A new approach to quantifying soil temperature responses to changing air temperature and snow cover

Michael C. Mackiewicz

Department of Biological and Physical Sciences, College of Southern Maryland, 8109 Cobden, Ct, # 202, Manassas, VA 20109, USA

Received 29 January 2011; revised 17 April 2012; accepted 22 June 2012

Available online 14 July 2012

Abstract

Seasonal snow cover provides an effective insulating barrier, separating shallow soil (0.25 m) from direct localized meteorological conditions. The effectiveness of this barrier is evident in a lag in the soil temperature response to changing air temperature. The causal relationship between air and soil temperatures is largely because of the presence or absence of snow cover, and is frequently characterized using linear regression analysis. However, the magnitude of the dampening effect of snow cover on the temperature response in shallow soils is obscured in linear regressions. In this study the author used multiple linear regression (MLR) with dummy predictor variables to quantify the degree of dampening between air and shallow soil temperatures in the presence and absence of snow cover at four Greenland sites. The dummy variables defining snow cover conditions were $z = 0$ for the absence of snow and $z = 1$ for the presence of snow cover. The MLR was reduced to two simple linear equations that were analyzed relative to $z = 0$ and $z = 1$ to enable validation of the selected equations. Compared with ordinary linear regression of the datasets, the MLR analysis yielded stronger coefficients of multiple determination and less variation in the estimated regression variables.

© 2012 Elsevier B.V. and NIPR. All rights reserved.

Keywords: Soil; Air; Temperatures; Snow cover; Regression

1. Introduction

Air temperature is the primary factor influencing shallow soil temperatures. However, in regions having seasonal snowfall, the absence or presence of snow cover significantly influences shallow soil temperatures. Zhang (2005) reported that snow accumulation rates, snow thickness, cover duration, and the low thermal conductivity and internal structure of snow also play roles in controlling soil temperature, in addition to the presence or absence of snow. These factors make snow an effective insulation barrier that creates an observable lag in the thermal response of a soil relative to changing air temperature. Superimposed on seasonal snow cover and meteorological conditions, soil temperatures are also influenced by irregular episodes of cloud cover, cold or warm spells, precipitation (rain and snow) events, and droughts (Beltrami, 2001; Decker et al., 2003; Goodrich, 1982; Sokratov and Barry, 2001; Thorn et al., 1999). The central importance of these factors is that they affect soil temperature, which in turn is a major factor affecting soil physical, chemical, and biological processes in cold regions.

An understanding of the thermal regime of shallow soils in response to air temperature and snow cover conditions is central to understanding surface and near surface microbiological activity; plant root growth; soil
and pore-water chemistry; and impacts on engineered structures, surface water flow and infiltration, potential contaminant migration pathways and remediation alternatives; and most recently in predicting past and future climates. Therefore, the author quantified the effect of the degree and extent of snow cover on soil temperatures recorded at four sites in Greenland (Fig. 1). On-going analysis associated with the latter area of research is in progress.

2. Data presentation

Temperature–time plots of soil and air data similar to those in Fig. 2 are commonly used to illustrate a relationship between soil and air temperatures under differing meteorological conditions. However, their use as an interpretative tool illustrates only a causal relationship. Re-plotting the data (Fig. 3) demonstrated that other independent variable(s) were involved in influencing the strength and direction of this relationship. The next step in characterizing this relationship was to fit the cross-plot air and soil temperature data to a best fit linear regression (Fig. 4). However, from Fig. 4 it is evident that the dataset consisted of two distinct populations. While the presentation format (Figs. 2 and 3) and the linear regression method of data analysis was valid for a first order approximation of the relationship between soil and air temperatures, it did not enable identification of the parameter(s) contributing to the relationship. A better characterization of the relationship, which still included the cross-plot format and best fit linear regression, involved analysis of each data cluster by calculating two distinct linear equations to describe each dataset (Fig. 4).

Fig. 1. Location of the study sites (Ilulissat, Sondre, and Sisimiut) in Greenland.
However, it was unclear to what extent the two separate linear equations differed, and whether the single linear equation was less meaningful than the two best fit linear equations. To address these concerns a multiple linear regression (MLR) analysis incorporating dummy predictor variables was used. MLR facilitates presentation of the data, and analysis of the data in a single best fit linear equation. This best-fit linear equation can then be reduced to two or more ordinary linear equations, the number of linear

---

**Fig. 2.** Plot of air and soil temperatures illustrating apparent casual but lagged association between the response of soil temperature to changes in air temperature. Temperature data (a subset of the population recorded) from the NSDIC dataset, recorded at Ilulissat, Greenland (Olesen, 2003).

**Fig. 3.** Cross-plot of the air and soil temperature data shown in Fig. 2. A slightly stronger association between soil and air temperature is evident in the poor to moderate linear response of soil temperature to changes in air temperature. The solid line represents the best-fit linear regression line for all 928 data points. The data pattern strongly suggests that two different data populations are present. Based on database information, the two populations include temperatures recorded in the presence and absence of snow cover.
equations being dependent on the number of assigned
dummy predictor variables. MLR analysis also enables
the investigator to validate the analysis, assess its
strength, and evaluate whether the resulting equations
are statistically different. Multiple regression using
dummy variables yields the same inferences, and is
statistically equivalent to multiple analysis of variance
(MANOVA). Other parameters involved in the
observed variations, including snow cover thickness
and soil-water content, will be investigated separately.

3. Equation selection

MLR incorporating dummy variables can be used to
quantify the effects of snow cover on shallow soil
temperature. MacLean and Ayres (1985) used regression
models based on data from Barrow, Alaska, to demon-
strate that the independent variables mean air temperature,
mean daily sky cover, and the fraction of a day that a site
receives no direct sun exposure can be used to predict soil
temperature. They concluded that their model has general
applicability in estimating soil temperatures from avail-
able meteorological data. Other studies (Bockheim and
Hall, 2002; Hoelzle et al., 2002; Sokratov and Barry,
2001; Thorn et al., 1999; Zhang et al., 2003) have shown
that soil temperature is influenced by air temperature, soil
moisture, soil thermal properties, and snow cover. A
multiple linear regression analysis (MLRA) incorporating
dummy predictor variables was performed on soil and air
temperature data and snow cover data recorded at three
sites in Greenland: Ilulissat (1968–1981); Sondre A
(1968–1976) and B (1968–1970); and Sisimiut
(1969–1982). The data collected by Olesen (2003) and
Olesen et al. (2003a,b) are included in, and available from,
the database of the National Snow and Ice Data Center
(NSIDC) World Data Center for Glaciology.

The author quantified the magnitude of the effect of
snow cover, or absence of cover, on the temperature
response of shallow soil (approximately 25 cm depth)
as air temperature changed at the three study sites. The
validity of the MLRA results was assessed by testing
the null hypothesis by comparing the slopes, intercepts,
and tests of coincidence for each intra-site and inter-
site derived equation.

4. Study dataset

In 1967 the Greenlandic Geological Survey (GEUS)
began monitoring soil temperatures in permafrost and
seasonal frost areas at four sites (including one dupli-
cate monitoring station) in western Greenland (Olesen,
1967; Van Tatenhove and Olesen, 1994), as part of the
UNESCO International Hydrological Decade program.
The four Greenland sites included one permafrost
and strength of the estimated relationship between soil temperature (dependent variable) and air temperature (independent variables) as affected by the occurrence of snow cover. Regression analysis is often used when the independent variable cannot be controlled. This was the case in the present study, which relied on observational data including snow cover and air temperature. However, misinterpretation of results can occur in attempting to quantify a relationship between two or more variables using the estimated correlation coefficient. The computational methods may be correct, but the estimate itself may be biased. Biased results can arise from parameter selection, data errors, or from unidentified variables unaccounted for in the analysis. Introduction of the variable $z$ in the analysis enabled the author to represent subgroups of the sample population of soil temperatures recorded in the presence ($z = 1$) or absence ($z = 0$) of snow cover, and facilitated the use of a single regression equation to represent multiple groups.

Regardless of the risks of misinterpretation associated with biases noted above, the analysis provided a reliable means for identifying relationships among variables. Results of such analyses can be used to quantify the strength of an observed relationship by estimating the coefficient of determination ($r^2$), which defines the percentage of the variance observed in the dependent variable (in this study, soil temperature) that can be explained by a linear relation with the independent variable (air temperature). The apparent relationship between two or more sets of random numbers is determined by estimating the significance of the correlation that calculates the probability of exceeding the observed correlation coefficient by chance. The significance test for $r^2$ is an analysis of variance, which tests whether the variance modeled by the linear regression is significantly greater than the variance not modeled, or the residual variance. In addition, determining whether the overall regression is significant requires testing of the null hypothesis, $H_0$. In this study $H_0$ was that soil temperature cannot be predicted from the air temperature and snow cover thickness any more accurately than by guessing the soil temperature. The MLRA methodology used in this study followed that reported by Mendenhall and Beaver (1991).

6. MLRA method

The MLRA relates a response variable $y$ to a set of predictor variables $x_1, x_2, \ldots, x_n$ using a multiple regression equation, which facilitates estimation and prediction of the mean value of $y$ given known values of $x_1, x_2, \ldots, x_n$. Predictor variables contributing information to the...
estimation of \( y \) can be quantitative and/or qualitative. The MLRA the author used to estimate the soil temperature response to air temperature for periods of snow cover and no snow cover included air temperature as a quantitative predictor variable, and snow cover occurrence as qualitative predictor variable. The qualitative variable (snow cover) could only assume two values in the equation; snow cover and no snow cover. Therefore, the predictor variable ‘snow cover’ was introduced into the equation using the dummy variable (\( z \)) as:
\[
T(s) = \beta_0 + \beta_1 x + \beta_2 z
\]
\[
z = \begin{cases} 
1 & \text{snow cover present} \\
0 & \text{no snow cover} 
\end{cases}
\]

where \( T(s) \) is the soil temperature (°C); \( \beta_0, \beta_1, \) and \( \beta_2 \) are regression coefficients; \( x \) is the air temperature in °C; and \( z \) is a predictor variable introduced into the equation using a dummy or indicator variable, as noted.

Therefore, in presence of snow cover \( z = 1 \) and the regression line is:
\[
T(s) = (\beta_0 + \beta_2) + \beta_1 x 
\]
while in the absence of snow cover \( z = 0 \) and the regression line is:
\[
T(s) = \beta_0 + \beta_1 x
\]

The result is that both lines have different intercepts but the same slope, and therefore define parallel lines. To illustrate the effect of snow cover on soil temperature, the slopes of the two lines must be unequal. This indicates that the two predictor variables interact; that is, the change in \( T(s) \) corresponding to a change in ‘\( x \)’ depends on whether snow cover is present or absent. To allow for differences in slopes, the interaction term \( xz \) was introduced into Eq. (1), which becomes:
\[
T_{\text{soil}} = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 xz
\]

This yields the following separate linear equations:
\[
z = 1 \quad T(s) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x \quad (5)
\]
\[
z = 0 \quad T(s) = \beta_0 + \beta_1 x \quad (6)
\]

Therefore, the regression coefficients for Eqs. (2) and (3) in terms of the coefficients of Eqs. (5) and (6) can be written as follows:

Snow present : for intercepts, \( (\beta_0 + \beta_2)_{\text{Eq}2} = (\beta_0 + \beta_2)_{\text{Eq}5} \); for slopes, \( \beta_{1\text{Eq}2} = (\beta_1 + \beta_3)_{\text{Eq}5} \)

Snow absent : for intercepts, \( \beta_{0\text{Eq}3} = \beta_{0\text{Eq}6} \); for slopes, \( \beta_{1\text{Eq}3} = \beta_{1\text{Eq}6} \)

Thus, Eq. (4) incorporates the two separate linear equations within the single regression model, and allows for different intercepts (\( \beta_0 \) for no snow cover; \( \beta_0 + \beta_2 \) for snow cover) and slopes (\( \beta_1 \) for no snow cover; \( \beta_1 + \beta_3 \) for snow cover). This process enabled Eq. (4) to be used to estimate the influence of snow cover on soil temperature. Regression models for the soil temperature response (in the presence or absence of snow cover) relative to air temperature change were applied to the combined dataset from the three sites (intra-site) and for each site individually (inter-site).

7. MLR equation validation

Verification that each intra-site and inter-site MRA equation contributed to the prediction of \( y \) (i.e. soil temperature) under snow cover and no snow cover scenarios required testing of the null hypothesis. The null hypothesis was that the calculated regression lines for predicting soil temperature, regardless of whether snow cover was present or absent, have equal slopes or intercepts and are coincident. Performing the test of parallelism, intercept, and coincidence on each model provided a means of interpreting the selected equation results. The appropriate null hypotheses for comparing the slopes, intercepts, and tests of coincidence for each intra-site and inter-site MRA equation are presented below.

7.1. Test of parallelism

The null hypothesis, that the two regression lines for each MRA are parallel, is equivalent to \( H_0: \beta_1 = 0 \). If \( \beta_3 = 0 \), then the slope of the regression line estimating soil temperature under snow cover conditions, \( \beta_1\ \text{SNOW COVER} = \beta_1 + \beta_3 \), is simplified to \( \beta_1 \). This is equivalent to the slope of the equation for predicting soil temperatures under no snow cover scenarios; consequently, we accept the null hypothesis, which simply states that snow cover has no influence on soil temperature. The null hypothesis for the test of parallelism is determined by calculating the partial F-statistic using:
\[
F(x|z,xz) = \frac{\text{SSR}(x,z,xz) - \text{SSR}(x)}{\text{MSR}(x,z,xz)},
\]
where \( \text{SSR} \) is the regression sum of squares; \( \text{MSR} \) is the mean square error; and \( x, z, \) and \( xz \) are the variables defined in Eq. (1).
7.2. Test of equal intercepts

The null hypothesis that the two intercepts for the regression lines determined for each intra-site and intersite MRA are equal (which allows for unequal slopes) is equivalent to H0: \( \beta_2 = 0 \) for the overall MRA. This test compares Eq. (4) with the reduced equation:

\[
T_{soil} = \beta_0 + \beta_1x + \beta_2xz
\]

The null hypothesis is then tested by calculating the partial F-test using:

\[
F(z|x) = \frac{SSR(x,z) - SSR(x)}{MSR(x,z,xz)}
\]

where SSR is the regression sum of squares; MSR is the mean square error; and \( x, z, \) and \( xz \) are the variables defined in Eq. (1).

The hypothesis that the two regression lines overlap is H0: \( \beta_2 = \beta_3 = \beta \). When both \( \beta_2 \) and \( \beta_3 \) equal 0, the equation for snow cover, Eq. (2), becomes the no snow cover equation, Eq. (3), which demonstrates that both lines coincide. The null hypothesis for performing the test of coincidence is determined by calculating the partial F-test using:

\[
F(xz|x) = \frac{[SSR(x,z,xz) - SSR(x)]/2}{MSR(x,z,xz)}
\]

The results of the tests of parallelism, intercept, and coincidence provide the means to interpret the selected results. Interpretation of a particular intra-site or intersite linear equation leads to one of four likely outcomes when assessing the validity of the derived regression equation. The first outcome is that the regression lines calculated under each snow condition (absence and presence) are parallel. This implies that estimates of soil temperature will be consistently greater under one of the snow cover conditions, but that the rate of temperature change will be equivalent regardless of whether snow cover is present or absent. The second outcome is that both the estimated regression lines share the same intercept but have differing slopes. Under this scenario, estimates of soil temperature in either snow cover scenario will have a common temperature point, but at other points the temperatures will differ because of other factors. The third outcome is that the calculated lines have different slopes and intercepts, which suggests that the relationship between soil temperature and air temperature differs under conditions of snow cover and no snow cover. The fourth outcome is that the regression lines are coincident, indicating that there is no apparent effect of the presence or absence of snow cover on soil temperature.

7.3. Intra-site data analysis

The combined three-site soil and air temperature dataset (\( n = 929 \)) under snow and no snow conditions is shown in Fig. 4; the data were fitted using Eq. (4). The MLRA statistics are shown in Table 1, and the calculated regression line is included in Fig. 4. The combined dataset for the three sites was modeled using Eq. (4), and incorporation of the dummy variables (\( z = 0 \) and 1) indicated that there was a statistically significant positive correlation (\( r^2 = 0.652 \)) between air and soil temperatures, with 65% of the observed variation in soil temperature able to be explained by air temperature. If the combined dataset is representative of soil responses to fluctuations in air temperature in the area encompassing the three sites, then it is estimated that a 1°C change in air temperature in the area will result in a 0.38°C change in soil temperature (\( \beta_1 \) in Table 1) during periods of no snow cover, and a 0.19°C change when the surface is covered by snow. The results for each of the three sites (including all data points under the \( z = 0 \) and \( z = 1 \) scenarios) are presented in Table 1, and illustrate that the MLRA provides a statistically reliable means for predicting soil temperature responses relative to air temperature in the presence or absence of an insulating cover of snow. However, the intra-site model equation provided only

<table>
<thead>
<tr>
<th>Site</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \beta_3 )</th>
<th>MSE</th>
<th>( F )</th>
<th>( r^2 )</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined</td>
<td>0.96*</td>
<td>0.38*</td>
<td>−0.19*</td>
<td>−3.05*</td>
<td>7.82</td>
<td>578.7</td>
<td>0.652</td>
<td>928</td>
</tr>
<tr>
<td>Ilulissat</td>
<td>0.52*</td>
<td>0.42*</td>
<td>−0.02*</td>
<td>−2.84*</td>
<td>6.37</td>
<td>242.5</td>
<td>0.699</td>
<td>316</td>
</tr>
<tr>
<td>Sondre A</td>
<td>0.803*</td>
<td>0.440*</td>
<td>−0.285*</td>
<td>−4.063*</td>
<td>9.878</td>
<td>163.6</td>
<td>0.723</td>
<td>192</td>
</tr>
<tr>
<td>Sondre B</td>
<td>1.038***</td>
<td>0.409*</td>
<td>−0.278*</td>
<td>−4.052*</td>
<td>6.125</td>
<td>86.0</td>
<td>0.799</td>
<td>69</td>
</tr>
<tr>
<td>Sisimiut</td>
<td>1.66*</td>
<td>0.23*</td>
<td>−0.11*</td>
<td>−3.34*</td>
<td>7.080</td>
<td>129.6</td>
<td>0.528</td>
<td>352</td>
</tr>
</tbody>
</table>

Table 1
Soil temperature MLRA results for the combined data from the three Greenland sites, and for each site individually. Temperature data include site conditions under both snow and no snow scenarios. *\( \beta \) values are significant at 99% confidence level by \( t \)-test; **\( \beta \) values are significant at 95% confidence level by \( t \)-test; ***\( \beta \) values are significant at 90% confidence level by \( t \)-test; +\( \beta \) values are significant at <90% confidence level by \( t \)-test. All \( F \) values are significant at 99.99% probability. MSE is the residual mean square error. \( r^2 \) measures the proportion of variance in the dependent data explained by the regression, and \( n \) is the number of data points.
soil medium. The insulating effect of snow varies with
as an insulating factor, and the thermal properties of the

7.4. Inter-site data analysis

The MLRA of the combined temperature dataset showed a good positive correlation between soil and air temperatures in the presence and absence of snow cover, but provided no information regarding the insulating effects of snow cover at each site. Therefore, the defined conditions when \( z = 0 \) and \( z = 1 \) were introduced into Eq. (4) for each site, and the resulting Eqs. (2) and (3) were compared with a separate ordinary linear regression analysis (OLR) of air temperature versus soil temperature in the presence and absence of snow cover for each site. The linear equations resulting from the MLRA and the OLR were identical.

The results of the MLRA of the air and soil temperatures and snow cover data from the Ilulissat site indicated that although the linearity among the variables was poor \( (r^2 = 0.237, \text{ Table 2}) \), there was a positive correlation between air and soil temperatures. This suggests that in the presence of snow cover there will be a soil temperature change of \( 0.23 \degree C \) \( (\beta_1 \text{ in Table 2}) \) in response to a \( 1 \) \degree C change in air temperature. In the absence of snow cover there will be a soil temperature change of \( 0.42 \) \degree C \( (r^2 = 0.443; \beta_1 \text{ in Table 3}) \) with a corresponding \( 1 \) \degree C change in air temperature. The MLRA data for the Sondre A and B sites and the Sisimiut site (Tables 2 and 3) all showed similar responses, in that there were greater changes in soil temperature in response to changes in air temperature when snow cover was absent \( (z = 0) \). These observations, and the greater soil temperature \( (\beta_1) \) change in response to air temperature when \( z = 0 \), are attributable to the absence of snow as an insulating factor, and the thermal properties of the soil medium. The insulating effect of snow varies with

<table>
<thead>
<tr>
<th>Data set</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>MSE</th>
<th>( F )</th>
<th>( r^2 )</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined data: 3-sites</td>
<td>-2.09*</td>
<td>0.18*</td>
<td>9.45</td>
<td>140.87</td>
<td>0.189</td>
<td>586</td>
</tr>
<tr>
<td>Ilulissat</td>
<td>-2.32*</td>
<td>0.23*</td>
<td>8.13</td>
<td>67.37</td>
<td>0.237</td>
<td>208</td>
</tr>
<tr>
<td>Sondre A</td>
<td>-3.26*</td>
<td>0.16*</td>
<td>12.65</td>
<td>15.55</td>
<td>0.129</td>
<td>107</td>
</tr>
<tr>
<td>Sondre B</td>
<td>-3.01*</td>
<td>0.13**</td>
<td>8.76</td>
<td>4.88</td>
<td>0.126</td>
<td>36</td>
</tr>
<tr>
<td>Sisimiut</td>
<td>-1.68*</td>
<td>0.12*</td>
<td>7.8</td>
<td>26.9</td>
<td>0.1</td>
<td>236</td>
</tr>
</tbody>
</table>

Table 3
MLRA results for soil and air temperatures recorded under site conditions of a blanket of snow. Notes on *\( \beta \) values, **\( \beta \) values, ***\( \beta \) values, \( F \) values, MSE, \( r^2 \), and \( n \) are as provided in Table 1.

Further analyses were performed to verify that the derived regression equations defined differences in observed soil temperatures in response to changing air temperature in the absence or presence of snow cover. These included comparisons of calculated slopes, intercepts, and coincidence of the regression equation under \( z = 0 \) and \( z = 1 \) conditions. For each intra-site and inter-site MRA, Eqs. (7), (9) and (10) were used to test the null hypothesis comparing the slopes, intercepts, and tests of coincidence, respectively.

The results of the tests of parallelism for each intra-site and inter-site equation indicated that for each model the null hypothesis was rejected (Table 4, \( F > 1 \)), and that the two slopes for each equation used for estimating soil temperature in the presence or absence of snow cover were statistically different. This indicates that soil temperature responds less to changes in air temperature when the ground is covered by snow \( (\beta_1 \text{ in Table 2, snow cover present}; \text{ Table 3, snow cover absent}) \).

The null hypothesis for the test of equal intercepts was that for each intra-site and inter-site model the intercepts of both regression equations for estimating soil temperature under different snow cover conditions are equivalent or equal to zero. The calculated partial \( F \)-test statistic (Table 3) for each model shows that the null hypothesis was rejected, as the intercepts were significantly different. The null hypothesis for the test of coincidence was also rejected for each model (Table 4, \( F > 1 \)). The calculated \( F \)-statistic showed that regression lines for each model did not overlap, which indicates that the response of soil temperature to changes in air temperature differed in the presence and absence of snow cover.
The 929 data points used in analyses of the soil and air temperatures are represented in Fig. 5. The cross plot shows a moderate—poor linear relationship between the recorded air and soil temperature for all data points, which is depicted by the solid linear best fit line (n = 929, $r^2 = 0.652$; Table 1). It is also apparent from Fig. 5 that the data in the lower left quadrant of the graph (representing the presence of a snow cover at the four sites) are more scattered than the data in the upper right quadrant. The observed scatter in the data is not attributable solely to the presence of snow cover. Snow characteristically has a low thermal conductivity relative to other materials, thereby making it an effective thermal insulator (Sokratov and Barry, 2001; Zhang, 2005) that retards heat exchange between the soil and the overlying atmosphere. Therefore, the snow cover at the four Greenland sites results in the underlying soil maintaining temperatures several degrees warmer than the overlying atmosphere (Fig. 5). Other factors controlling soil temperature include the timing and duration of seasonal snow cover and snow density (Ling and Zhang, 2003), and changes in snow depth (Ge and Gong, 2009; Ling and Zhang, 2003).

From Fig. 6a (Ilulissat), b (Sondre A), c (Sondre B), and d (Sisimiut) it is apparent that the insulating effect of snow dampens the soil response to air temperature changes. The MLR analysis (Tables 1−3) enabled quantification of the dampening effect of the snow cover. The soil temperature response to a 1°C change in air temperature at the four Greenland sites was 31−45% lower during periods of snow cover compared with soil temperatures recorded

### Table 4
MLRA results for tests of parallelism, intercept, and coincidence on each equation, verifying that each intra-site and inter-site MLRA contributes to the prediction of y (i.e. soil temperature) under snow cover and no snow cover conditions.

<table>
<thead>
<tr>
<th>MLR model validations</th>
<th>Site</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow:no snow</td>
<td>Combined</td>
<td>30.08</td>
<td>0</td>
</tr>
<tr>
<td>slope comparison</td>
<td>Ilulissat</td>
<td>7.61</td>
<td>0.0061</td>
</tr>
<tr>
<td></td>
<td>Sondre A</td>
<td>13.82</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>Sondre B</td>
<td>13.11</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>Sisimiut</td>
<td>2.8</td>
<td>0.0951</td>
</tr>
<tr>
<td>Snow:no snow</td>
<td>Combined</td>
<td>163.9</td>
<td>0.000</td>
</tr>
<tr>
<td>intercept comparison</td>
<td>Ilulissat</td>
<td>58.73</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Sondre A</td>
<td>30.18</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Sondre B</td>
<td>11.08</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>Sisimiut</td>
<td>77.6</td>
<td>0.000</td>
</tr>
<tr>
<td>Snow:no snow</td>
<td>Combined</td>
<td>958.62</td>
<td>0.000</td>
</tr>
<tr>
<td>coincidence of</td>
<td>Ilulissat</td>
<td>33.134</td>
<td>0.000</td>
</tr>
<tr>
<td>equations</td>
<td>Sondre A</td>
<td>23.017</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Sondre B</td>
<td>139.32</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Sisimiut</td>
<td>39.009</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### 9. Graphical presentation

The 929 data points used in analyses of the soil and air temperatures are represented in Fig. 5. The cross plot shows a moderate—poor linear relationship between the...
... during periods without snow cover. It is important to note that these results are for the generalized site condition of soil temperature response to changes in air temperature when snow cover is absent or present. As discussed earlier (Fig. 5), other soil and snow properties influence the scale of soil temperature responses. These factors include seasonal changes in heat capacity of frozen versus unfrozen soil, thermal conductivity and diffusivity, soil moisture, solar radiation, surface albedo and the timing and duration of seasonal snow cover (Ling and Zhang, 2003), and snow thickness and density (Ge and Gong, 2009; Lawrence and Slater, 2010; Ling and Zhang, 2003).

10. Summary

Soil and air temperature data recorded at four Greenland sites were analyzed using multiple regression analyses that incorporated dummy predictor variables. The MLRA considered the temperature response of shallow soil to changes in air temperatures under conditions of snow cover presence or absence. The MLR equation was reduced to two separate linear equations, based on selection of the dummy predictor variable \( z \), which represented site conditions having an absence \( (z = 0) \) or presence \( (z = 1) \) of snow cover. Use of the predictor variable enabled hypothesis testing in relation to the regression equations, and assessment of differences in slope, intercept, and coincidence between the two reduced linear equations.

The results obtained for the four sites indicated that under conditions of snow cover the temperature of shallow soils changed by \( 0.12 \pm 0.23 \) °C in response to a corresponding \( 1 \) °C change in air temperature, while in the absence of snow cover the soil temperature

---

**Fig. 6.** Cross-plot of soil and air temperature data recorded under varying snow cover conditions \( (z = 0 \text{ and } 1) \) at each of the four Greenland study sites: a) Ilulissat, b) Sondre A, c) Sondre B, and d) Sisimiut. The regression lines shown were calculated by substituting the dummy variables \( z = 0 \) or \( z = 1 \) into the MLRA model. The dashed regression line represents the air and soil temperatures recorded in the presence of snow cover \( (z = 1) \), while that in the absence of snow cover \( (z = 0) \) is represented by the solid line. The regression coefficients \( b_0 \) and \( b_1 \) are shown in Table 2, and the MLRA regression coefficients \( b_0 \), \( b_1 \), \( b_2 \), and \( b_3 \) are shown in Table 1. The estimated model statistics are provided in Tables 1–3.
changed by 0.22—0.44 °C. Analysis of the combined data for all four sites showed that in response to a 1 °C change in air temperature there was a shallow soil temperature change of 0.18 °C in the presence of snow cover, and a change of 0.38 °C in the absence of snow cover. Relative to ordinary linear regression analysis of the datasets, the MLRA yielded stronger coefficients of multiple determination, and less variation in the estimated regression residuals.

The increase in the calculated $r^2$ between the MLRA and the individual regression equations determined in the presence and absence of snow cover is attributable to the addition of the product variable ‘xz’ in Eq. (4). Increasing the number of variables in a multiple regression model almost invariably increases the $r^2$.

Acknowledgments

The author appreciates and acknowledges the use of data from the Danish and Greenlandic Geological Survey (GEUS), and the cooperative agreement between GEUS and the National Snow and Ice Data Center (NSIDC) data repository, which enabled access to the data. The author also expresses appreciation to Peter Veneman, Baoshan Xing, and Stephen Simkins for their review of very early stages of this research.

References


National Snow and Ice Data Center (NSIDC). World Data Center for Glaciology Data Repository. http://nsidc.org/.


