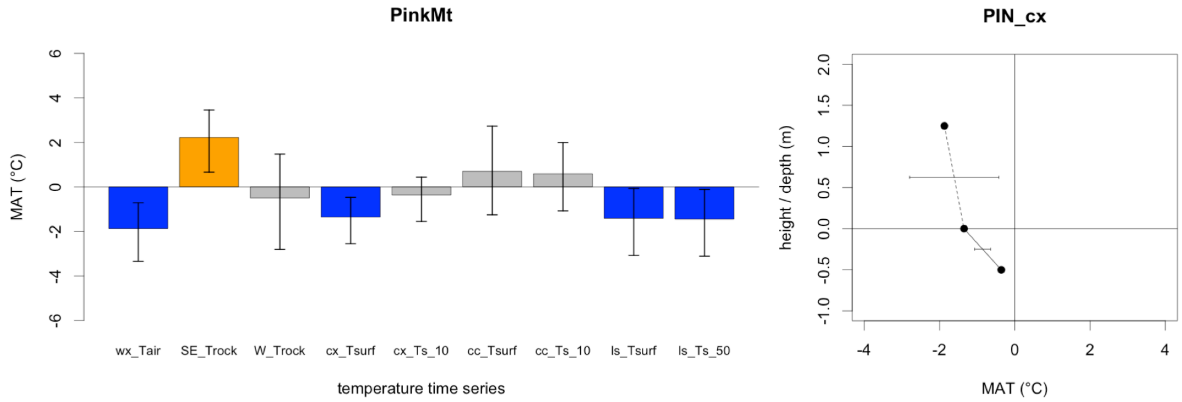

Figure 5: Examples from Pink Mountain of variation in MAT and uncertainties in offsets. left: Mean annual temperatures and spread for general intercomparison, orange bars are RMAT values above zero of enire measurement period, blue bars are entirely sub-zero RMATs; right: MAT profile with spread (U_{offset} /2) of the surface / thermal offsets. Solid lines indicate offsets that are larger than the spreads, dashed lines are used if offset is equal or smaller than the spread. The temperature at 1.4 m height is the mean annual air temperature (MAAT).

The annual temperature amplitudes (ATA) used in this article are the differences between mean July and January temperatures divided by 2. The N-factors used in the discussion of the snow cover influence are calculated on a seasonal and biweekly basis by dividing the mean surface temperature (or freezing index If) through the mean air temperature (T_{surf}/T_{air} or If_{surf}/ If_{air}). Because the relative errors of these calculations are much smaller, we do not detail these calculations and their uncertainties here.