UNIVERSITY OF CALGARY

Four Applied Methods for Spatial Visualization in Snow Avalanche Forecasting

by

Cora Shea

A DISSERTATION
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOSCIENCE

CALGARY, ALBERTA
October, 2011

Abstract

This thesis presents four applied methods for seasonal snow observation with respect to avalanches. Previous avalanche-related spatial variation and scale studies have shown a clear need for observation and methods to focus on the scale of interest to human triggering. These methods have the common goal to reveal spatial variation of interest to avalanche formation and human triggering in an efficient, accessible manner.

The four methods are: (1) A minimally destructive slope-scale sampling method, (2) A method to relate Google Earth terrain images to surface hoar formation in sparse trees, (3) A method of accessibly presenting complex GIS warming model data over real terrain, and (4) A method of measuring heat in the snowpack using a thermal imager. Despite their common goal of spatial visualization, each new method draws on a different subset of background literature and employs very different methods in development and use. Thus, each method is presented as a self-contained paper with independent results. Of note, these methods have all subsequently received active use, and conclusions from such use are discussed at the end of the thesis.

For Bruce, who always believed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

In modern times and Western countries, avalanches are comparable in their average yearly fatalities to other weather hazards. In the United States, avalanches have caused a moving five-year average of fatalities hovering at 25-35 per year between 1995 to 2010 (Colorado Avalanche Information Center, 2010), with other reference hazards including floods (71 fatalities), lightning (39 fatalities) and tornadoes (56 fatalities), on a 10-year average (National Weather Service, 2010).

Some transportation corridors, including rail and highways, face danger from avalanches. *Avalanche forecasting* (discussed further in Section 1.2.1) can be used to protect highways with substantial commerce utilizing them and requiring them to remain open through the winter. Closures are costly to commerce, and open highways during times of avalanche hazard threaten the safety of vehicles on the road. Highway operations use explosives to trigger what are often smaller, situation-controlled avalanches for maintenance in a similar – yet much more powerful and safer – way than would a person travelling on the snow surface. The successful protection of highways by these avalanche forecasting and control agencies is at least partly responsible for the relative dominance of recreational fatalities – fewer people are dying on highways (McClung and Schaerer, 2006, pg 18).

To some extent, avalanche hazard can be predicted. Avalanche forecasting programs operate for this purpose around the world. The stability of seasonal snow is directly influenced by the weather, which is itself a complex phenomenon. The weather then leads to secondary effects such as loading stress by new snow, formation of *weak layers* (discussed further in Section 1.2.2) at the surface such as surface hoar or surface crusts, and internal metamorphism leading to weak layers due to cool temperatures and a shallow snowpack. The complexity of these interactions means that, at present, no model or method exists to completely predict the avalanche phenomenon, especially for an individual slope (Schweizer et al., 2008).

The slope *spatial scale* (discussed further in Section 1.2.3) is of particular interest to skiers, snowmobilers, and other recreationalists. Decisions about the stability of a particular slope can be necessary many times during a day of travel. This happens via *recreational decision making*, which is discussed alongside forecasting in Section 1.2.1. Both organized avalanche forecasting and recreational decision making depend heavily on data collection to assess the ever-changing and complex avalanche phenomenon.

Hence, studying and forecasting avalanches, for all activities from research to assessing the snowpack stability for a day of skiing, are strongly rooted in observation. The snow itself is observed and tested to assess its current condition, and weather is observed to help forecast future snow conditions and assess certain immediate factors such as loading and warming. Observers watch for clues of instability, dig pits whereby the vertical layering and horizontal variation of snow is exposed for study, and watch for differences in observational data as time, weather, and terrain vary.

Unlike many other fields, it is common for professionals and recreationalists alike to take time from their day and figuratively stick their noses in the snow to observe, record, test, and analyze it, an occupation in other fields reserved solely for researchers. Observers who work or play for many seasons in avalanche terrain often have logbooks filled – formally or informally – with observations including weather, evidence of avalanches, tests that have been performed on the snow, internal snow layering and crystal types, and internal snow temperatures. Therefore, the methods and tools by which we are able to observe the snowcover directly influence our study of avalanches and hence our safety.

More recently, deserved attention has been given to the idea of *spatial variation*, discussed alongside weak layer formation in Section 1.2.2, whereby the seasonal snowcover forms and metamorphoses in a highly spatially variable manner with respect to avalanche formation (Schweizer et al., 2008). Understanding this variation – that a single slope can contain many different trigger points for skiers or snowmobilers, as one instance of this concept – helps ski guides and recreationalists make informed decisions in the backcountry. Many other observations of interest regarding spatial variation – and thus directly or indirectly trigger points over terrain – remain as-yet undiscovered. To complicate matters, these would-be observed qualities are different at different scales (Haegeli, 2004). For example, the way a weak layer forms on an open slope will not be true across a larger scale that also includes terrain with sparse trees.

1.1 Objectives

To address this complex and newly emerging conceptual structure, methods to observe variation in the snow across the scales that matter for human triggering for either recreation, access, or explosive control are desirable. Furthermore, due to the involvement of educators, forecasters, and recreationalists in this exercise of snow observations, methods are needed not only for researchers but also for these additional groups.

Therefore, the four methods presented in this thesis have the primary objective to be spatially-oriented, practical methods of visualization for conditions useful in avalanche forecasting. Second, these methods strive to be scientific and technological with respect to using statistics, modelling, and use of both terrain and thermal imagery to provide a useful perspective on conditions, while still fulfilling the primary goal. These methods are applied, and contribute to the field of avalanche science in their practicality rather than theoretical aspects.

1.2 Further background

For the reader unfamiliar with the italicized terminology above, this section provides greater depth of background.

1.2.1 Avalanche forecasting and recreational decision making

Currently, the action which increases avalanche-related safety for any user is correct decision making. Any user – guide or recreationalist – after collecting as much useful information about the conditions as possible, must actually put that information into practice and act upon it.

To provide an example, at some point the user will be in the field, facing a decision: Ski the slope, ski another slope, or go home. Highmark a slope, or just go for a conservative tour. Choose the open slope, the gully, or the trees. For a user to be set up for success in these examples is, in fact, a long process beginning with understanding what information may be needed (i.e. research), proceeding through educating that user on how to use the information and how to sort through the extraneous noise of peer pressure, irrelevant information, and the temptation of good skiing or sledding (i.e. education), actually obtaining, processing, and predicting useful information (i.e. forecasting), and finally using some repeatable and successful method to make the decision itself. The flow of information for recreational decision making may be seen in Figure 1.1.

A professional avalanche forecasting program has its advantages because it can (a) provide a central place to collect relevant information about the snow conditions, (b) recruit experienced forecasters to process this information, and (c) provide complete focus to avalanche hazard, thus being particularly attuned to the patterns and specific traits of each forecast area. This type of setup lends its experience to all possible decisions for the end user, but must remain general if the forecasters are not making the actual individual decisions, as is the case for a regional forecaster. A combined effort of forecasting, education, and research can help a user know not to ski steep south-facing slopes in the spring during warm temperatures, for example, but ultimately the user still must decide whether a specific slope fits that criteria.

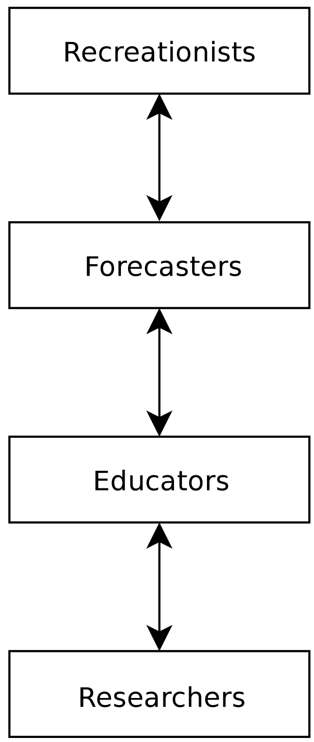


Figure 1.1: The bi-directional flow of information and technique within avalanche forecasting and decision making. Information flows both ways – that is, research should ultimately fulfill the needs of the end user. Information can also flow directly from any of these groups to the end user as well, though less often.

Avalanche forecasting is a process by which the current and short-term future avalanche hazard is assessed. Currently in North America, a forecast is normally produced once or twice per day, and the hazard or danger is rated on a five-point scale between low and extreme (Statham et al., 2010). This scale, even when produced as a general danger rating for a large area, can provide good information about the avalanche hazard. From 1996 to 2007, of the 74 avalanche fatalities in Canada for which the danger rating forecasted by a public avalanche bulletin is known, 83 percent of them occurred with a danger rating of Considerable or higher (Jamieson et al., 2010), which includes the ratings ranging from human triggering likely to human triggering certain (Statham et al., 2010).

However, due to the bulletin information being only a part of the decision process discussed above, avalanche forecasting programs often also have a vested interest in the issues of general safety and decision making for the end user. Education also has this interest for the same reasons, as does research. Also because of this, forecasting, education, and research all wish to treat the avalanche hazard in the same context as the end user. In other words, all of these audiences greatly desire to provide tools and information applicable and usable at the particular slopes a user will face, and about the particular triggers and layers on those slopes that the user should be concerned about. These are known, respectively, as the slope scale and the trigger scale, and they are discussed more in Section 1.2.3 below. General area forecasts greatly assist, but do not solve, these slope-specific decision problems. Finally, this common goal enables research to contribute new applied methods and observation techniques to any audience with an eventual positive effect on the end user travelling or working on or below snow covered slopes.

1.2.2 Snowpack processes and weak layer formation

The seasonal snowpack forms in layers due to variations in snow storms (Colbeck, 1991). These layers then metamorphose and may either disappear over time or persist as distinct stratigraphy within the snowpack. To produce an avalanche, the snowpack needs a cohesive *slab* which transmits stress to fail a *weak layer* buried below it (Heierli et al., 2008). This weak layer collapse is assisted and focused by the often stress-concentrating effects of the *bed surface* below the weak layer, which is left in place after the slab and weak layer have avalanched (McClung and Schaerer, 2006).

The variation of these three components across terrain creates distinct trigger points, or places on the slope where a buried weak layer is supportive of the snow above it but unable to support additional load such as a human or explosive. When triggered, the weak layer fails, the failure can propagate within the snowpack beyond the trigger point, resulting in a slab avalanche. Other types of avalanches exist, but it is this type – dry slab avalanches – that are of most concern, accounting for 95 percent of the Canadian avalanche fatalities between 1996 and 2007 (Jamieson et al., 2010).

As a slab may be detected through ski travel, hands-on testing, and observation, and many different snow types may serve as a bed surface, the majority of forecasting attention is often given to weak layers. This variation of these weak layers across terrain directly relate to the slope-scale decision making process discussed above. If a weak layer is highly variable across a large area, overall the probability of that weak layer being present within the snowcover of a particular slope is lower than if the weak layer was everywhere, but the difficulty to assess the danger of any particular slope is higher. To avoid the safe but extremely limiting decision to always treat slopes with buried weak layers as if they will fail (highway operations using this type of decision process would be forced to keep highways closed most of the winter, for example) it is of interest to know how these weak layers vary, both over terrain, and over time with regards to their strength.

The formation and metamorphism of seasonal snow weak layers is an extremely complex subject. In bulk (Bartelt and Lehning, 2002) or in very small quantities (Kaempfer and Plapp, 2009; Gravner and Griffeath, 2009) snow metamorphism can be modelled; however, there is still much at the microscale of ice crystal formation and change that is not understood (Libbrecht, 2005). Therefore, studies of weak layer formation and variation over time and terrain are primarily observational and empirical, emphasizing the continuing need for quality applied observation methods.

1.2.3 Spatial scale

The public avalanche forecasting bulletin areas discussed above in Section 1.2.1 are on the order of 100 to 30,000 square kilometers (Bakermans et al., 2010). As anyone who has seen storm clouds stop and precipitate on one side of a mountain ridge can attest, conditions can vary wildly within such a large area. Furthermore, the recreational decision process also discussed in Section 1.2.1 operates on only a handful of slopes within those larger areas.

This concept of different useful scales is a dissertation topic in itself, and an excellent reference is therefore Haegeli (2004). As a definition, the term *scale* refers to the *characteristic length or time of a process, measurement or model* (Blöschl and Sivapalan, 1995).

Therefore, for example, by the term *slope-scale* in avalanche research, one refers to the decisions and information pertaining to a particular slope and its avalanche possibility. The term *trigger scale* refers to individual potential trigger points on a slope, but not the whole slope. *Basin scale* refers to a basin with many slopes and varying aspects and elevations, and *layer scale* refers to a scale that captures observations and processes pertaining to buried individual snow layers. The key to this concept is that things generally true at one scale may not be true at a larger or smaller scale. For example, if the danger rating is Considerable at the large, multi-square-kilometer *forecast area scale*, that rating may or may not be accurate on a single slope, even within that forecast area (Bakermans et al., 2010).

This commands attention, and even a particular level of respect, from the audiences of Figure 1.1, as the scale differences tend to create great difficulties for the end user. For example, if a user reads a forecast of Considerable danger (which may indeed be true over the majority of an area) but encounters and safely skis a slope within that area for which the avalanche danger is low, without respect for the scale differences he or she may reduce their confidence in the bulletin and ski more aggressively in the future. Therefore, observational methods which operate at similar scales to which their intended audiences operate are desirable.

1.3 Contributions

The four methods presented in this thesis are:

1. An efficient and minimally-destructive method of spatial point measurement on snow (Chapter 2, Shea and Jamieson 2010d)

2. Use of Google Earth terrain imagery to estimate the continuity of surface hoar formation (Chapter 3, Shea and Jamieson 2010c)

3. A method of presenting GIS model data without requiring a GIS, for operational avalanche forecasting (Chapter 4, Shea and Jamieson 2010a)

4. Use of thermal photography to visualize temperature-related snow processes (Chapter 5, Shea and Jamieson 2010b)

Contributions one, two, and four have appeared as manuscripts in peer-reviewed publications; contribution three appeared in conference proceedings.

As the primary author on contribution one, I was responsible for all simulation, mathematics, and the bulk of the writing.

As the primary author on contribution two, I was responsible for most of the observations, all model development, and the bulk of the writing.

As the primary author on contribution three, I was responsible for the user interface design including background research, the GIS code development, the mathematics adaptation from the model used, and the bulk of the writing.

As the primary author on contribution four, I was responsible for the development of the method in the field, background research, the limited mathematics present in the paper, analysis of the data, and the bulk of the writing.

Each contribution, its scale, and its audience, are discussed below. Following that, the next four chapters contain the four contributions above, followed further by conclusions and references for all chapters combined.

1.3.1 Spatial Sampling

The presented sampling method efficiently allows many observations to be made with minimal destruction to the snow. This method works on the slope scale, and allows observations to be made on the scale of individual trigger points, and thereby allowing one to observe the variation of those points across a slope. Its intended use is research, and the presentation of it in this thesis includes a comparison to other sampling methods with respect to the semivariogram.

1.3.2 Google Earth Photography and Surface Hoar

The presented surface hoar size prediction method for sparse forests utilizes easily accessible terrain imagery. This method also works on the slope scale, showing surface hoar variation across a single slope. Based on Google Earth photography, the method can be visually used by recreationalists during route planning. The concept of the effect of sky view on crystal size and hence potential trigger points can also be applied ad-hoc to recreational travel in areas of sparse forest.

1.3.3 Operational GIS Model Development

The presented warming model displays the results of an existing warming model over recognizable, real terrain. This method works on the basin scale, showing the variation across a day of ski touring or snowmobiling, for example. The method is intended for forecasters and recreationalists. It presents modelled Geographic Information System (GIS) data over the web, that is, without requiring a GIS. The data presentation also allows *hypothesis testing* for determining warming based on a range of possible future conditions.

1.3.4 Snow Surface Thermography

The presented temperature measurement method shows the surface temperature of snow – whether on a surface exposed by digging a pit wall or on the natural surface – by using a handheld thermal imager. The method operates on the layer and trigger point scale, where variation internal to the snowpack may be examined via pit wall exposure, or spatial variation across a small portion of terrain near buried or exposed terrain features may be examined. The method is intended for researchers and educators, but may, in time, find application to forecasting programs.

Chapter 2: Spatial Sampling

To better observe snow spatial variation on a small scale and over time, more efficient and less destructive sampling methods are needed. This paper, entitled *Star: An efficient snow point-sampling method* appeared in the journal Annals of Glaciology as Shea and Jamieson (2010d).

I am grateful for the permission of the International Glaciological Society to reprint the paper from Annals of Glaciology in its entirety. The manuscript here differs with small modifications from the original version.

For more information on use of this sampling method in the field, see Section 6.1 in the concluding chapter.

2.1 Abstract

The changeable, variable, and fragile nature of snow creates unique sampling challenges. In this paper, we present Star: an efficient, field-usable sampling method for use in point-sampling spatial studies. This paper validates the accuracy of the Star method via a comparative Monte Carlo simulation using 1024 detailed samples of elevation data. As spatial snow studies generally want to find spatial continuity in layers and other properties, we used variogram ranges to compare the ability of four sampling methods to accurately reveal such spatial correlation. The other three methods compared to Star represent gridded, gridded random, and pure random methods, whereas Star can be called a linear random method. The simulation showed Star's reproduction of spatial range to be comparable to both gridded and gridded random methods. From this comparative process we introduce a new measure of appropriateness for sampling methods: the correct convergence on a variogram model, which we

call correct spatial correlation detection. This directly measures how many sampled areas become correctly classified with either spatially correlated or non-correlated variance for a given variogram model fit. In this measure, Star performed equivalently to the other methods, and in correct convergence it performed equally to pure random sampling.

2.2 Introduction

Snow sampling methods have a multitude of applications and challenges. As snow properties changes over days or even hours, and sampling can be destructive to the snow properties being measured, sampling methods must be as efficient as possible. Time spent laying out a sampling grid or moving from point to point can affect the number of observations that can be made as well as possibly affect the snow properties if not done properly.

This paper compares point-sampling methods for surveys such as depth, penetration resistance, surface conditions, and so on. These can also be thought of as minimal-support observations (Blöschl and Sivapalan, 1995), and they commonly enable the observer to make many more observations in a day than large-support tests such as the Rutchblock, for example.

Performing such point observation surveys usually occurs with the intent to spatially describe snow qualities of an area or spatial processes affecting the area (Schweizer et al., 2008). Examples include finding the spatial extent of snow layers (Kronholm, 2004) for use in avalanche forecasting or obtaining good spatial visualizations of water storage in the snowpack (Cline et al., 2001).

2.3 Related Work

A few previous evaluations and comparisons of sampling methods for spatially measuring avalanche-related conditions exist. Such spatial sampling methods typically focus on many points within small areas on the order of a slope or basin, and are subject to strong time and access constraints. They are distinct from sampling networks such as distributed weather stations (Gray and Male, 1981) in that they aim to provide a spatially visualized field of data within a small area rather than a representative point. They are also distinct from more general field and landscape surveys in that they do not aim to adapt the method to follow a particular trait of interest – such as elevation in the case of land surveys – and instead serve as a discovery method rather than a survey one.

Currently, the most common way to compare and defend such spatial sampling methods are lag-bin distributions, as given in Kronholm (2004) and Bellaire and Schweizer (2008). These distributions can be thought of as histograms for how many point pairs in a sample exist in a given lag bin for possible variogram-type analysis.

Variograms measure the spatial correlation at different distances, or lag bins, within a field, and therefore they measure the continuity of similar measurements well. This can be useful for tracing the two-dimensional extent of traits such as snow layers (Kronholm, 2004) or wind and terrain effect (Deems et al., 2006). The lag-bin type of evaluation, then, comes from the understanding that the more points in a lag bin, the better that sampling method will capture spatial correlation at that lag.

A much more in-depth analysis, performed by Kronholm and Birkeland (2007), analyzed different sampling methods and how they reproduced a known range, sill, and nugget for a spherical variogram model generated from a small subset of generated random fields. This type of analysis directly analyzed histograms of error, which addressed the accuracy question much more thoroughly than lag-bin analysis.

Development of new spatial sampling methods of this type seems limited. Most current methods derive from grid-type structures, with their grid spacing varying to capture information on different spatial scales. These include the LH grids (Birkeland et al., 2004), the MT grids (Birkeland et al., 2004), and the Swiss grid (Kronholm, 2004; Kronholm et al., 2004). Some studies (Bellaire and Schweizer, 2008); (Cline et al., 2001) utilize a sampling method with random locations placed within and organized by an overall grid, but provide minimal analysis of the method.

Gridded methods can be easier to divide into an orderly day of work, but they can require extensive layout. Pure uniform random distributions are near impossible to divide up and sample logically without destroying the area in the process. However, they continue to be highly desirable due to their accuracy (Kronholm and Birkeland, 2007). We feel this leaves a gap, which then defined our objectives – design a sampling method with:

- Efficient and minimally destructive layout and sampling
- Similar spatial modeling accuracy when compared to other current methods

We felt that such a method would have an orderly implementation of random points, yet still without needing to lay out a grid; thereby obtaining its accuracy via randomness, and its efficiency via some imposed order.

2.4 Methods and Data

Using the gstat package (Pebesma and Wesseling, 1998) in the R Project for Statistical Computing (R Development Core Team, 2006), we performed a Monte Carlo simulation to compare the fitted variogram models for 1,024 real datasets to fitted variogram models of samples of that data. We describe the details of the datasets, variograms, sampling methods, and inclusion of randomness in the following subsections.

2.4.1 Dataset

We used naturally occurring datasets. Digital Elevation Model (DEM) data, having both spatial correlation and occasional fractal dimension, shows the same general qualities of snow cover (Deems et al., 2006).

We used four 1:50,000 parcels of DEM data from Geobase.ca (Government of Canada, 2007): 93b, 93e, 93f, and 93h. Each parcel contains 16 grids with 1201 x 1201 elevation points. We trimmed each 1201 x 1201 grid down to 1000 x 1000, and then due to the $O(n^2)$ computational demands of the variogram, we additionally split that into sixteen 250 x 250 point grids.

This gave 1,024 grids of data, each with 62,500 points. Each can be thought to model a 25 x 25 m grid with possible samples every 10 cm. When sampling a point between two known values, we used the closest known value. And, as discussed below in the subsection entitled *Variogram*, we operated on the residuals of the elevation values left after removing linear trends from each grid.

This dataset enabled us to examine enough data points to assess each sampling method over a variety of spatially correlated data, including a wide variance of ranges and many fractal and linear variograms (e.g. Deems et al., 2006).

Generally, normal distributions do not model snow data well, and indeed log-normal distributions do only somewhat better (Kronholm, 2004). Hence, although terrain data do not necessarily adequately represent all snow variable distributions, they do represent a primary forcing variable in all aspect-dependent snow data such as weak layer formation and therefore can be argued to provide an advantage over normally distributed distributions.

While investigating a subset of the data for lognormal trends, we compared the trends in the real data with the Swiss grid samples (described in the section below) of those data. We found that although between eight and eighteen percent of the Swiss grid samples showed a significance at $\rho < 0.05$ fit for a lognormal distribution, (dependent on fit method, KS/Lillefors, Anderson-Darling, and Cramer-von-Mises tests were used via the Nortest R package (Gross, 2006)), *none* of the corresponding real data sets showed $\rho < 0.05$ fit to a lognormal model. Thus, we did not include normal or lognormal fit as a basis for choosing the dataset.

2.4.2 Sampling

We compared gridded, gridded random, linear random, and pure random sampling methods. Figure 2.1 shows visual layouts of the four sample methods compared in this paper.

Due to the extensive analysis provided by Kronholm (2004), we used the Swiss grid as the representative grid method. For the gridded random method we chose the L-Grid (Bellaire and Schweizer, 2008; Cline et al., 2001) which enjoys relatively wide use. A simple uniform random sample distribution served as the pure random sample.

To approximately equalize its number of points and maintain its original spacing, we added an additional outer layer (16 points) to the Swiss grid. The 16 additional points give it the most sample points (129) of any method. Random sampling consisted of 125 points, Star consisted of 126 points (21 points per line over six transects), and L-Grid consisted of 125 points (five points per grid over 25 equal grids).

Our method, Star, fills the linear random niche as we could find no others. It consists of six transects which always divide up the area the same way, as shown in

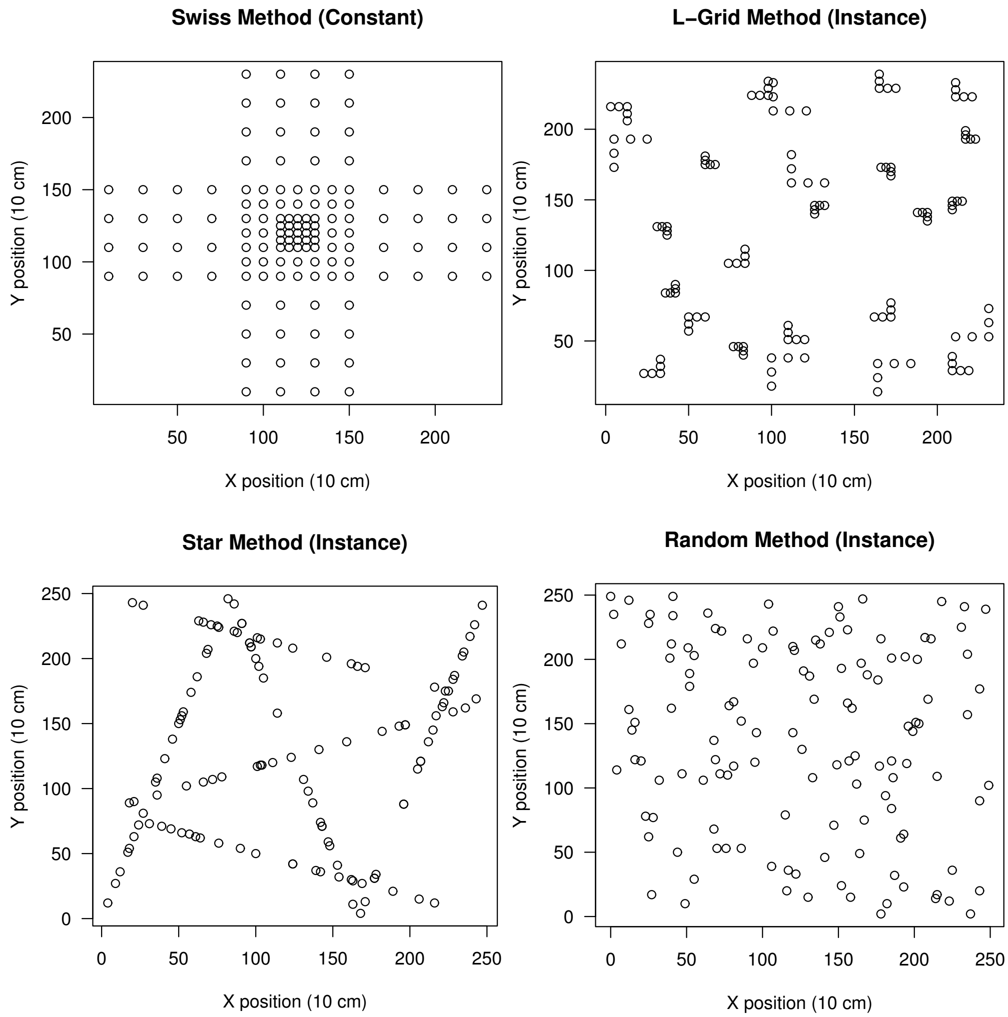


Figure 2.1: The four sampling methods used. Note that the three methods that contain randomness – Random, Star, and L-Grid – varied with every instance. Each grid consists of 250 x 250 points, which we show here as 10 cm spacing.

Figure 2.1. Each transect has the same number of sample locations. Only the spacing between points varies randomly from sample to sample, transect to transect.

After one winter of use, we found the effective minimum spacing varies by the type of sampling being performed (Shea and Jamieson, 2009). For point crystal size measurements, our smallest usable spacing was 10 cm. For measurements that require larger support or equipment, the smallest spacing may be larger to prevent overlap of measurement effects.

To use Star, one begins at the top of observer's left and traverses across the area to one third of the way down the opposite (right) side, sampling at uniform random intervals along the way. This forms transect one. From there, the user repeats the random sampling process as he turns and traverses again to a point two-thirds down the left side, and turns again – still sampling – to traverse to the lower right corner. This forms transects two and three, respectively. The same 'by thirds' spacing structures transects four, five and six which are essentially a 90 degree counter-clockwise rotation of the first three transects over the same area.

Star's efficiency comes from a number of qualities. Most notably, the user always travels a known and reasonable distance including a small number of turns to sample an area. Here, on the 25 m squares, the user would traverse about 184 m total with only six turns, including the traverse on the bottom from the end of transect three to the start of transect four. To sample the Swiss grid in the most efficient way would take 191 m of travel with 18 turns.

As for the L-grid, since the positions of the L's in the method vary it does not have a constant travel distance. As a lower bound – when all of the L's are as small as possible and ideally stacked linearly with respect to one another – the user would travel just under 160 m. However, the user would make 75 turns: one for each L, one to head toward the next L, and one to align oneself along the new L. At the upper distance bound – using the largest L size and with L's placed in opposite corners of their grids – the user would travel over 240 m, again with 75 turns.

As the random sampling method creates very difficult problems for finding the shortest distance to travel and minimal required turning, we do not present its efficiency here. However, the high number of turns required to perform L-grid should give the reader an intuitive sense for why pure random sampling methods remain essentially unusable in the field.

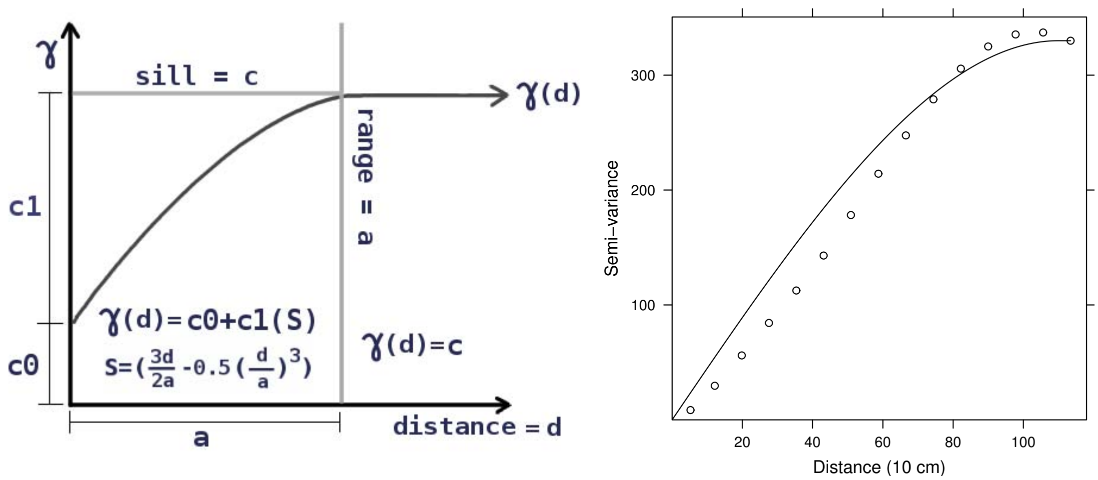


Figure 2.2: A diagram showing (on left) how the equations of the spherical variogram model determine range, and (on the right) an example variogram and fit spherical model. The spherical model shows a reasonable – but non-perfect – fit to the variogram.

Thus, we feel the structure of Star adequately fulfills our objective of efficiency and minimal destructiveness as the user can stay on one line at a time via skis or foot travel; the remainder of the paper focuses on the comparative accuracy of the method.

2.4.3 Randomness

For the three methods containing randomness – Random, L-Grid, and Star – the simulation varied that randomness for *every* sample layout. For Star, the spacing along each sampling transect varied. For Random, the x and y coordinates for each point varied. For L-Grid, the axis point of each L within each grid, the spacing between points, and the orientation of the L varied. So, the layouts in Figure 2.1 are but one instance of many random variations used.

When formative process varies, so should the measurement spacing (Blöschl and Sivapalan, 1995). This variation in spacing ‘catches’ and identifies correlation by being itself an independent function of what it measures. Thus the attractiveness of random methods comes from process-independent variance in the sample method. In other words, randomness does not oscillate in step with any measured data (other than random data), and it may help discover the operation of unknown formative process scales.

However, this means that methods which use randomness must maintain it. Even if the locations of a sample are decided randomly once, one cannot easily know whether that single instance lies within the extreme end of a random distribution, or the more desirable median of a random distribution. Thus, if we selected a single instance of Star, L-Grid, or (pure) Random methods, it could introduce bias or skew and our results would be very different.

2.4.4 Variogram

For each real data grid and its four samples, we produced omnidirectional variograms by Cressie's (1993) robust method. Each semi-variance, calculated over 15 bins with respect to residuals left after linear trend removal, then provided a basis for fitting a variogram model. We removed all linear trends regardless of significance because we prioritized removing all linear anisotropy over retaining the original data values.

The general variogram, shown in Equation 2.1 (O'Sullivan and Unwin, 2003), simply finds the squared difference between all n residual value pairs (z_i, z_j) within a lag bin of width 2Δ and given lag distance $d \pm \Delta$ wide from each other:

$$\hat{\gamma}(d) = \frac{1}{2n} \sum_{i,j \in d \pm \Delta} (z_i - z_j)^2 \qquad (2.1)$$

Cressie found that although this general form has no bias, it can have skew because the squared factor amplifies large outliers. His robust model calculates the

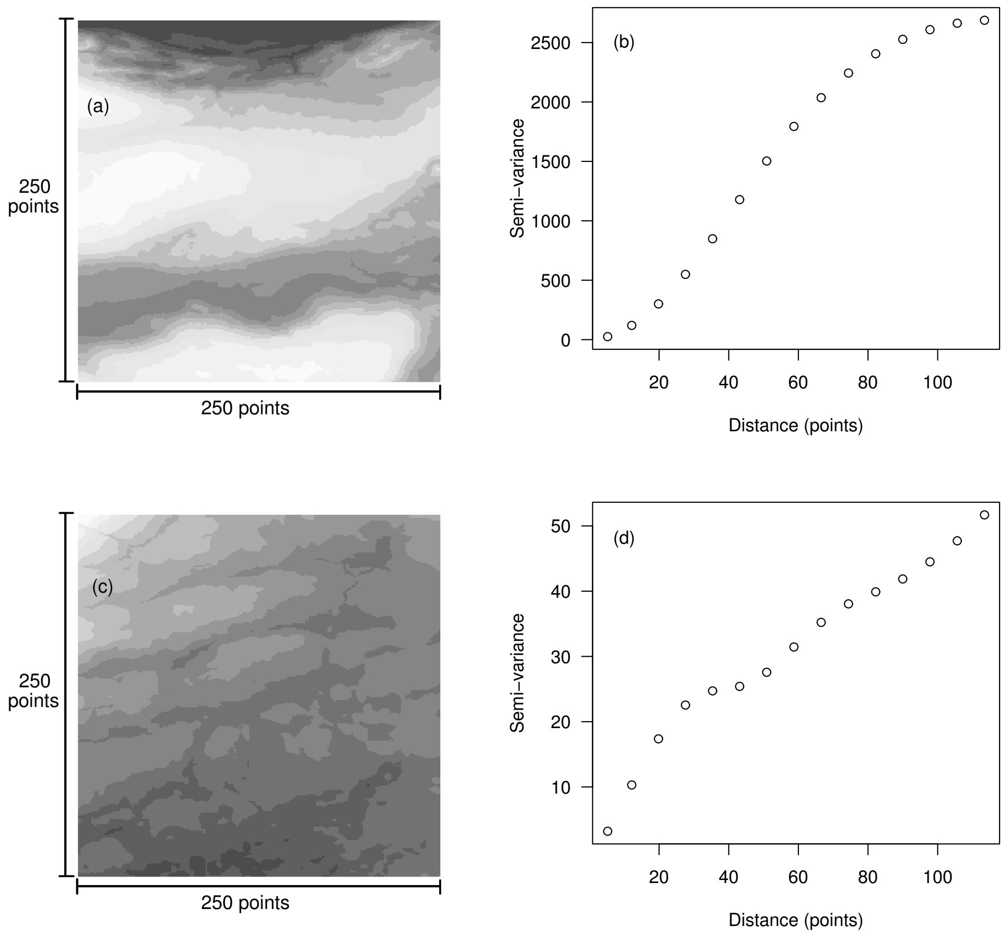


Figure 2.3: Two example datasets, both from Geobase grid 093b04, presented in greyscale varying with the point values. Each has the semi-variance rise of their corresponding variograms to the right. The upper variogram in (b) should be more properly fit with a Gaussian model rather than a spherical one. Compare to the lower variogram plot in (d), which presents a complex fractal character of multiple ranges, or 'spatial correlation within spatial correlation'.

variogram based on transformed differences of $|z_i - z_j|^{\frac{1}{2}}$ (Cressie, 1993) with a numerical denominator to account for the bias this introduces, as shown in Equation 2.2:

$$2\hat{\gamma}(d) = \frac{\{\frac{1}{n}\sum_{i.j\in d\pm\Delta}|z_i - z_j|^{\frac{1}{2}}\}^4}{0.457 + \frac{0.494}{n}} \tag{2.2}$$

After calculating the variogram, we fit a standard spherical model γ as shown in Equation 2.3. The model has range a, sill c, and nugget c_0 and gets defined up to the range $d = a$. After that, one uses the linear function $\gamma(d) = c$ where $c = c_1 + c_0$. This means that c_1 represents the y-axis distance over which the semi-variance rises from nugget c_0 to sill c (O'Sullivan and Unwin, 2003):

$$\gamma(d) = c_0 + c_1[\frac{3d}{2a} - 0.5(\frac{d}{a})^3] \tag{2.3}$$

We present the model and its use on a single real data variogram in Figure 2.2. One may see that the variogram can serve as a tool for relating to physical process scales, as the range demonstrates the physical extent at which measurements cease being similar to one another.

Due to its current prevalence in snow literature as a measure for evaluating data as well as sampling methods (Kronholm, 2004; Bellaire and Schweizer, 2008), we selected the omnidirectional variogram and a spherical model. We utilized least squares to fit the spherical model form to each variogram.

Use of these methods allows us to compare Star in relation to Swiss and Random which have been previously compared (Kronholm and Birkeland, 2007). However, such a setup – and the variogram in particular – has many limitations. Later, in the section entitled *Common Non-Convergence*, we briefly discuss the limitations of the spherical model and omnidirectional variogram.

As previous work has already been done on the variances of all spherical semivariogram model attributes – range a, sill c, and nugget c_0 – across different sampling methods excepting Star (Kronholm and Birkeland, 2007), we chose to simply present the error in range a to comparatively demonstrate Star's fitness as a sampling method. The range has been of interest in snow sampling for purposes of finding layer extent

(Kronholm, 2004), but it is of course not an exhaustive measure of fitness.

One may note that in Kronholm and Birkeland (2007), those sampling methods which performed well in range comparisons also performed well in sill comparisons. Here, after removing linear trends individually for each data set, the remaining sill values complicate comparison across data sets, and thus we refer to the relation present in that work. Then, the nugget can be thought of as a measure of how much small-scale process detail a method can capture. It may also be interesting for additional comparison in the future, but here we focus on the larger perspective of spatial correlation detection.

2.5 Results

When we completed the Monte Carlo simulation, we had four possible cases of how each sample's variogram model could compare with that of the original real data's model. With a least squares fit method for the spherical model in Equation 2.3, the real data model could *converge* on a spatial correlation range, the sample data could *converge* on a spatial correlation range, and the two did not necessarily happen together.

Non-convergences imply that the variogram would be better served by a linear model, a fractal model, or any other number of other possibilities. Fractal variograms rise up to a leveling-off point (pseudo-sill) before rising again to another leveling-off point. We show an example from our dataset in Figure 2.3.

These fractal variograms imply spatial correlation within spatial correlation, and although they can be interesting and useful (Deems et al., 2006) they often reduce to a non-converged model when spherical models attempt to fit to them. This comes from the spherical model's inability to fit multiple ranges and sills and thus reducing, by least squares, to fitting none of them.

Even more visually near-spherical variograms, such as the semi-variance rise of the Gaussian-type variogram shown in Figure 2.3(c), can reduce to a linear, non-converged fit. The fitting mechanism - in this case least squares - must attempt to fit the lesser rising slope at the low distances and thus may undershoot and miss the sill when it does occur at higher distances.

As the spherical model must curve eventually, these non-convergences presented with very large ranges - often several orders of magnitude larger than the data extent. Thus, unless otherwise noted, we defined a *converged* spherical model as one having a range of 500 points or fewer. This definition includes spatial correlation within the 25 m area and twice that width in a hypothetical prediction beyond it.

This hypothetical extension, though not statistically significant, appears in use in other sources in practice (Kronholm, 2004), and thus for comparison consistency we include those extents in our analysis here. Also, where possible, we also present results for different definitions of convergence at less than 500 points.

This gave us four possible categories, each with some unique subset of the 1,024 tests, as not all the real data converged on a good fit to the spherical model:

1. Common Non-Convergence (CNC): Where neither real data nor sample data variograms converged on a spherical model.

2. False Convergence (FC): Where the sample data variogram converged on a spherical model but no spherical model fit the real data variogram.

3. False Non-Convergence (FNC): Where the real data variogram fit a spherical model, but the sample data variogram did not converge on a spherical model.

4. Common Convergence (CC): Where both the real data and sample data variograms converged on a spherical model, though not necessarily the same one.

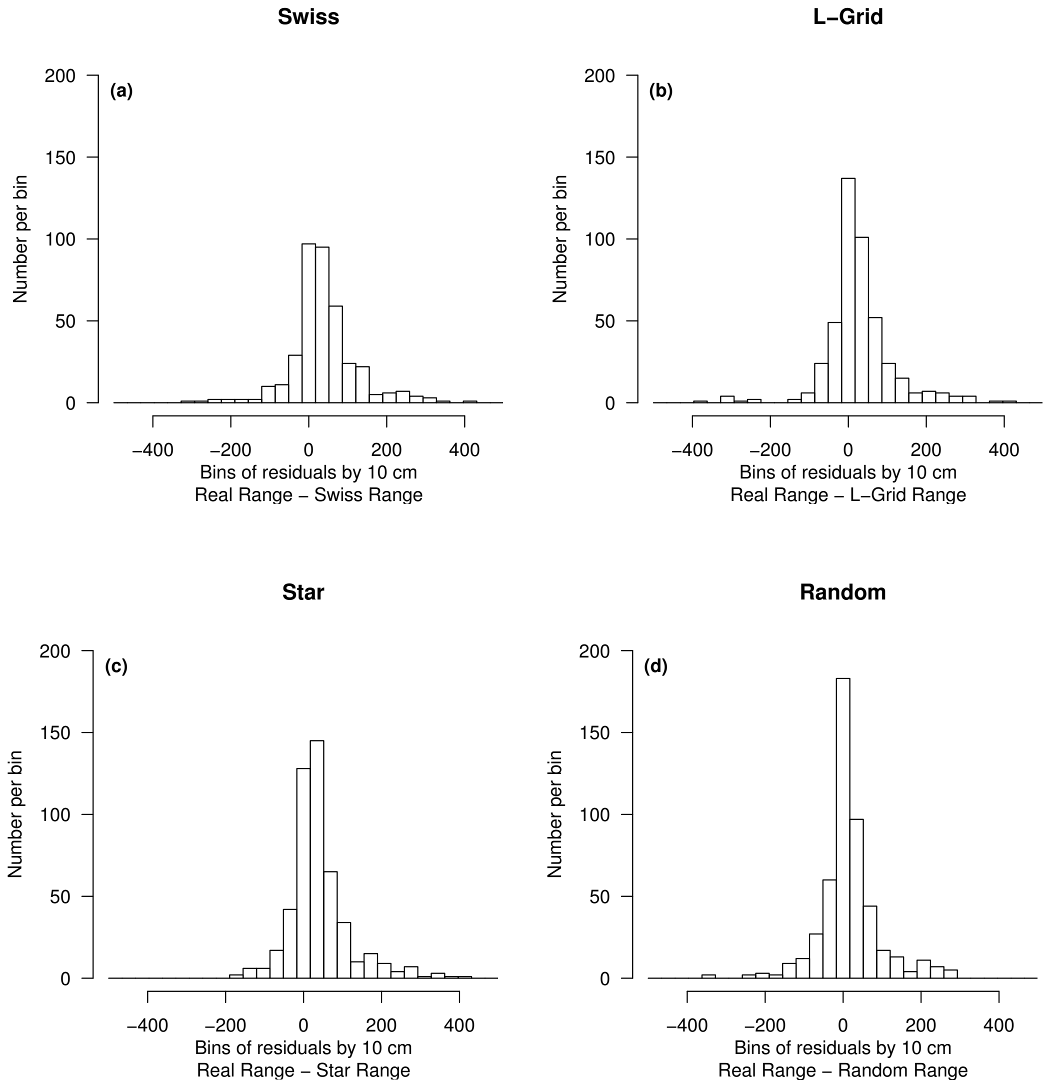


Figure 2.4: Histogram distributions of range errors for n correct Common Convergences by sampling method. (a) Swiss: n = 384, Median = 31.6 points (3.1 m), Standard deviation = 86.5 points (8.7 m). (b) L-Grid: n = 447, Median = 16.3 points (1.6 m), Standard deviation = 85.8 points (8.6 m). (c) Star: n = 496, Median = 25.2 points (2.5 m), Standard deviation = 77.3 points (7.7 m). (d) Random: n = 498, Median = 8.2 points (0.8 m), Standard deviation 77.9 points (7.8 m).

The main validation of accuracy lies in the common convergences and common non-convergences, but we examine the particulars of each category in turn. How often a sampling method ends up in the right category can be as much or more important than how well it performs in any one category.

Without knowing how often a sampling method correctly or incorrectly detects the existence of spatial structure in the underlying data, one cannot put faith in the structures that the sample does detect. Because, as this paper shows, some instances of 'detected' structure may not have actually been present at all. And conversely, some structures that should have been found went unnoticed. We consider spatial correlation to be a measure of spatial structure, and we consider spherical model fit to be a measure of detection, but false spatial structure detection – whether by false presence, or false absence – should be of concern by any measure.

2.5.1 Common convergence

Finding the total number of correct convergences involves a simple intersection. If *C(Real)* represents the set of real data sets that converged on a spherical model fit, and *C(Sample)* represents the set of sample data sets that converged on a spherical model fit, then the intersection of the two gives the Common Convergence (CC):

$$CC = C(Real) \cap C(Sample) \tag{2.4}$$

Though only part of the picture, the number of times a sampling method shows spatial correlation correctly can be a measure of its performance. The correct convergence of a sample measures how often it detects spatial correlation when it actually exists in the real data.

Consider, for example, having an ideal sampling method which has a common convergence percentage of 100 percent, and a common non-convergence percentage of 100 percent. Then, every time our sample converged, we could know the real data actually demonstrated some kind of spatial structure - here, variance such that it fits a spherical model with a reasonable range. In reality, when samples have common convergence only 75 percent of the time or less on these elevation data, we know that at least 25 percent of our linear variograms *should* have been reasonably spherical – but which data compose that 25 percent is unknown.

When both real data and sample data fit a spherical model well, the next question became: how well? To answer this, we found the residuals in ranges for the common convergences. For example, if the real data presented with a range of $a_r = 200$, and the sample with a range of $a_s = 400$, although they both converge, the sample data range is not very accurate with an error of -200 points (-20 m).

Figure 2.4 shows histograms for the range differences for all common convergences with ranges less than 500 points (50 m). If the ranges of the sample set model variograms can be denoted as *a(Sample)* and the ranges of the real data set model variograms as *a(Real)*, then the histograms display:

$a(Real_i) - a(Sample_i), \forall (Real_i, Sample_i) \in CC$

for each sampling method. Note that for range errors, medians represent bias; in Figure 2.4, all medians are positive and thus imply that these four sampling methods generally underestimate the range.

One can see that both more accurate convergence and higher incidence of common convergence should be highly desirable, as should higher incidence of common non-convergence, as presented in the section entitled *Common Non-Convergence* below. The common convergence numbers are shown over different definitions of convergence range in Figure 2.5(a).

2.5.2 False convergence

We can derive a sample's False Convergence (FC) by the set difference of converged samples *C(Sample)* with the set of common convergences, CC from Equation 2.4:

Table 2.1: Chi-squared analysis for categorical FC and FNC tendencies. $O(FC)$ and $O(FNC)$ represent observed False Convergence and False Non-Convergence rates out of 1024 samples for each sample method, with convergence being a model fit at range $a < 500$ points. $E(FC)$ and $E(FNC)$ represent the weighted expected FC and FNC rates out of the $n = 1391$ total false results represented by the four samples. Finally, $(O - E)^2/E$ represents the standardized squared difference between observed and expected values, which when summed yield the χ^2 statistic of 52.05. Which, with degrees of freedom $df = 3$, implies categorical distinctness at $\rho < 0.001$ across incorrect spatial correlation detection results per method.

Method	$O(FC)$	$O(FNC)$	Total False	E(FC)	E(FNC)
Random	137	164	301	122.59	177.87
Star	181	166	347	141.31	205.04
L-Grid	147	215	362	150.39	218.22
Swiss	103	278	381	155.50	225.63

Method	FC $(O - E)^2/E$	FNC $(O - E)^2/E$
Random	1.69	1.08
Star	11.15	7.43
L-Grid	0.76	0.05
Swiss	17.73	12.16

$$FC = C(Sample)/CC \tag{2.5}$$

Of all the methods, Star had the most false convergence: the error of finding, via a spherical model, spatial trends where no spherical trends exist in the real data. The Swiss method erred on the side of more false non-convergences, whereas both L-Grid and Random more evenly distributed their false convergence and false non-convergence. Table 2.1 shows the FC values for ranges less than 500 points.

The reason for Star's false convergences come from its pockets of concentrated points near the outer areas of the sample. In Figure 2.1, the reader may observe sections of greater and lesser concentration of points. In the few instances where small spatial correlation happened to exist in those areas in the real data, Star emphasized it enough to cause the spherical model to converge falsely.

To examine these tendencies between samples, we performed a simple type of cluster analysis through quadrat counts (e.g. O'Sullivan and Unwin, 2003) every 50 points square (5 x 5 m). This means, by definition, the L-Grid contained five points per quadrat. The Swiss Grid showed quadrat count ranges between 0 (the outer corners) and 41 (center quadrat). The random grid ranged from 4.88 - 5.09 samples per quadrat, with no readily apparent pattern. The Star method ranged from 0 - 12.68 per quadrat, with all four quadrats with more than 12 samples on the outermost rim. One can see these clusters near the edges in Figure 2.1.

When examining the categorical Chi-squared analysis of the FC and FNC categories in Table 2.1, one can note that a main distinction between the samples can be quantified by the most populated quadrat on the outer rim. In the case of Random and L-Grid, the most populated quadrat in the outer rim ≈ 5. In Star, it goes over 12, and in Swiss, it only equals 4.

For the low bias in distribution (an unbiased distribution would have every quadrat count equal 5), we see low contributions to the χ^2 total. For the two methods with bias – Star and Swiss – we see high contributions to the χ^2 total, with the Swiss method demonstrating the most categorical bias. Note that here bias does *not* mean error; rather, it demonstrates a method's tendency to have more FNC or FC within its total false results ($FNC \cup FC$) relative to the other samples.

Also of interest may be that although four quadrats of the Star method along the outer edges averaged more than 12 samples over the 1024 runs, all of the outer edge quadrats together averaged to 5.22 samples per quadrat, explaining the success of Star to find spatial correlation over the dataset.

2.5.3 False non-convergence

We can derive a sample method's False Non-Convergence (FNC) by the set difference between the set of converged real data *C(Real)* versus the set of common convergences, CC from Equation 2.4:

$$FNC = C(Real)/CC \tag{2.6}$$

The numbers for false non-convergence appear in Table 2.1. Overall, the Swiss grid presented the greatest tendency to not converge on a good spherical model fit when the real data did. This means it was more likely to present a linear semivariogram (or a very large, poorly fitting spherical model) when the real data presented stronger spatial correlation. We feel that this occurred due to the concentration of points in the center and very few near the edges, as discussed above via quadrat counts in the section above entitled *False Convergence*. Such a method may only detect spatial correlation in that area well. Indeed, other sampling methods, such as the MT2004 grid as described in Kronholm and Birkeland (2007), have been developed to try to spread the points over a larger area while retaining the field advantages of gridded construction.

2.5.4 Common non-convergence

To be a correct model of the real data, a sampling method should not only converge upon a range with a given model (here, the spherical one in Equation 2.3), but also not converge when the model does not fit the real data well. We can find the Common Non-Convergence (CNC) via the set difference of all real data models, $\{Real\}$, with the union of both convergence sets, $C(Sample)$ and $C(Real)$, as given in Equation 2.7:

$$CNC = \{Real\}/(C(Real) \cup C(Sample)) \tag{2.7}$$

Though all real data shows visual spatial correlation, some real data instances did not fit a spherical semivariogram model well. Two examples are shown in Figure 2.3, where one can see obvious spatial patches which we may wish to discover via our sampling method. And yet, the corresponding variograms next to the images are obviously non-spherical.

In fact, from the low numbers of real data instances which did not converge, the reader may observe that – although common in snow data analysis – spherical variograms and even variograms in general may not necessarily be good solve-all tools for detecting patterns and process effects.

Regardless, CNC, along with CC, completes the set of correct spatial correlation answers a sampling method can produce via the variogram and a model. Though technically uninteresting as CNC does not actually discover any spatial range, it assists in calculating the Correct Spatial Correlation Detection total for a sampling method, as discussed in the next section.

2.5.5 Correct spatial correlation detection

When we know the common convergences from Equation 2.4 and the number of common non-convergences from Equation 2.7, we know the total number of *correct spatial correlation answers* a given sampling method finds. This total gives the measure of Correct Spatial Correlation Detection (CSCD) for a sample:

$$CSCD = CNC \cup CC \tag{2.8}$$

Figure 2.5b gives the numbers of Correct Spatial Correlation Detections for each sample over various ranges. In this case, rather than simply using 500 points (50 m) as a general convergence measure, we found how many correct answers each sample detected for ranges 500-100 points, at intervals of 50 points (5 m). Since linear and other non-spherical variograms forced to fit spherical models can be identified by extremely large range values, this initially assists in sorting out which *should* be converging on a good model fit, and for lower ranges simply reduces the window of good fits we get to analyze.

As the number of converged simulation data points become fewer and fewer, it becomes easier and easier to use a so-called 'ignorant' sampling algorithm. For example, being picky enough to only call ranges of less than 100 points (10 m) leaves us with so few real data convergences (154 of 1024 samples) that we could theoretically not sample anything at all, and thus never converge, and still achieve the 'correct' answer (non-convergence) over 80 percent of the time. This should give some intuitive sense of the instability demonstrated at left edge of Figure 2.5b at ranges of 150 points and fewer.

Finally, one may note from Figure 2.5b that, even in a best case, no sampling method detects spatial correlation correctly better than three out of four times. Considering various definitions of range, the three non-Random sampling methods perform comparatively.

2.6 Discussion

Above all, this paper demonstrates that different sampling methods have different strengths and weaknesses. Of course, all statistical strengths and weaknesses presented here depend greatly on the spherical model fit.

But given that, for the smallest range error bias per known converging variogram

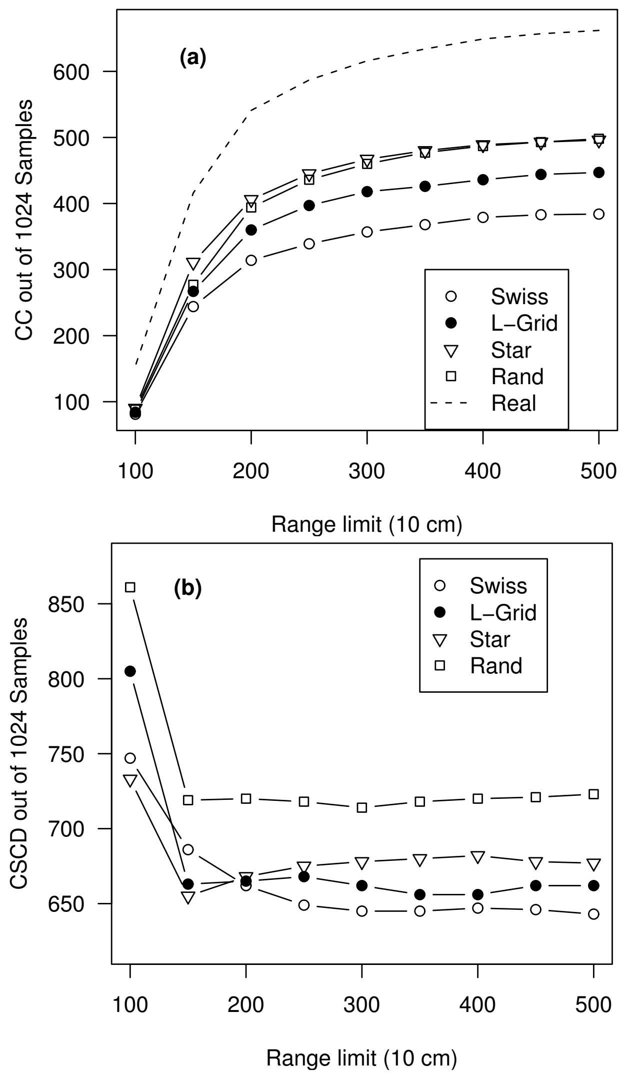


Figure 2.5: (a) Common Convergence (CC) and (b) Correct Spatial Correlation Detection (CSCD) graphs. Definitions vary by what range limit we choose to define as a good spherical model fit, and we have shown the results over various definitions of convergence 10 – 50 m. Note the instability at 150 points (15 m) and less for CSCD; the corresponding ranges in the CC graph show that relatively few data points exist at this definition and thus we do not obtain a good model.

model, the L-Grid seems the obvious choice. For focus at the center of an area, the Swiss grid may assist in revealing details there with its point concentration. For detection of unknown ranges of a process, the Star sample properly matches or rejects a spherical fit when the actual data does more times than the other methods, implying better correlation across varied data sets. In addition, Star has efficient design and a small standard deviation of range error in its common convergences. Of course, the pure Random model presents the best correlation and smallest error of all but remains very inefficient to properly implement in the field.

2.7 Conclusions

Given the strengths and weaknesses of each sampling method, and given the range of applications for each method, one cannot simply say one method presents the best mix. Efficient ease of use allows one to obtain the most points in a given time. Statistical robustness allows one to feel more confident in the results.

As an initial design of a sampling method intended to make randomness usable and efficient in the field, the Star method shows promise for linear random sampling methods in general due to its ease of layout plus comparably accurate spatial correlation detection. One could spend some time minimizing the clustering effects in the corners to improve Star and better approach a purely random distribution.

However, when measuring spatial variability on snow, we cannot really know what the variability range is without measuring every single point. Thus, when using the variogram, we only know the spatial correlation range of our measurements, and a little about how good our measurement methods probably are. Furthermore, when we find spatial correlation in our sample data without knowing every measurable point, False Convergence and False Non-Convergence will occur with any sampling method, and Figure 2.5b shows that even in the best case FC or FNC will occur at least one out of four times for these sampling methods using this dataset.

The greatest question revealed by this paper centers around how to model spatial correlation in snow. The real data we used had visual spatial correlation in every sample; however, the spherical model converged on a reasonable fit to the semivariograms of those data fewer than two out of three times. One might benefit from scaling a sampling method to extend well beyond the expected range, which would then obtain more robust measurements of spatial correlation. However, the destructiveness of snow sampling seems to often present a barrier to adaptive sample method scaling.

Furthermore, when treated by hand to detect anisotropy, to discover better fits with different models, or to use techniques other than the variogram, the sample data will probably be much less limited than presented here. Thus, there seems to exist a great deal of possibility in investigating how variograms – or other process detection and correlation tools – can be best used with the particulars of snow science sampling.

2.8 Acknowledgements

The authors wish to thank Dr. L. Bentley and Dr. G. Margrave of the University of Calgary for their thoughtful comments on a draft of this paper. In addition, the paper greatly benefited from careful review by two reviewers, including Dr. K. Kronholm.

For financial support, the authors gratefully thank the Natural Sciences and Engineering Research Council of Canada, HeliCat Canada, the Canadian Avalanche Association, Mike Wiegele Helicopter Skiing, Teck Mining Company, Canada West Ski Areas Association, the Association of Canadian Mountain Guides, Backcountry Lodges of British Columbia, and the Canadian Ski Guides Association.

Chapter 3: Google Earth Imagery and Surface Hoar

Remote sensing has opened new avenues and methods for many fields, and avalanche science is among them. The paper which appears here addresses an applied use of Google Earth imagery, and is slightly modified from the one entitled *Spatial distribution of surface hoar crystals in sparse forests* which appeared in the journal Natural Hazards and Earth Systems Science as Shea and Jamieson (2010c).

For more information on further use of this method, see Section 6.2 in Chapter 6.

3.1 Abstract

Surface hoar size and location relate directly to avalanche initiation trigger points, and they do so in small-scale spatial distributions. Physically, surface hoar will grow where the snow surface is cold relative to the air and water vapour is plentiful. Vapour aside, snow cools at night primarily by longwave radiation emittance. Emittance can be restricted by clouds, trees, and terrain features. With 96 independent spatial point samples of surface hoar size, we show the extreme small-scale size variation that trees can create, ranging from 0 to 14 mm in an area of 40^2 m^2. We relate this size variation to the effects of trees by using Google Earth images to estimate the amount that trees impinge on sky view for each point. Though physically related to longwave escape, radiation balance can be as difficult to estimate as surface hoar size itself. Thus, we estimate point surface hoar size by expected maximum areal crystal size and dry terrain greyscale value only. We confirm this relation by using it at a different area and in a different formation cycle. Then, its overall average error was 1.5 mm for an area with surface hoar sizes ranging from 0 to 7 mm.

3.2 Introduction

Surface hoar crystals grow on the surface of snow from direct deposition of water vapor. The crystals, once formed and buried under additional snow load, form a persistent and brittle weak layer (McClung and Schaerer, 2006) with a distinct ability to propagate fractures within the snowpack and thus release avalanches.

As the surface hoar crystals collapse and slide in layer failure, the size of the crystals themselves may contribute to increased potential energy release and increased propagation of layer failure (Jamieson and Schweizer, 2000; Heierli et al., 2008). In addition, larger crystals maintain an unstable separation between slab and bed surface longer during the process of settlement and bonding (Jamieson and Schweizer, 2000).

In other words, an ability to predict where surface hoar crystals form and whether they are large or small would directly help predict and clarify where skiers may trigger slabs overlying a surface hoar layer. The scale of interest for such predictions lies at a small enough scale where we may say something useful about trigger points.

This paper focuses on the *where* rather than the *when* of surface hoar – often, avalanche forecasting operations have a good idea about whether or not surface hoar has formed, but desire additional information about the spatial distribution of the layer. Here, we examine the size of surface hoar crystals over small areas in sparse forests.

Besides being common skiing terrain in North America, trees provide control for many environment variables. Sparse forests offer protection from high winds, and they create extreme variation in net snow surface longwave radiation escape. Thus, such areas are known for being protected enough to grow surface hoar, but variable enough to not grow it homogeneously. Logging cuts serve as a prime example (McClung and Schaerer, 2006). In addition, understanding the variation of surface hoar in sparse forests has direct useful value: between 1984 and 1996, of the 13 fatal avalanches that released on buried surface hoar for which vegetative area is known, six of them happened at treeline or below (Jamieson and Geldsetzer, 1996).

Physically, we expect that the size of surface hoar can be directly and spatially correlated with sky view, a measure of how much atmospheric view and exposure a point on the snow surface has, and through which it may emit longwave radiation. In forests, close trees over a point on the ground decrease sky view, whereas widely spaced trees increase it. As sky view is difficult to measure point-by-point over large areas (Brown et al., 2001), we approximate it by using greyscale values in dry land terrain imagery – where darker pixels correspond to trees.

We begin with an overview of previous studies of physical parameters in surface hoar formation. Our methods, including use of land imagery, follow. Then, we present options for analysis: as this study is the first to attempt spatial surface hoar size prediction *without* accompanying weather data, we show the avenues that led us to using Google Earth. Finally, we present our empirical model, its validation, and a summary and outlook.

3.3 Previous Work

Whereas previous studies have linked surface hoar growth to bulk transfer of water vapour, few have examined net longwave radiation effects and none have done so in a two-dimensional spatial setting at the trigger scale.

Many factors affect surface hoar growth, from wind (Hachikubo and Akitaya, 1997; Hachikubo, 2001; Föhn, 2001), to radiation balance (Cooperstein, 2008), to air-snow temperature gradients (Lang et al., 1984; Hachikubo and Akitaya, 1997), to humidity (Feick et al., 2007). At the most simple physics level, one might say surface hoar grows when we have cold ice, warm air, and lots of water vapour. However, some factors affect growth both positively and negatively. Clouds, for example, supply vapour but reduce effective longwave radiation escape (Colbeck et al., 2008).

Surface hoar needs a substantial amount of water vapour. This explains the success of bulk transfer models (Colbeck, 1988; Hachikubo and Akitaya, 1997; Höller, 1998; Föhn, 2001; Lehning et al., 2002): wind can provide moisture to the snow surface via bulk transfer. However, humidity does not always serve as the main determining factor. Previous work showed similar air temperatures and relative humidity both within and out of forests (Höller, 1998), but we observe high surface hoar size variance within forests. So, if air temperature and vapour presence do not change spatially within sparse forests, the remaining varying factors should be wind – or bulk transfer capacity – and longwave radiation escape.

Wind presents its own particular set of challenges. Some studies found light wind to be beneficial to surface hoar growth (Föhn, 2001; Hachikubo and Akitaya, 1997; Colbeck, 1988). Yet, there are recorded examples of wind (with relative humidity close to that of vapour pressure over ice) sublimating ice rather than depositing vapour upon it (Hachikubo, 2001; Feick et al., 2007). In addition, the physical necessity of wind in the deposition process itself remains unresolved, with conflicting opinions in Lang (1985) compared with Hachikubo and Akitaya (1997) and Föhn (2001). At the very least, high ($> 3\ m\,s^{-1}$) wind speeds at the snow surface will affect the shape of deposition into something other than surface hoar (Cheng and Shiu, 2002), and low wind speeds will present large barriers to bulk transfer of adequate vapour.

Wind affects both major factors in surface hoar growth – temperatures and vapour – and it does so in a currently un-modellable spatial fashion at useful scales (Hägeli and McClung, 2000; Campbell et al., 2004). For this reason, most of the studies mentioned above involve measurement points only very near weather stations or other monitoring equipment such as radiometers. Here, then, we have come to an apparent impasse: our best mechanisms for estimating surface hoar size require accurate wind and humidity measurements, and neither may be predictively mapped, spatially, within a 40^2 m^2 area. Thus, for this first spatial predictive effort, we chose an environment that controls the variance of both.

This leaves the option for a spatial study on the variance in effective longwave radiation escape. Longwave radiation effects have already been studied in relation to surface hoar growth by aspect (Cooperstein, 2008) and integrated with bulk transfer-based estimates for better predictions in open areas (Lehning et al., 2002).

An excellent location to study these longwave effects is in sparse forests. Trees block sky view for a point on the ground. Trees are detectable – individually – over large areas using modern basic remote sensing. Trees are usually stationary over long time scales, have high emissivity which makes them quite good at emitting longwave to snow (Ellis et al., 2011), and vary spatially on scales that humans can access, measure, and visualize. Furthermore, trees offer the conditions desired here – that is, protection from high winds.

Finding a clarifying situation, such as sparse forests, with which to study sky view and longwave blockage effects can be quite helpful. The self-compounding nature of physical factors involved in surface hoar growth – as well as their small-scale variability at any give point – have put up great obstacles to studying surface hoar growth by point sampling sparsely over a large area (Feick et al., 2007). Also because of this, very few studies have tried to map the variance of these effects over terrain, and none – until now – have attempted spatial prediction.

3.4 Methods and Data

This paper contains two types of field measurements: surface hoar crystal sizes and snow surface point temperatures. We also compare these field data to terrain images to obtain a measure of closeness to trees. The methods for both obtaining the data and correlating the data and images are presented here. In our methods, we sought to:

- Provide a reasonable estimate of surface hoar crystal size variation over terrain at the trigger scale
- Utilize only information easily available to the practitioner: i.e., single point samples of surface hoar size and basic dry land imagery from Google Earth

In other words, we intentionally present a method to estimate spatial distribution which uses limited input, few necessary resources, and no field instrumentation.

3.4.1 Field method and data

In our area of operation in the Columbia Mountains, Canada, we point-sampled surface hoar sizes during the surface hoar formation periods between 16 January to 22 January, and 17 February to 20 February, 2009.

We spatially distributed the point samples using the Star sampling method (Shea and Jamieson, 2010d). This ensured randomized spacing between points and snow surface preservation between sampling days. The random placement of points prevented the operator's preference from being influenced by terrain and other attributes of the area. In addition, the method allowed random spacing between analysis points; such randomness generally improves the robustness of spatial analyses later done with that data (Kronholm and Birkeland, 2007). The slope preservation enabled by Star's layout allowed us to measure the change in surface hoar size at the same locations from day to day.

Sample point distance separation ranged randomly from 10 cm to the full width of a sample area. Each sample area contained 48 crystal size samples and covered approximately 40 x 40 m of terrain. Due to the fractal nature of snow (Deems et al., 2006), we chose our extent carefully to be one that could capture the scale of skier triggering.

At each point, we recorded the minimum and maximum size of surface hoar crystal found there, which we later averaged into a mean size for that point. We noted a size of 0 mm only if no crystals were found. The same points were sampled repeatedly on consecutive days during the January cycle.

The efficiency of the sampling design allowed us to obtain more spatial points than any previous surface hoar study. In short, 96 spatially separate and semi-random points were obtained in the January cycle, and each point was visited three times over four days. In the February cycle, 48 spatially separate points were obtained, and they were each visited once.

The two sample areas in January, from which we developed the relation with vegetative greyscale in visual terrain imagery, need be distinguished. We will call them Area I (primarily north facing) and Area II (primarily northeast facing). Both lie at an elevation of 1900 m, with generally horizontal orientation but containing small slopes up to 25 degrees of incline. Neither sample area had dense forest, and neither was entirely open, although Area II had more open area than Area I. One can see a photo of the general terrain in Figure 3.1. Each contained a mix of trees and rolling terrain, which meant any immediate obstructions to sky view were of a smaller size than the sample area.

The one sample area in February, which we used to confirm the relation with

Figure 3.1: Photo of a portion of Area I. Notice the ski tracks across the area for scale. This area is quite sheltered from wind. Also, during the January surface hoar cycle, the winds in the area ranged from only 1 to 12 km h^{-1}.

greyscale, is located at a different location (150 km to the northwest), and a different aspect (south facing). The attributes of the area were similar to Area I and Area II in all but aspect, that is, a mix of trees and small terrain features.

3.4.2 Temperatures

For our temperature measurements, we measured the surface temperature of the snow at each point with two different infrared thermometers – a Thermohawk 400 hemispherical 1:1 thermometer used at 1 m above the surface, and a Testo 825-T4 3:1 thermometer used at 30 cm above the surface. This occurred only for the last January cycle day, for both Area I and II. As both measurements had similar results, we present the downward facing measurements from 30 cm above the snow surface. What we call *point* snow surface temperatures, then, are effectively ten centimeter diameter averaged snow surface temperatures.

3.4.3 Imagery

Google Earth photos of Area I and Area II appear in Figure 3.2b and 3.2d, with the blurry dark areas corresponding to trees. We obtained our land cover imagery from Google Earth. Individual trees can be distinguished; for freely available data, these photographs are quite good. The images exhibit no snow on the ground or trees; we use them only for greyscale values to estimate sky view. The trees are primarily conifers, which can provide dependable greyscale shading distinctness.

From a screenshot of the images overlain on Google Earth terrain, we extracted the terrain-projected (warped) image area corresponding to each of the two sampled areas. We then used cubic interpolation to expand the projected areas back into flat rectangular form via an image processor. No other image modifications besides a projection to a flat rectangular surface occurred. Area II required more interpolation than Area 1 to make it rectangular, with approximately 25 degrees of stretch in the lower right hand corner, and smaller adjustments in the other corners.

Each Star sample area of approximately 40 m x 40 m translated to a terrain image with about ten pixels per meter resolution. This includes pixels generated by cubic interpolation when re-projecting the image to a rectangular form.

Greyscale shading in our photographs comes primarily from trees, less so from rocks and terrain rolls, and nearly nothing from steep slopes or aspect within a given small area. This gives a clean, qualitative link to sky view: areas of lighter grey in a small area generally have more, and areas of darker grey within a small area generally have less.

3.4.4 Mapping

To make visual maps of the crystal size and surface temperature measurements shown in this paper, we created a continuous image of probable values from our real-world point measurements. This step is simply to make it easier to visually comprehend the measurements; all data points involved in calculations in this paper are only those actually physically measured in the field.

To produce these visual maps, we used the standard inverse distance weighting (IDW) algorithm. IDW weights the values of closer known points more heavily than further known points, where the weight w of each known points with distance d from the point we wish to predict gets included in the average with a weight $w(d) = d^{-2}$. We used IDW to calculate size and temperature values at all locations on a 1 x 10^{-5} spaced latitude and longitude WGS84 grid (Pebesma and Wesseling, 1998). Two such example maps can be seen in Figure 3.2a and 3.2c.

3.4.5 Determining position

We determined the latitude and longitude of the sampled point locations using the following information:

- Between 600 and 1400 GPS reference points per visit to each sample area
- Pacing each transect of sample points (Star consists of six transects) using pace length and various measuring devices
- Recording reference information in addition to crystal size at each point, including patchiness of surface hoar crystals at each point, which was used for correct day-to-day ordering of points

These independent sets of information essentially served as checks to assess the accuracy we could obtain from any one of the measurement methods. Our base fitting method consisted of a custom-designed GIS fit program, which determined the position of each sample location by closest fit from the average of thousands of compiled track points.

We estimate that our total cumulative error per 8-point transect line with GPS and pacing error combined (and practically inseparable) ranged between 1 m and 5 m depending on the sample. We used a relatively accurate recreational handheld GPS unit to let us not only move quickly while sampling, but also to demonstrate that such a device can provide position for the field measurements needed for others to use the method outlined in this paper.

When determining position on the Google Earth photography, we moved the coordinates of Area II, as a unit, seven meters to the south. This occurred due to visual match up of tree location values recorded in the field with what we observed on the Google Earth imagery. This shift was most probably needed because of the different tilt of the trees in the photograph, creating longer shadows in the image. Area I required no translation, and we did not perform any translation on the sample from 19 February 2009 to let it serve as an evaluative dataset.

3.5 Options for analysis

Where weather data is concerned, many-point spatial studies face a choice of challenge. They may use measurement equipment at every point, which can be expensive and time consuming. Or, they may model and extrapolate values from nearby measurement equipment, which can result in loss of accuracy.

With our intent to develop a model which uses readily available inputs, we discov-

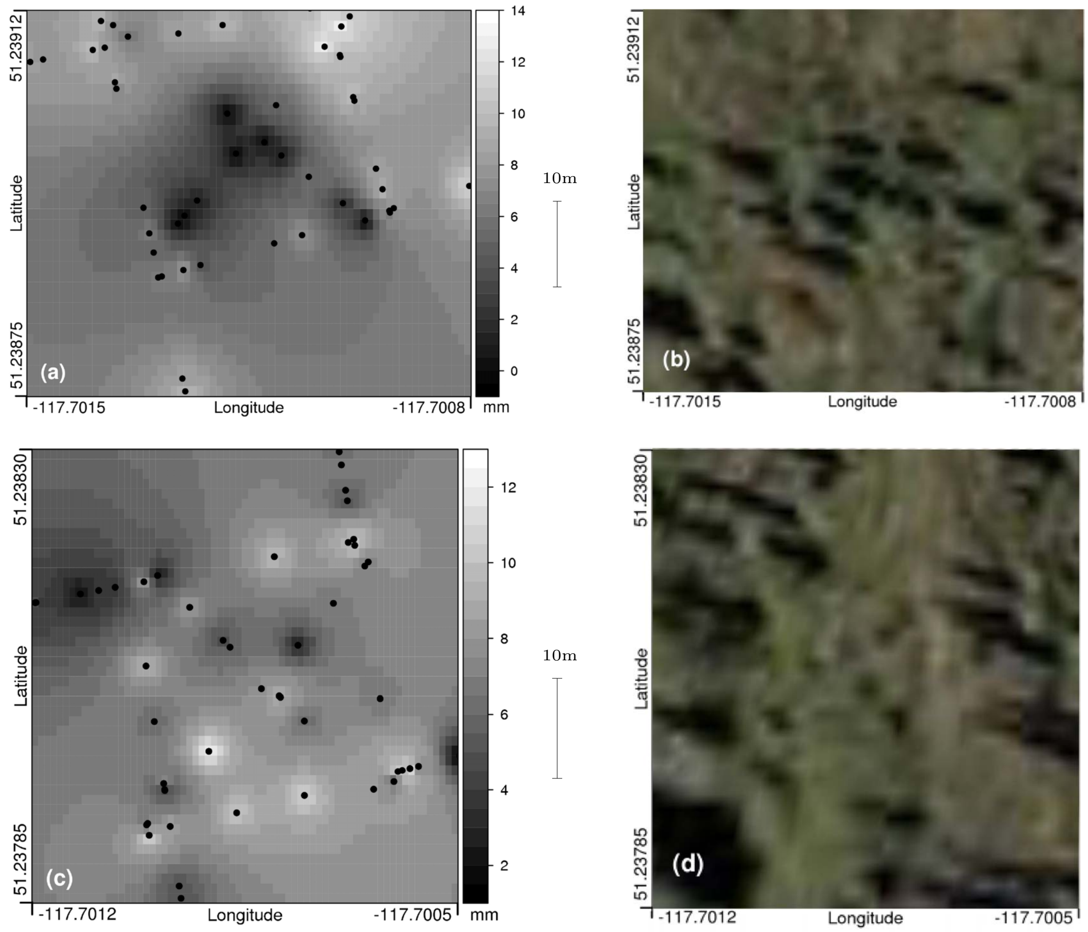


Figure 3.2: Visual overview of the two sample areas. (a) IDW map of mean crystal sizes for one day at the Area I location, (b) Google Earth image corresponding to Area I. (c) IDW map of mean crystal sizes for one day at the Area II location, (d) Google Earth image corresponding to Area II. For (a) and (c), black dots indicate the actual physical sample locations. For (b) and (d), images are provided under the Google Earth terms of use, copyright 2009 TeleAtlas, 2009 British Columbia, and 2007 Google Earth.

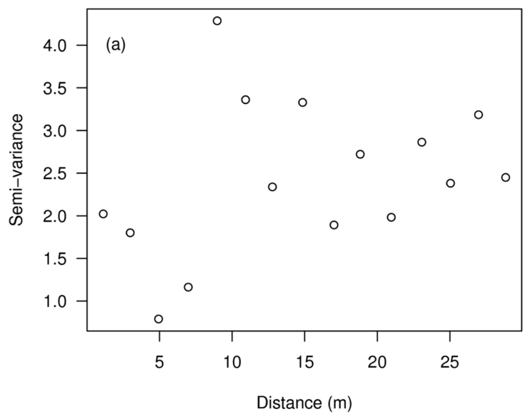
ered that, in our data, crystal size generally scales with Google Earth photography greyscale. This is not to say that greyscale scaling will *always* work. Rather, we show here that it worked in our conditions, which we consider to be typical of sparse forests. Furthermore, the use of greyscale shows that other, non-weather data may be used to creatively augment, improve, or even provide estimates. We did not arrive at such a solution immediately, and the following subsections outline our process.

3.5.1 Variograms

As an exploratory method for discovering spatial structure and process scales, the variogram (Cressie, 1993) currently has no peer. Conservatively, one may determine the process scales for a given process and area, and then return to that area and scale a sampling method to capture that process on its scale of operation as accurately as possible. Oftentimes, however, one does not have that luxury as snow sampling can be destructive and the desired conditions last only a short time.

We could not re-scale our sample without further destruction to the slope. However, as we could return from day to day, we could find variograms both for one day and for change across multiple days. Searches for both isotropic and anisotropic spatial correlation in the single-day surface hoar size sets did not reveal any obvious patterns. However, the multi-day variograms had more defined ranges, as these distill the data down to only the processes which change crystal size after formation. Even these variogram ranges were non-definite. Figure 3.3 shows two such variograms.

The variable and weak nature of these variograms revealed two issues with the analytical use of the variogram here. First, the process scale does not necessarily stay constant from day to day, or area to area. As we can see in Figure 3.3, even the weak ranges vary for the same analysis from day to day. One may intuit that trying to chase a constantly varying set of process scales would be quite time consuming, and



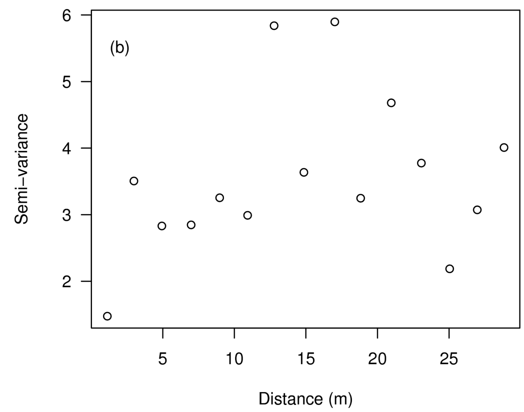


Figure 3.3: Two examples of size-difference semi-variograms with weakly defined ranges. (a) Overnight size change from 16 to 17 January, Area II. (b) Overnight size change from 17 to 18 January, Area I.

would also be quite spatially consuming when the sampling involved is a destructive process.

Second, the type of sampling method and the scaling of that sampling method greatly affect the outcome of a variogram. Due to the fractal nature of some snow measurements (Deems et al., 2006), a sampling method with points that oscillate with the process scale can greatly affect the outcome, and even pick up and mix up ranges of unexpected interrelated processes.

We cannot know whether either of the above issues caused the lack of good variogram results. However, our sampling method evaluation (Shea and Jamieson, 2010d) and close measurement point spacing give weight to the variogram not being useful in these conditions. Furthermore, a variogram cannot help one spatially predict any value, unless the process creating the range can be very precisely mapped over the desired area.

3.5.2 Surface temperature

We also assessed the physical processes. We measured surface temperatures, since surface hoar grows on ice when it is cold relative to the air (Lang et al., 1984; Hachikubo and Akitaya, 1997). At the end of the January formation cycle, we took a nighttime point surface temperature (0400 to 0600 local time) and daytime point surface temperature (1000 to 1200 local time) at every one of the 96 spatial points in Area I and Area II. We show crossplots of surface temperature versus crystal size in Figure 3.4.

The correlation between nighttime surface temperature and crystal size ranged from $r = -0.64$ to -0.74, both $p < 0.001$. The correlation between daytime surface temperature and crystal size ranged from $r = -0.39$ ($p = 0.03$) to $r = -0.31$ ($p = 0.06$). This supports the findings of Cooperstein (2008): a single temporal measurement of the effects of daytime shortwave matters less than the overall sum of shortwave, which a night of longwave escape must then overcome. However, this result does not necessarily help with size prediction, for the following reasons.

First, the process of directly predicting such point temperatures at small scales over terrain would be very complicated. Though current work has begun to tackle this very problem (Morstad et al., 2007), much work remains. The magnitude of those efforts should demonstrate the complexity of building and verifying a usable physical predictive model for temperatures alone, much less for surface hoar size.

One also may think to use trees via greyscale to predict the more physical variable of surface temperature, and then use the temperature predictions to estimate size. However, the relationship between distance from trees and point surface temperature do not necessarily correlate. For example, during the day, open and non-treed snow absorbs more incoming shortwave radiation, and thus would have to emit that extra absorbed energy at night to achieve the same surface temperature. Snow close to

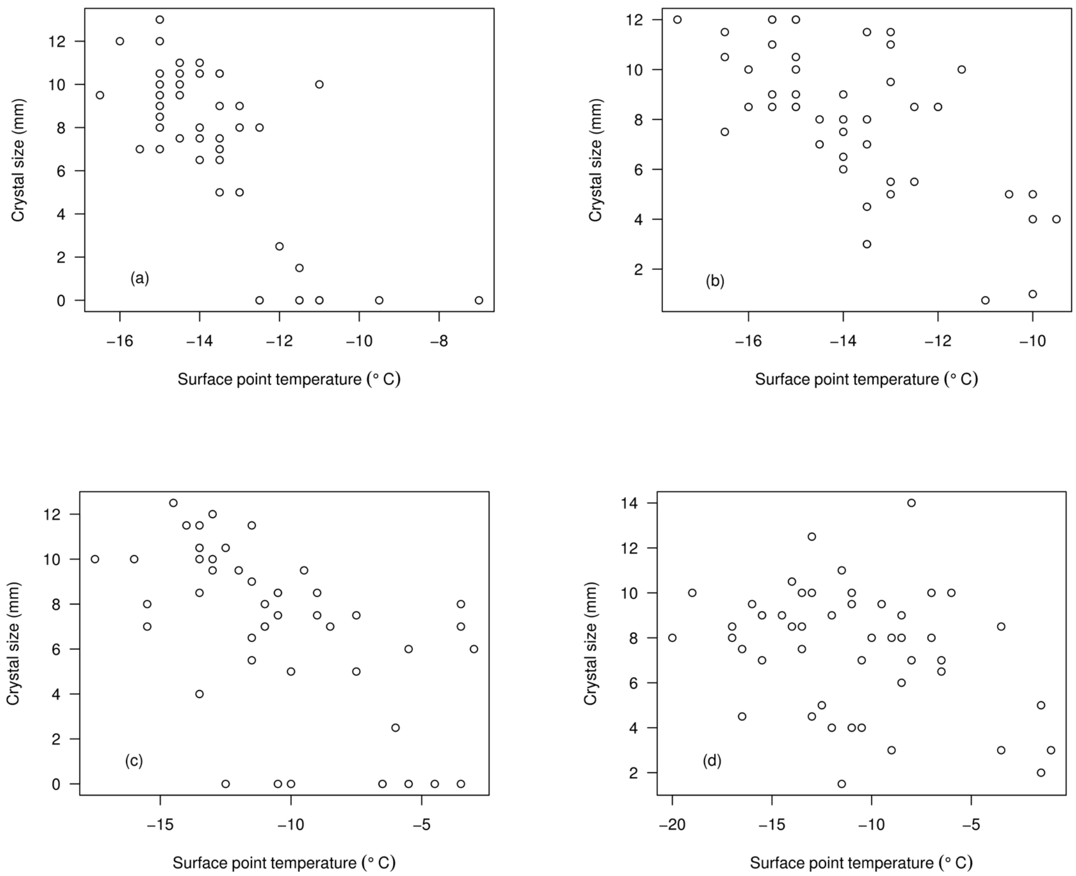
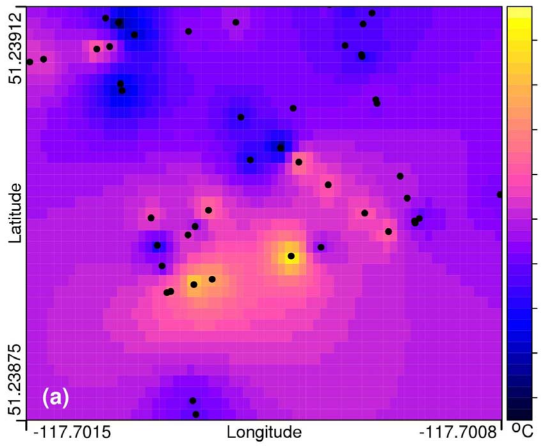


Figure 3.4: Cross plots of point snow surface temperatures versus mean surface hoar crystal size found at each point during both night and day. (a) Night, Area I. Pearson correlation $r = -0.74$, $p < 0.001$. (b) Night, Area II. Pearson correlation $r = -0.64$, $p < 0.001$. (c) Day, Area I. Pearson correlation $r = -0.39$, $p = 0.03$. (d) Day, Area II. Pearson correlation $r = -0.31$, $p = 0.06$.

trees, on the other hand, gains some protection from incoming shortwave during the day. From only these effects, larger surface hoar would grow within trees, which we do not observe in practice.

The more dominant effect of trees on nearby snow is probably a reduced ability for the snow to effectively vent heat via longwave radiation escape at night due to tree blockage and re-radiance. Trees also can create canopies which keep the air warm



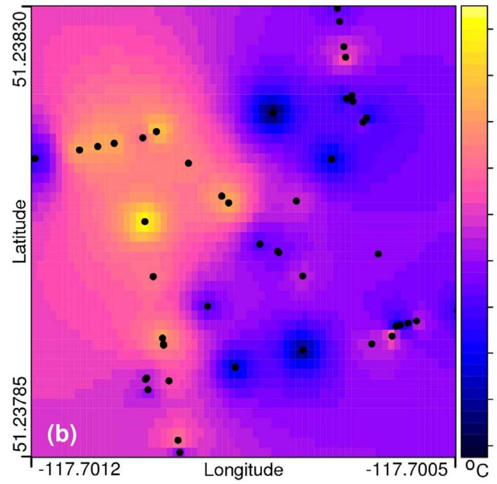


Figure 3.5: Inverse-distance weighted plot of temperature *change* from night (approximately 0500 hrs) to day (approximately 1100 hrs) on 21 January, 2009. Black dots indicate the locations of the actual sample points. (a) Area I. (b) Area II.

by buffering it from temperature changes (Sicart et al., 2004), an important factor in near-surface temperature gradients but not accounted for by surface temperatures alone.

Even if we were to be able to estimate point surface temperatures at small scales over terrain, the strong size correlation *solely* with nighttime temperatures indicates that a radiation balance estimate would need to not only be accurate in space, but also quite accurate in time, compounding the difficulty. One may examine the spatial IDW maps of the *change* in surface temperatures from night to day in Figure 3.5. Even during the daytime, some areas cooled and some areas warmed in a complex interplay of shadows and radiation balance. The most noticeable surface warming actually occurred at shaded points within the trees, as one may see by comparing the temperature plots in Figure 3.5 with their corresponding images in Figure 3.2.

Furthermore, this sets aside the issue that surface temperatures account for only part of the variance in size, as most natural variables do. Estimating temporal surface temperatures at such small scales with all of the complexity mentioned above presents a great deal of difficulty for the predictive benefit of only one variable. So, we sought an empirical approach.

3.5.3 Greyscale

Trees affect skyview – which in turn influences diurnal surface temperature changes – and trees also affect air temperatures, wind, and so on. They capture the effects of many variables in one physical construct. Due to surface hoar needing efficient longwave radiation cooling, we hypothesize that, all other things being equal, surface hoar will be smaller near trees and larger in open clearings.

To verify this empirical relation, we used the greyscale values of dry land photography. The point crystal size measurements and corresponding terrain imagery for the two January formation cycle areas are shown in Figure 3.2. We expect that a single pixel on the terrain photo, due to noise and extreme sensitivity to local effects, will have poor correlation to a single crystal point measurement. Rather, for purposes of both reducing the effects of noise and accounting for a finite area around a geographical point, we averaged a given radius of greyscale pixels.

To find the best such radius, we constructed eighty different greyscale point sets, each containing one pixel value per measurement point, and each corresponding to a different radius – one to eighty – in linearly averaged greyscale pixels. For a radius of 10 pixels, for example, we averaged the block of 21 by 21 pixels surrounding the sample point to produce our averaged point greyscale value. This occurred independently of nearby or overlapping sample radii. We then compared the correlation between the mean surface hoar size at each measurement point and the averaged greyscale values within a given radius in the terrain image at that point. The fluctuations in correlation values by radius can be seen in Figure 3.6.

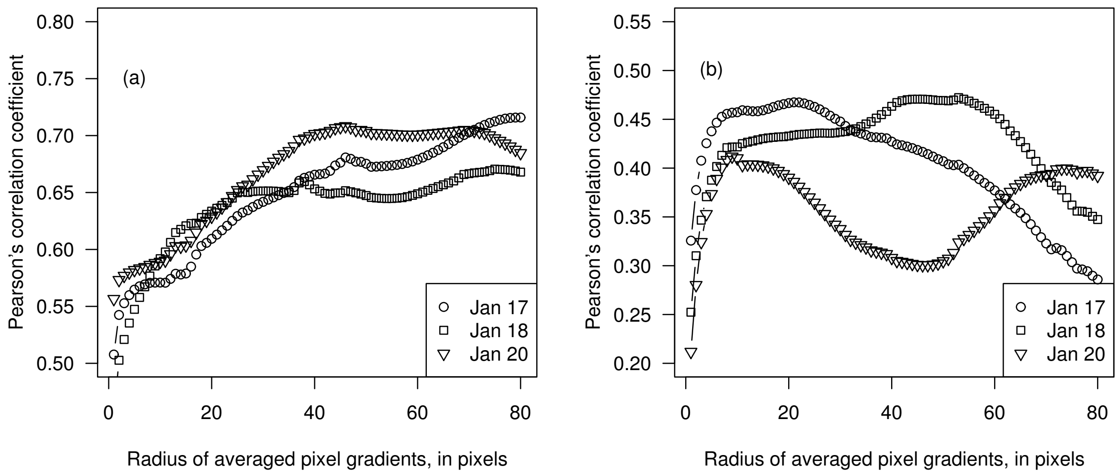


Figure 3.6: Pearson's correlation coefficient variation by radius of Google Earth image pixels included in a point averaged greyscale value. Each correlation coefficient shown comes from the 48 points per given day. The values for $grey_{40}$ correlation lie at $x = 41$. (a) Area I. All correlations have significance $p < 0.05$. (b) Area II. All correlations except the three points furthest left on each line have significance $p < 0.05$.

Qualitatively, one may think of this as the area of effect that trees might have on a surface hoar crystal or, inversely, as an empirical measure of sky view. In Figure 3.6, one can see that an obvious maximum exists on some days but not always, and the maximums do not always occur at the same radius value.

Our choice of radius was the maximization of correlation over the six days, which lies at approximately 40 pixels. Though this general maximum is obviously not individually true for 20 January 2009, 40 pixels gives us an easy to use and easy to generate value which has reasonable correlation. Compared to the effort of predicting spatial surface temperatures, for example, this radius-of-tree-effect variable gives straightforward values. The 40 pixel radius corresponds to an 81 pixel diameter, or an area with a radius of about four meters on the imagery used here. We call this averaged greyscale value over a 40 pixel radius $grey_{40}$.

The potential drawbacks to this method are also readily apparent. In open areas, where this qualitative relation between skyview and greyscale variation does not hold, this radius of effect also would not hold. However, for the purpose of sparse forests, it cleanly isolates the amount of tree cover around a point and has good correlation to surface hoar size.

3.5.4 Regression

Having found an easily and spatially obtainable value – $grey_{40}$ – we wish to use it to do something spatially useful. We first constructed six different linear models using basic linear regression with unit least squares and utilizing R (R Development Core Team, 2006). These six models correspond to using $grey_{40}$ as the independent variable on each of the three sample days and two areas. We captured the slope a and intercept b of each linear regression model.

The impracticality of using an ideal day as a model for other days became readily apparent. The slope and intercepts of each of the six models varied at least partially because the average maximum surface hoar size for each sample area varied (14.7 mm for Area I, and 12.5 mm for Area II). Determining which of these maximum sizes represented the ideal could not be easily done. In addition, some days appeared to have a more curved (exponential) relationship with greyscale, whereas the overall relationship with greyscale was linear.

Further, the range of $grey_{40}$ varied (0.177 for Area I, 0.156 for Area II), as did the other averaged grey point values. Even when slopes and intercepts were scaled by maximum crystal size and $grey_{40}$, the slopes of greyscale luminosity versus crystal size varied from 26.23 to 69.49. Regression over the whole 288 point dataset and individual areas fared somewhat better, with the standard deviations of the residuals just over 3 mm.

Regression depends very strongly upon each individual smaller dataset being affected by the independent variable in the same numerical way. For example, if trees affect crystal size very strongly one day and less strongly the next, regression across the two datasets would average the effect and potentially produce a non-useful result. Conversely, regression can overfit to the effect of a single day and area. We see that here: even with a generally good relationship to the independent variable ($grey_{40}$), both slopes and normalized slopes were so overfit to a particular day as to not be useful.

Trees stay constant from day to day, and so their *type* of effect on snow should also stay more or less constant from day to day. Exceptions exist: when the entire area is covered by a cloud, for example, trees would not be the dominating factor. But for the clear conditions which surface hoar benefits from (and which existed in this study) we expect trees to have the same general effect from day to day, area to area. So we turned toward developing a relationship for a given size increase per change in $grey_{40}$.

3.5.5 Constant increase

Efforts with regression implied that we needed a way to scale, by area, for the range in both crystal size and greyscale. Both of those values do not stay constant across large areas, but as here we focus on small areas and even smaller scales of size estimation, they functionally serve our purpose. To capture the observed and generally linear relationship with $grey_{40}$, we split the greyscale range for each area by the observed range in crystal size, e.g. $range(grey_{40})/range(mean\_sizes)$. This gives an amount of greyscale brightness increase per unit crystal size increase.

We applied this to the two areas sampled in January, 2009. To do so, we needed a minimum mean crystal size (0 mm) and corresponding $grey_{40}$ value, as well as a maximum mean crystal size (14.7 mm for Area I or 12.5 mm for Area II) and a greyscale range (as stated above, 0.177 for Area I, 0.156 for Area II). With these values, we can determine a slope for a linear model to predict crystal size *for each area*. This type of scaling, though simple, adapts the expected crystal sizes to (a) the greyscale range of each area, and (b) the crystal size range in the area.

Both the January sample areas each had the same averaged minimum $grey_{40}$ at 0 mm size locations. This dark greyscale value was $grey_{baseline} = 0.22$. When all of this data was applied numerically to the January areas, it resulted in a surprisingly cross-day, cross-sample result of $\Delta grey_{40} = 0.012/mm$, or a 1.2 percent increase in greyscale whiteness for every additional mm of crystal size.

This constant increase strategy implies that a spatial size prediction within the area depends on:

- The dark, usually 0 mm crystal-producing $grey_{40}$ value called $grey_{baseline}$. 0 mm does not necessarily have to serve as the minimum, but did for both surface hoar formation periods in this paper.

- The value of $grey_{40}$ – the 40 pixel radius greyscale average – on the image at the point we wish to predict, i.e. $grey_{40}(lat, lon)$

- $\Delta grey_{40}$, which is the change in $grey_{40}$ value per expected mm of surface hoar size for the specific surface hoar layer and area. This part requires at least one (and, for these areas, only one) field measurement.

3.5.6 Model

A generalization of the constant increase relation gives our size prediction $size_{predict}$ at location (lat, lon) within the area:

$$size_{predict}(lat, lon) = \frac{(grey_{40}(lat, lon) - grey_{baseline})}{\Delta grey_{40}} \qquad (3.1)$$

Using the above model means that we may estimate the mean size of a surface hoar crystal at the coordinates (lat, lon) by using only terrain image data and two empirical scaling numbers for the area: $grey_{baseline}$ and $\Delta grey_{40}$.

For the case of Area I and Area II during the 16-22 January 2009 surface hoar cycle, Equation 3.1 can be rewritten as:

$$size_{jan\_16-22}(lat, lon) = \frac{(grey_{40}(lat, lon) - 0.22)}{0.012} \tag{3.2}$$

To develop such a specific equation for an area, one primarily needs to find $\Delta grey_{40}$ for that area. Since it equals the amount of positive change in $grey_{40}$ per expected mm growth of surface hoar, we need to obtain bounds on both the size and greyscale values for an area. Greyscale $grey_{40}$ minimum and maximum bounds for the area are $grey_{baseline}$ and $grey_{max}$, respectively. Surface hoar crystal size minimum and maximum values for the area are $size_{min}$ and $size_{max}$, respectively.

With these values, one may obtain the $\Delta grey_{40}$ for an area:

$$\Delta grey_{40} = \frac{grey_{max} - grey_{baseline}}{size_{max} - size_{min}} \tag{3.3}$$

Physically finding these values is likewise intuitable. For the greyscale values, $grey_{baseline}$ and $grey_{max}$ may be generally found via a histogram of $grey_{40}$ for an area. More accurately, they may be found in the image at the corresponding locations of the $size_{min}$ and $size_{max}$ field samples.

As for crystal sizes, field sampling may be done in the brightest $grey_{40}$ area of a 40^2 m^2 area to obtain $size_{max}$, and in a $grey_{baseline}$ dark area to obtain $size_{min}$. Even more generally, two samples of size and greyscale may probably be obtained from *any* two locations varying in size and $grey_{40}$ values; however, using the expected minimum and maximum values for an area allows the capture of the widest range and thus a potentially more accurate coverage of that range.

Our experience indicates that the January surface hoar formation cycle was typical for the general area, and it allowed a simplifying step. For all days and both areas, the minimum surface hoar crystal size $size_{min}$ was 0 mm. This can be confirmed both visually in the images in Figure 3.2, and by noting that, physically, trees emit longwave radiation and can block sky view enough to cause such 0 mm values.

This gives a more specific version of $\Delta grey_{40}$ which we use in this paper:

$$\Delta grey_{40} = \frac{grey_{max} - grey_{baseline}}{size_{max}} \tag{3.4}$$

Finally, when Equation 3.4 is combined with Equation 3.1, we obtain a more intuitive relation to the size obtained from an open area field sample:

$$size_{predict}(lat, lon) = size_{max} \frac{grey_{40}(lat, lon) - grey_{baseline}}{grey_{max} - grey_{baseline}} \tag{3.5}$$

As demonstrated by this relation, with this surface hoar formation pattern one needs minimal local knowledge to obtain the point sample of minimum or maximum mean size for a given surface hoar cycle. They may be found in most grey and least grey areas of a terrain image, respectively. Of course, one may already see instances where large scale processes would interfere with large scaling of this model. But as assessed in the next section, for our 40^2 m^2 sample areas with trees as the dominant factor, this concept held.

3.6 Results

We found that the model in Equation 3.2, when used to predict day-to-day surface hoar sizes for the 16 to 22 January surface hoar cycle, showed reasonable results. The standard deviation of error for the January size predictions are less than half of the mean crystal sizes, generally giving 2.5 mm average error on 7 mm average crystals. A more intuitive interpretation would be that if one were to think of categories of crystal size, e.g. *bigger* or *smaller*, the greyscale results usually predicted the same category for a given point. Still, the model performed better than the regression approach described above.

Statistical summaries about the residuals between predicted mean sizes and actual mean sizes for all days in the cycle can be seen in Table 3.1. The higher mean size with lower maximum size in Area II reveals its more generally open terrain than Area I. In other words, crystals grow larger in open areas, so the closer the mean is to the maximum size the more we may expect it to be an area with mostly open terrain.

One may note that the positive mean residual values for both areas in Table 3.1 imply that model estimates generally underestimate the real crystal size. This gives support to the implicit assumption of the model that for any greyscale value brighter than the baseline value, surface hoar is assumed to grow. The general underestimation of crystal sizes indicates that the model applies the assumption in a relatively conservative manner.

Crossplots for single days which have been spatially estimated using the model may be found in Figure 3.7. In Figure 3.7a, one can see a distinct widening trend in real-world size variance for the lighter greyscale values. This demonstrates that while trees account for much of the variance in the area, open areas with light greyscale values produce a range of real sizes not able to be predicted by this model. We expected this, as our model only accounts, via greyscale, for the spatial effects of trees.

We then confirmed this relation outside of the January formation cycle. During the surface hoar formation cycle in February 2009, we obtained an average single

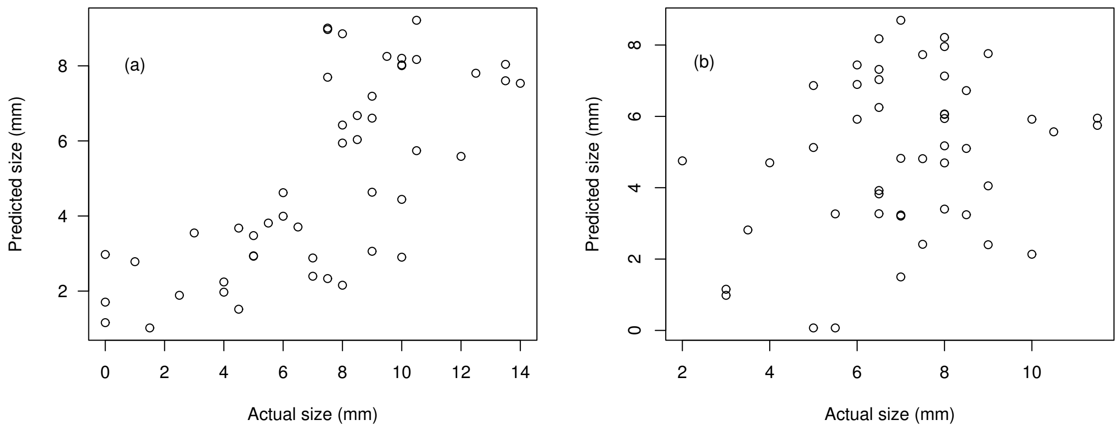


Figure 3.7: Crossplots of predicted versus actual surface hoar crystal size for both sample areas. (a) Area I. (b) Area II.

Table 3.1: Comparison of residuals for both sample areas. Residuals are the difference *(Actual Point Mean Crystal Size) - (Predicted Point Mean Crystal Size)*. Each sample area contains 144 actual point versus predicted point comparisons. Mean and maximum size values are from all observed sample points for each area.

	Residual mean (mm)	Residual std dev (mm)	Mean crystal size (mm)	Maximum crystal size (mm)
Area I	2.07	2.75	7.36	14.7
Area II	2.59	2.82	7.41	12.5

sample from an average open location and 47 additional points in the area, all using the same sampling method as in January. This new area was at a different location (150 km to the northwest), and a different aspect (south facing) to demonstrate the model holding across differing time, aspect, and location.

The single point contained surface hoar with mean size 5 mm, and that point on Google Earth had a $grey_{40}$ value of 0.65, much whiter than the images from Area I or Area II. Coincidentally, the area also had the same $grey_{baseline} = 0.22$, which we found via a histogram of greyscale values of the image. To obtain this $grey_{baseline}$, we took the average dark value of the lowest peak of dark $grey_{40}$ values in a detailed histogram.

We used the simplified $\Delta grey_{40}$ from Equation 3.4. As mentioned above in the

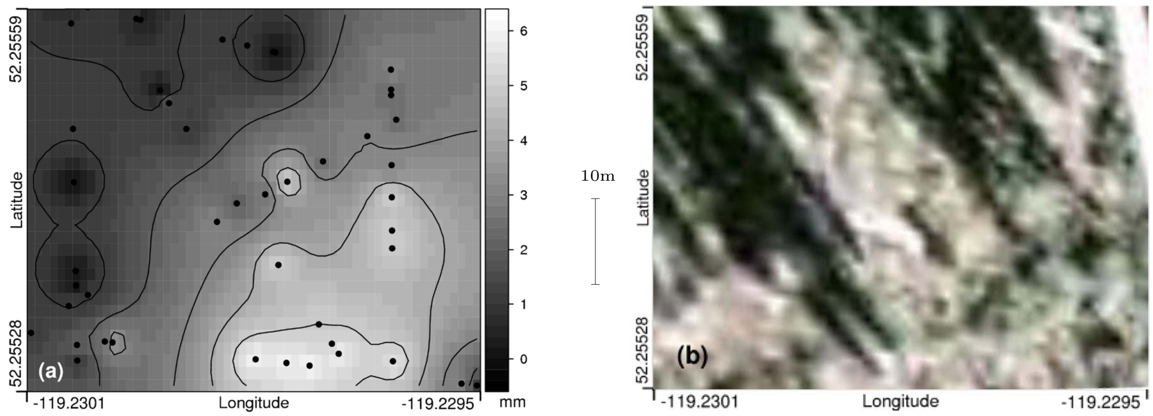


Figure 3.8: Area used for verification of the greyscale relation. (a) IDW map of measured mean crystal sizes in the new area sampled on 19 February, 2009. Black dots indicate the 47 actual physical measurement and model verification points. Contours have been added for visual clarity. (b) Google Earth image corresponding to the sample area used. Note the extreme slant in tree shading projection on this south aspect image. Imagery is provided under the Google Earth terms of use, copyright 2009 TeleAtlas, 2009 British Columbia, and 2007 Google Earth.

Model section, field methods would also ensure that the dark $grey_{baseline}$ value did, in fact, correspond to 0 mm so the model can be properly scaled.

This gives us a grey change per mm growth $\Delta grey_{40} = 0.086/mm$. From Equation 3.1, we adapted the following model to predict mean sizes for the surrounding area, for this new February layer:

$$size_{feb\_17-20}(lat, lon) = \frac{grey_{40}(lat, lon) - 0.22}{0.086} \tag{3.6}$$

Using Equation 3.6, we then predicted the sizes at each of the 47 physical points that we sampled, and we compared the predictions to the actual sampled values. The model predicted the correct size of surface hoar to within 1.5 mm for 60 percent of the points, and to within 2 mm for 70 percent of the points. As the average crystal size from those 47 points also equaled 5 mm, these results are similar to those of the January cycle.

The model produced a mean of absolute value residuals equal to 1.52 mm – meaning that, as before, actual crystal sizes were generally larger than predicted – and a standard deviation of actual residuals equal to 1.80 mm. Visually, the IDW map confirms the same shading effect; the Google Earth greyscale image and corresponding map are shown in Figure 3.8.

The similarity in accuracy between the model-building data from January and the predicted February values indicates that the constant increase strategy of using greyscale is extracting as much as can be used. No one variable can account for all of the variance in surface hoar size. But greyscale does rather well, especially considering its ease of use.

3.7 Discussion

When talking about millimeter-sized changes in surface hoar crystals, separating the amount of change caused by large scale processes versus those associated with small scale variance (which appears as stochastic variance at best when captured at a resolution of one sample per 100 m, for example) can be extraordinarily difficult, if not impossible (Hägeli and McClung, 2000; Campbell et al., 2004; Campbell and Jamieson, 2007).

This highlights the fundamental distinction between *measuring* a process which affects surface hoar formation, and *mapping* the result of that process over terrain. The latter, especially, may occur both as a descriptive and as a discovery method.

In this paper, we do not measure a process and then extrapolate it over terrain to create our model; rather, we focus only on mapping and prediction of surface hoar crystal size via Google Earth imagery greyscale values. And we focus on that only for specific instances: sparsely treed, relatively sheltered, and small areas.

Qualitatively, the basis of this model lies upon capturing a radius-of-effect for trees at a single point. Radiation balance, air temperatures, humidity, surface temperatures – the variation of all of these well-known physical quantities are exceptionally difficult to estimate spatially at these small scales. A physically successful method of spatial prediction may use these values in the future. However, we cannot yet quantitatively obtain these values on small spatial scales for non-instrumented areas, much less with the same ease and accuracy with which we can for greyscale.

A result of this study is the demonstration of a possibly useful strategy to find relations with spatial effects over terrain. This strategy, as outlined throughout the paper, has two parts. First, as instrumentation for huge numbers of spatial points can be impractical at best, one may select areas where the environment influences key variables. Sparse forests, though in and of themselves interesting as skiing areas, also represent the control of wind and relative consistency in the major formative process – skyview – from day to day. Second, though weather data may be most desirable for estimation purposes, it remains spatially elusive at small scales, and so using other augmenting sources of data such as greyscale can help estimate a large and useful amount of variance in crystal size.

Here, we demonstrate the accuracy of the model within a 40 m x 40 m area surrounding the single point sample. Though this produces estimates at the most useful – i.e. trigger – scale, larger scaling may be limited. Scaling the model into a significantly different and totally unsampled drainage or aspect would probably not only affect the relative shading in the terrain photos, but also introduce the effects of the more large-scale formation processes which have not been directly accounted for here.

For an illustrative example of these larger scale effects, examine the large grove of trees in the Google Earth photo for Area I (Figure 3.2). The cluster visually provides shade in the photos and appeared as a mechanism of storing and enhancing daytime heat in the left temperature map in Figure 3.5. However, the lack of trees in the photo on the slope immediately to the north in Area I does not help predict its convex north facing structure which so effectively allows the slope to continue cooling in daytime. Aspect has already been shown to have an effect on surface hoar size (Cooperstein, 2008), but aspect does not get captured by greyscale values. A bright greyscale area on a South slope may have the same value in our model as a bright greyscale area on a North slope, but the two areas will probably produce different crystal sizes. Sky view, though partially captured by shading in this study, ultimately depends on more factors than greyscale shading can provide.

As the number of trees per area increases, shading increases and this model will indicate small surface hoar growth, if any. As the number of trees per area decreases, the mean size of surface hoar will increase according to the model, and also by the intuitive concept of increased net outgoing longwave radiation.

However, with no trees or nearly no trees the effect of shading will become less applicable and other effects such as aspect (Cooperstein, 2008), moisture supply (Colbeck et al., 2008), and wind (Hachikubo and Akitaya, 1997; Feick et al., 2007) may become the main determining factors of spatial variance in crystal size. As discussed above, the notably larger actual surface hoar sizes in lighter greyscale valued areas – larger than our model could capture, that is – indicate that in areas with no large shading variance, other formative factors dominate.

Furthermore, when longwave-tree interactions are not the dominating factor within forests – say, a cloud covers the sparse forest in the evening, or very high winds push through the trees – entirely different conditions will result. This fact, true for all poorly understood natural processes, indicates that this model should not be used for multi-kilometer scale *modelling*, per se.

These open area formative factors oftentimes do not even scale within mountain ranges, much less across them (Hägeli and McClung, 2007). Thus, a general physical solution to the spatial prediction problem remains challenging and elusive.

Even non-formative sources of error such as photography resolution, angle of tress relative to the photography, accuracy of the GPS used to obtain the relation between $size_{max}$ and the greyscale value, and other factors can all contribute to how well the model works at any given location.

3.8 Summary and outlook

This work shows the small-scale spatial variability of surface hoar crystal size in sparse forests, and accounts for part of that variability using one factor: sky view. We developed a spatially predictive linear model which estimates surface hoar crystal size using averaged greyscale values in a dry land Google Earth image. Physically, this mechanism relates the size of surface hoar to its amount of sky view, or the amount of tree shading around that point within a 4 m radius. The relation held across different days and aspects in a January 2009 formation cycle, and also worked in an entirely different area, aspect, and formation cycle in February 2009. The results of the model generally do well at categorizing larger or smaller crystals, and had reasonable error for prediction using only one independent variable.

No previous work has done a two-dimensional spatial examination with this many surface hoar measurement points, much less developed a spatial model for surface hoar size that works at all, on any scale, in any conditions. As a field, we have yet to even set eyes upon extensively measured spatial variance of surface hoar and understand it qualitatively. So, though the highest goal would be to have a model that works everywhere all the time, less than 30 percent average error for a new cycle and area represents a fairly reasonable start.

Independently of predictive uses, our data shows the previously unknown extreme crystal size variance to be found at these 40 m scales. The greyscale relation shows the radius of effect of trees when they are the dominating factor, and it represents the first focus on the *spatial* surface hoar formation variable of sky view.

Also, this work demonstrates that surface hoar formation research need not necessarily be limited to single points near accurate weather stations. Other augmenting information may be found and coaxed into helping find trends and produce estimates in areas and at scales where precise weather data at each point would be impossible to obtain.

More generally, this study shows the extreme spatial variance of surface hoar sizes across single treed areas. Such variance within these small 40 x 40 m areas clearly shows the limitation of trying to project single point estimates over a very large area. Said another way: We cannot usefully spatially map surface hoar formation for skiers over large scales before being able to do so over small ones. With a general prediction of "5 mm mean crystal size" for a 40 x 40 m area, crystals will still range from 0 mm to 7 mm, or larger. For the January cycle, "8 mm mean crystal size" could mean a range from 0 to 14 mm, all within a small area. This could correspond to triggering ranging from very likely to very unlikely, within the same area described with a single mean. Measuring and noting small terrain effects, *and* how they actually fit within the spatial variance of a larger area seems more useful than simply striving for a larger scale mean size prediction. One can see the use in mapping and understanding – even one by one – the spatial factors that create such variance.

Our approach here limits wide applicability. But, the results of such an approach can still provide important uses, as this study does for understanding sky view for skiers in trees. And, such an approach – with a semi-controlled environment and additional augmenting data – maximizes the potential to find a useful relation that may be further built upon later. In the same way that a handful of factors may be identified and measured for one single point of surface hoar growth, finding conditions and areas that allow identification and measurement of only a few *spatial* factors at once may be key to pushing our physical and single-point surface hoar knowledge into the spatial realm.

3.9 Acknowledgements

For financial support, the authors gratefully thank the Natural Sciences and Engineering Research Council of Canada, HeliCat Canada, the Canadian Avalanche Association, Mike Wiegele Helicopter Skiing, Teck Mining Company, Canada West Ski Areas Association, the Association of Canadian Mountain Guides, Backcountry Lodges of British Columbia, and the Canadian Ski Guides Association.

For logistical support we thank the Avalanche Control Section of Glacier National Park and Mike Wiegele Helicopter Skiing. For careful field work, we thank Thomas Exner, Mike Smith, Cameron Ross, Ali Haeri, Mark Kolasinski and Jordan Stiefvater.

Chapter 4: Operational GIS Model Development

Like many fields, avalanche studies can produce substantial model-generated data, but it often lacks accessible and fast ways of displaying and analyzing such data. This paper is a modification of Shea and Jamieson (2010a), which appeared in the proceedings of a conference in 2010.

The main contribution of this manuscript is a clear and accessible presentation of GIS model data for use by non-GIS-experts in daily avalanche forecasting. The existing model SWarm and the corresponding zenith calculations are not new, only modified for this purpose.

For more information on operational use, see Section 6.3.

4.1 Abstract

Using mathematical models to predict and visualize snow processes over terrain with a Geographic Information System (GIS) takes a specific set of skills and significant computer processing power and time. These factors are often at odds with how such a model would be used for, say, daily avalanche forecasting. We present GSWarm, a simple and fast GIS-based model that uses the near-surface snow warming statistical and empirical model SWarm as a basis. GSWarm resulted from (a) published user comments on existing snow and avalanche computer tools, (b) published graphic design principles, and (c) direct forecaster feedback. Using GSWarm as an example, we present key ideas used to provide a simple interface to a GIS model, including: (1) Calculating many possible scenarios ahead of time, so hypothesis testing of different weather and snow conditions can be done quickly. (2) Allowing small previews of many results to be seen on one screen, for selection of specific conditions without

using input boxes. (3) Providing scaling and visualization help to the user rather than giving a single final result. These ideas represent a unique perspective on snow and avalanche computer model design.

4.2 Introduction

As a researcher, it often seems like the design of a new model for snow and avalanche forecasting follows the path on the left side of Figure 4.1: develop a model – physical, empirical, etc. – to obtain a satisfactorily useful result. And this rightly so, for the main scientific difficulty often lies in modelling the workings of the natural world rather than the needs of the user.

However, the *use* of the model, once a useful physical relation is identified, depends most primarily on two things: time, and ease. Once launched into the world of daily use, all steps on the right-hand path in Figure 4.1 add to the time needed to use a model. And ease of use does not necessarily mean providing one clear result. A desired result can also be an ensemble-type comparison. Here are some pertinent quotes from other papers which designed computer systems models for snow and avalanche forecasting:

....Our primary goal is to create a tool to visualize, explore, and ask questions of weather and avalanche data sets, thereby allowing us to find spatial patterns and facilitate hypotheses generation. (McCollister et al., 2003)

....Neither of these programs has been widely adopted amongst veteran forecasters because they require substantial time and effort to operate, as well as vast regional backlogs of weather and avalanche data. (Cookler and Orton, 2004)

....The critical point here is that our model of the backcountry forecasting process

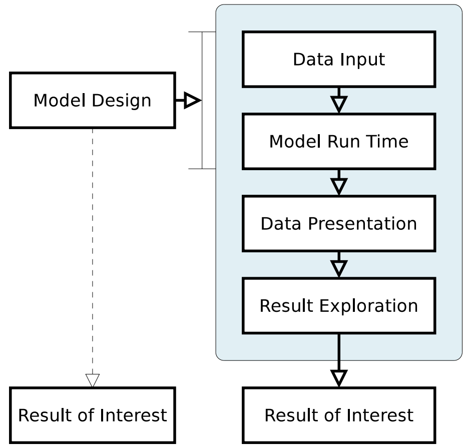


Figure 4.1: Two model development paths: the idealized path on the left, the reality on the right. This paper examines those components within the shaded box.

is one primarily based on hypothesis testing. Thus, the role of the model is not to provide the observer with the avalanche hazard for the following day or to identify the probability of avalanches. Rather, it is another part of the information gathering and hypothesis testing process....(Purves et al., 2002)

We realized that, although the mathematical models developed by our research group such as SWarm (Bakermans and Jamieson, 2009) and SAWLEM (Zeidler et al., 2006) enjoy use by practitioners, the amount of use is not necessarily what we expected given that a wider audience has expressed interest in the subjects. And so the question – and challenge – seemed to be whether a mathematical model could be made accessible to additional audiences.

To do this, we adapted the near surface warming model called SWarm (Bakermans and Jamieson, 2009) to another presentation medium: real-world digital terrain data. Many complications and difficult decisions arose from this development path, essentially at every stage of the right-hand path in Figure 4.1 except the final result. This paper outlines our process and the resulting system: a GIS-based implementation of SWarm called GSWarm. The results of this paper *are* the methods – the methods used to adapt a mathematical model continuously over terrain – and so we present previous work, the methods used, a few comments on accuracy, and conclusions.

4.3 Previous Work

This section describes previous work done in computer forecast systems. It also describes the SWarm model, upon which GSWarm is based.

4.3.1 Other systems

Computers play a diverse and somewhat unclear role in avalanche forecasting. Existing systems show almost as many design philosophies as systems themselves.

There are programs which provide a forecast answer for the day such as a danger rating (Merindol et al., 2002; Giraud et al., 2002; Floyer and McClung, 2002; Zeidler and Jamieson, 2004; Cordy et al., 2009; Schirmer et al., 2010). Other systems provide insight into physical snow conditions without providing a direct forecasting answer, and these include SNOWPACK (Bartelt and Lehning, 2002), SWarm (Bakermans and Jamieson, 2009), SAWLEM (Zeidler et al., 2006), and CROCUS and SAFRAN (Giraud et al., 2002), among many more.

Still others organize data – such as weather information, snowpack information, or history data for an avalanche path – into a recognizable and readable format. These include Cornice (Purves et al., 2002), and GeoWax (McCollister et al., 2003). This category also includes *spatial notepads*, which allow recording of snowpack and weather conditions over terrain and time. Although spatial notepads exist (Brabec et al., 2001; Canadian Avalanche Association, 1991), others have been called for, such as for surface hoar layers (Davis, 2010).

Some models fill a category in a specific culture, locale, or regime well. These include the Mammoth Mountain binary decision trees (Rosenthal et al., 2002) and the local-expert-weighted system in Gassner and Brabec (2002), among others. Few systems are different adaptations of existing models with new perspectives. One such system is a cellular automata model developed by Kronholm and Birkeland (2005) designed to show the conditions that create large avalanches.

Hence, not all models are mathematical, or even adaptable to terrain. Some mathematical models that predict snow conditions do run spatially at small scales, including Adams et al. (2010), which runs on the slope scale. Some data input libraries allow point data – a common type of data in snow and avalanche forecasting – to be intelligently interpolated across a map (Bavay et al., 2010). However, to our knowledge as authors, no other snow and avalanche-type forecasting systems currently run a mathematical and physically predictive model as a continuous (non-interpolated) field over real terrain larger than a slope.

Mathematical models, computers, and GIS use have gained a stronger foothold in non-forecasting applications such as predicting avalanche runout, impact, and designing defences and structures for avalanche terrain. Other, similar applications such as glacier mass balance calculations and terrain radiation models run over larger areas. However, these applications are often not nearly as time limited as avalanche forecasting.

The diversity of such models should show at the outset that the methods here for adapting SWarm into GSWarm cannot directly be applied to every one. However, the process of designing GSWarm should show that a new presentation of a model, even one as beneficially simple in design as SWarm, can be quite complex.

4.3.2 SWarm

GSWarm needs to be distinguished from SWarm, as they do not compete. SWarm is an empirical model – a mathematical description of field research results.

SWarm as most *users* think of it is a spreadsheet which allows the user to obtain expected warming down 10 cm in the snowpack based on latitude, longitude, date, current cloud cover, and days since snowfall. The user of the spreadsheet is then presented with expected warming for different slopes and aspects, and the data are presented on a familiar aspect-incline rose, on a Cartesian graph, and in a table. Figure 4.2a shows a screenshot of SWarm in use.

SWarm takes specific input and presents tangible numbers, but only at a single point. GSWarm does not give numbers as a result, instead it presents *warming maps* which display expected warming by using colour. One such map and its key are provided in Figure 4.2b.

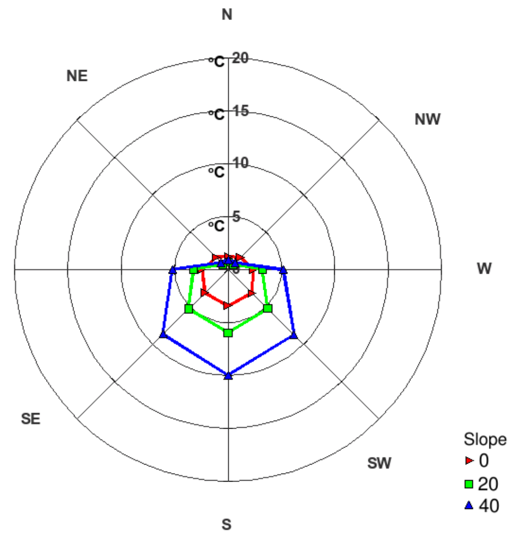
4.4 Methods

With the help of preliminary releases and user feedback, the goals of GSWarm became clear:

- Primarily visual rather than numeric results
- Minimal time and computer skills needed by user
- Easy to compare results from different inputs
- Free to the end user
- More spatial information, such as terrain shading

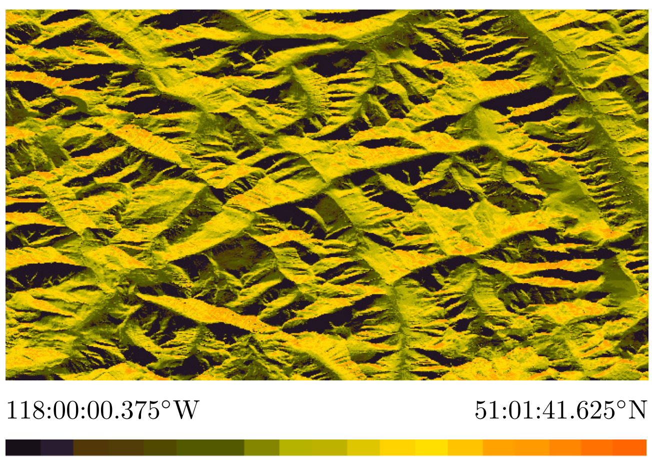
We walk through the adaptations in order of the steps shown on the right-hand

(a)



(b)

51:24:33.375°N 117:20:14.625°W



118:00:00.375°W 51:01:41.625°N

0°C 10°C

Figure 4.2: Output of expected warming 10 cm down from (a) SWarm and (b) GSWarm for a single day (October 29), albedo value (one day of no snowfall), and cloud cover value (4/8). A full size GSWarm warming map is 2272 x 1306 pixels for the Rogers Pass area. The legend in SWarm (a) has been moved and scaled to fit. In addition to the aspect-incline rose shown here in (a), SWarm also provides a table and Cartesian coordinate graph, both also containing expected warming listed by slope and aspect.

side of Figure 4.1. This is but one possible streamlining process of many, but without such a process the spatial implementation of SWarm would not have been possible.

4.4.1 Data input

Data input can be time consuming. Imagine a model that returns a point result value in time and space. It has a few fields of data that, in total, take 15 s to fill in. For a point model, examining all the different aspects would take one minute. Examining all aspects and all elevation bands would take three minutes. Examining all of that data and comparing it against the same data for the two days prior would take nine minutes. In a thirty-minute morning meeting, this can be an excessive time requirement. And that is just the data input, not anything else such as running the model or looking at the results.

Of course, this example shows poor design. Better design is straightforward, for example, SWarm returns results for nearly all aspect and slope values at once. And, it returns those results together on a graph for easy comparison.

To save time, GSWarm input involves clicking a mouse, not typing. For reasons described in Section 4.4.2 below, GSWarm was developed to run in pre-determined forecast areas. From a GSWarm area, a single click brings up all of the modelled day's data in one display. This display may be seen in Figure 4.3. The data displayed are maps of forecasted sub-surface warming for different possible conditions over the area for one day.

The data output style for GSWarm (Figure 4.3) is based on a theory of *small multiples* (Tufte, 1990). Small multiples allow a user to see a visual ensemble collection of data in one high-level view. In the case of GSWarm, the data ensemble is presented over a grid of image maps. The two axes of the image map grid are: days since snowfall, and fraction of cloud cover. Days since snowfall is then used to obtain an estimate of the albedo of the snow surface.

These variables are of course much simplified – rarely is there one value for cloud

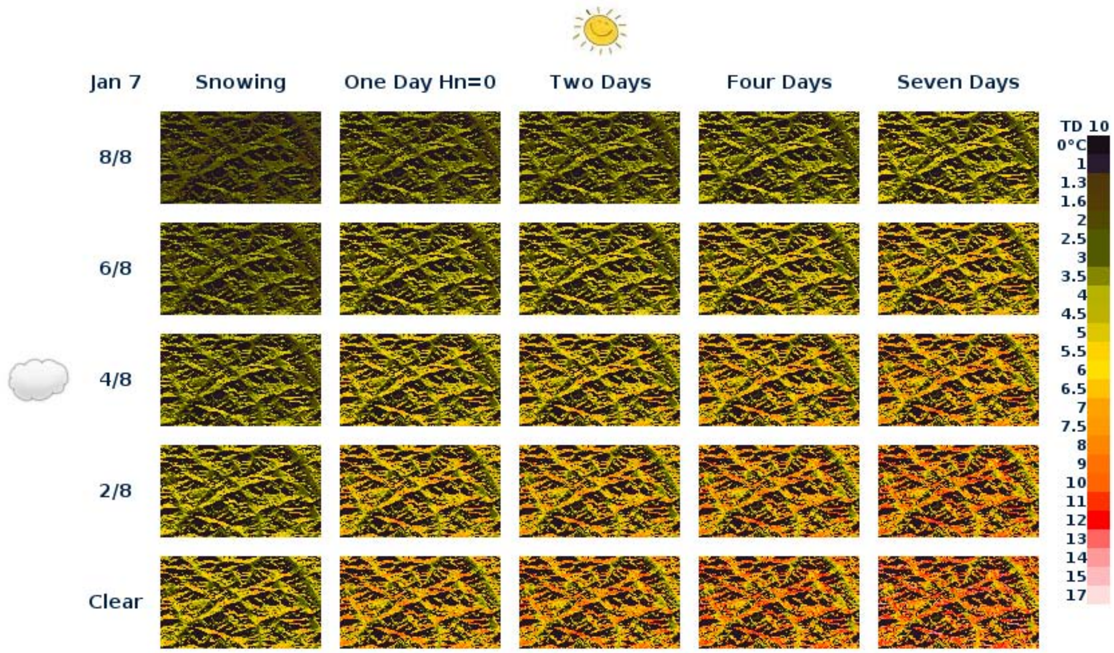


Figure 4.3: GSWarm's one day view. The preview images show a range of conditions spanning the least (top left) and most (bottom right) warming that one could reasonably expect. Clicking on previews allows the user to download images with more detail, i.e. *drilling down*. All images in this paper are best viewed in colour, either in the electronic proceedings or in the online model.

cover over a day or area, and rarely is there one value for albedo over an area – but their simplifications have been designed into the SWarm model to maintain reasonable accuracy while greatly increasing the ease of use (Bakermans and Jamieson, 2009). This, combined with the small multiples design principle, allows a visual overview of the warming ranges possible *without* any further input.

Thus, one click to select the area shows the small multiples for one day. A second click in the menu above the small multiples (not shown in Figure 4.3) brings up the small multiples for a different day. Clicking on a small map brings up a larger version, and clicking on the larger version brings up a full scale image warming map of that area. Including an intermediary sized image – and requiring a second click – seemed needed as the small image maps are quite small, and the large ones are a few MB in size.

4.4.2 Model run time

Model run time proved to be the most challenging part of the adaptation. To honour the goal of making the product free to end users, we chose to develop the image maps in GRASS, a free and open source Linux-based GIS system. GRASS is a stable and mature system that lends itself to research and has established documentation both online and in print form (Neteler and Mitasova, 2008). When adapting SWarm to terrain, the output method had not yet been finalized, and GRASS enabled us to keep the option open to write a free GIS-based program that would run on the computer owned by the user.

GRASS contains the shortwave modelling program `r.sun` (Hofierka, 1997) which allows for beam (direct), diffuse, and reflected shortwave radiation calculations over terrain. These calculations include terrain shading. The program has an option to minimize memory use by using pre-calculated horizon shading maps.

GSWarm essentially calculates the incoming beam and diffuse shortwave at every point over an area by using `r.sun` and calculated terrain shading maps, and then uses the SWarm model method outlined in Bakermans and Jamieson (2009) to modify the shortwave value into an expected warming down 10 cm in the snowpack. This raster-type calculation extends the SWarm model easily from one point to real terrain. There are a few modifications, as outlined below in Section 4.4.5.

Ideally, a GIS-based warming model would return a map of expected warming for arbitrary conditions, an arbitrary location, and an arbitrary date. However, SWarm uses the maximum shortwave input in a day, and hourly estimates of incoming shortwave must be made in order to find that maximum value. To perform twelve of these calculations on 3181 x 1829 pixels (the size of our first area of interest) and produce one day of warming maps on a server-grade machine takes 45 minutes of processing time and 1.3 GB of computational hard drive cache space. Performing these calculations on a personal computer takes between five and twenty hours.

Given this time and computing power requirement, it made more sense to limit GSWarm to pre-set forecasting areas of high interest so many calculations (such as terrain shading maps) could be calculated ahead of time. The first region we selected was that surrounding Rogers Pass, British Columbia, Canada, from 51:01:41.625° north to 51:24:33.375° north, and 118:00:00.375° west to 117:20:14.625° west in NAD83 with a resolution of 00:00:00.75 degrees (about 40 m) per pixel. Figures 4.2b and 4.4 show different warming maps with geographic extent labels. The Rogers Pass area is shown in all warming maps in this paper; however, other areas have subsequently been requested and modelled as well.

Although this pre-calculation solved many of the time challenges, it introduced a new data space challenge. Each full warming map requires 4 MB of space, and for the nine different cloud cover values (0/8–8/8) and eight different albedo values (0–7 days since snowfall) offered by SWarm, this produces 63 images and over 250 MB of images per day.

As the SWarm model depends on the Julian day for the date input, one year could be calculated and used for every year. But with eight months of interest (October through May) and around thirty days per month, a full image solution for just Rogers Pass would take up 60 GB of space. This would not be unreasonable for one area, but as we anticipated having many areas, this needed to be reduced.

Rather than nine different cloud cover values as SWarm offers, we display five (0, 2, 4, 6, and 8/8). Rather than eight different albedo values, we use five: snowing conditions and 1, 2, 4 and 7 days since snowfall. And rather than calculating every day of the winter, we calculate one day every two weeks October through February, and every week March through May. The weekly spacing in the spring captures both the quickly changing solar declination, and the return of declination to north of the Equator. The spacing of the cloud cover and albedo values show steps with a similar result spacing in output.

The pre-calculation method reduces the model run time down to only the time needed to download the image maps. The reduction of cloud cover and days since snowfall values not only reduces the required space but also makes the small multiples display described above (Section 4.4.1) easier to comprehend at a glance while still covering the physical value extremes.

4.4.3 Data presentation

The *small multiples* method described in Section 4.4.1 is a central part of the data presentation. A visual presentation that is clear, graphical, and intuitively efficient lends itself to the needs of avalanche forecasting (Atkins, 1992). However, it is not a complete answer. The presentation includes everything from the programs it uses, to the operating systems it runs on, to the colours used in the maps, to the feel of the interface. The presentation, in turn, directly affects the time needed to run the program.

To solve the platform problem, GSWarm was implemented as a web page, meaning it can run on any platform that can run Firefox or Chrome. The image maps are stored remotely to keep any updates to the program centralized and to save space on a user's computer; however, GSWarm could be run locally if necessary.

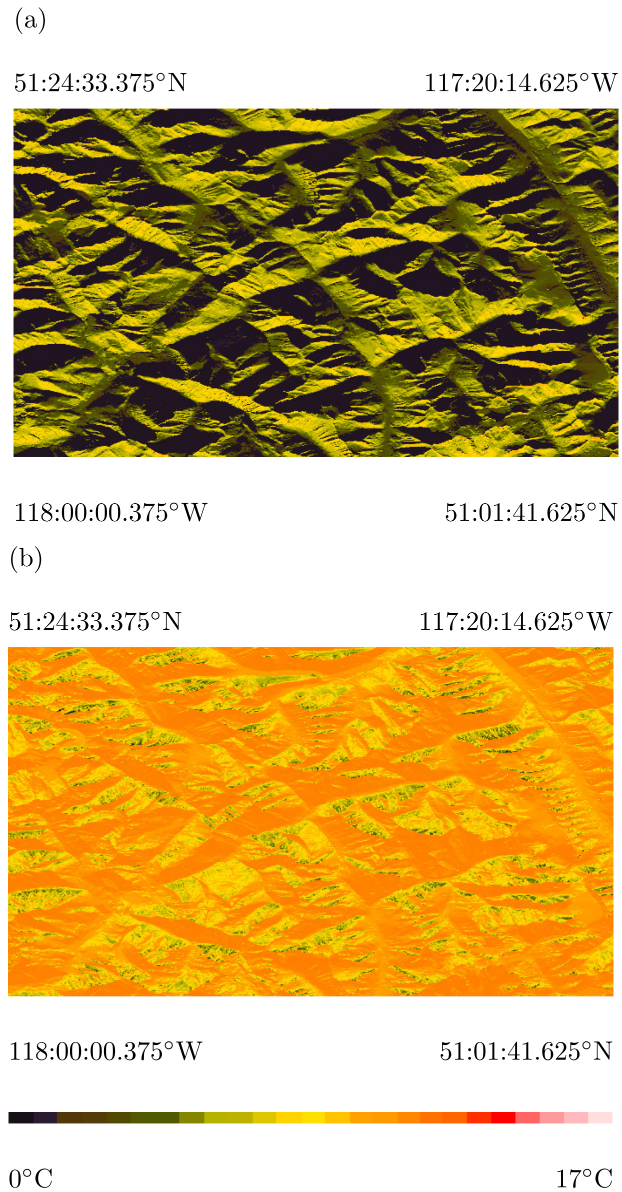


Figure 4.4: GSWarm warming maps from a clear day with fresh snow from (a) December 24 and (b) May 27, along with their common colour ramp.

One of the most important qualities of the image maps is the colours they use to represent temperature. Tufte (1990) emphasizes how colouring on any map should be clear and intuitive. However, what this meant for GSWarm took some experimentation. If the primary goal of GSWarm was to have single, easy-to-understand maps, a many-coloured and fully used colour ramp would be most desirable as it displays the most detail per single map.

However, we felt the real purpose of GSWarm is to allow direct comparison between days, months, albedo values, and cloud cover values. This meant that the GSWarm maps needed one single colour ramp common to all image maps in the entire program. Eventually, we developed a non-linear colour ramp which splits up the small warming values from 0 to 6 °C of expected warming into as many different colours as 6.5 to 17 °C of expected warming. This allows a question of, say, *how much more does terrain create shade in December versus May* to be easily answered with glance, as shown in Figure 4.4.

Finally, the presentation needed a helpful way for the user to landmark different places within the map. Many of the terrain features in an image map with significant terrain shading (such as Figure 4.4a) can be picked out, but a day without such terrain shading (such as Figure 4.4b) can be difficult to landmark in. Thus, shaded terrain overlays – produced with the traditional western location of the sun – were added that a user can turn on and off to visually find features within an image map. An example of a shaded overlay may be seen in Figure 4.5.

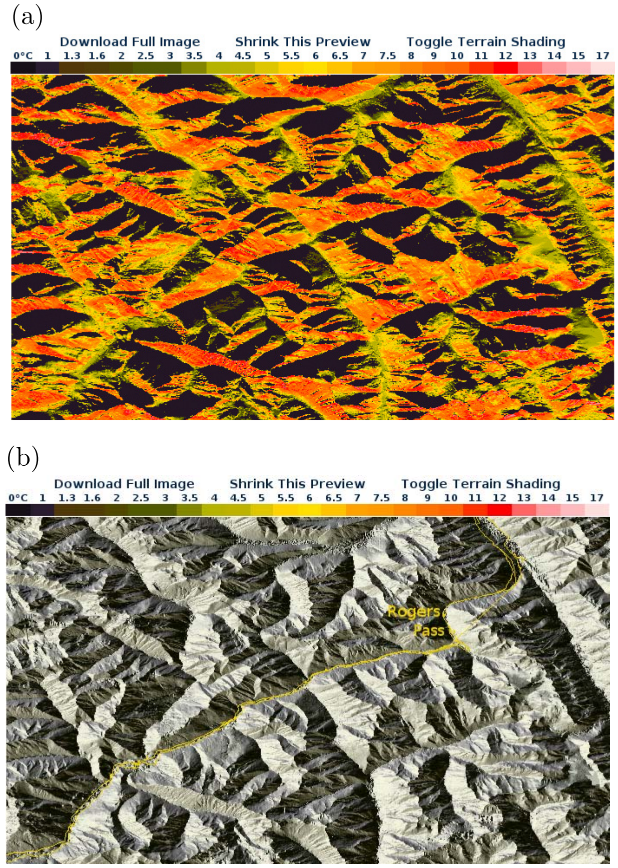


Figure 4.5: GSWarm views from 7 January after seven clear days and current 4/8 cloud. (a) Warming map screenshot, (b) Terrain overlay of the same area with shaded relief and landmarks. Rogers Pass is located one-third diagonally inwards from the upper right. Map width is approximately 130 km. Note that only (a) shows the expected warming at each pixel location; the overlay may be displayed with different opacities over the maps in GSWarm, and this changes the intensity and hue of the map output colours. For latitude and longitude geographical extent, refer to Fig. 4.2b.

4.4.4 Hypothesis testing

By the quotes presented in the introduction, the ease of exploration within a range of results seems to be a desired model trait. This type of exploration is also known as hypothesis testing, which requires a good method to present the data, as already discussed. Ease of hypothesis testing also depends on data input and model run time – if either are too long or unwieldy, obtaining multiple data to explore becomes difficult, and the model will not be used by the forecaster.

Yet, there is an aspect of hypothesis testing which goes beyond the clean presentation already discussed. Not only should results show quantitative data, but they should assist in understanding the *change* in data for small variation in conditions. For example, the ease of obtaining the data from last week for comparison allows one to intuit: *This is what it felt like last week, and this is the model output from last week, so if this is what the model says for this week, this may be what it will feel like this week.*

GSWarm shows these changes clearly. A day in December and a day in May appear quite different in their warming maps – for an example, see Fig. 4.4 – but two days in January (even two weeks apart) look quite similar. This aspect is exactly what we as designers desired out of the GSWarm model. Variables such as the set of completely shaded areas at a particular solar zenith are near-impossible for the human mind to calculate without aid, but easy for a computer to calculate. Furthermore, comparing these physical changes across many different days enables the human mind to do what it does best: extrapolate patterns and suss out the variables that matter most for decision making.

This is but one example where decisions in one design area can lead to complications in another. One solution to the data input problem is to automatically read data from a weather station into a model. But then, the creation of an ensemble-type result presentation – the ability to test hypotheses from this input data – is possible but must be done much differently than outlined here.

4.4.5 Comments on model accuracy

Although the purpose of this paper is to show the development method behind GSWarm, as it models a physical quality its accuracy should be discussed. The average error in the SWarm model was 1.6 °C over the development dataset. GSWarm, being a re-implementation of an existing model, has not been subsequently validated.

When using the published SWarm model – e.g. the shortwave modification coefficient 0.00542 from Bakermans and Jamieson (2009) – GSWarm overestimates the warming. This is because the GRASS `r.sun` program estimates physical shortwave input using specific physical parameters such as Linke turbidity and albedo in addition to the solar zenith (Hofierka, 1997). Hence, GSWarm uses the SWarm coefficient for physically measured maximum shortwave radiation input: 0.00448, also from Bakermans and Jamieson (2009).

In addition, GSWarm uses re-designed aspect and slope corrections which follow local time and apply to the entire estimated beam (direct) shortwave input at each raster pixel. These are very similar to those used in SWarm but do not include leap years:

```
decl = (23.45 * sin((day+284) * 360/365))
hour = 360/24 * (12-localtime)
azimuth = atan((-(cos(hour)) * cos(decl) * sin(lat)) +
    (cos(lat) * sin(decl)),
        sin(hour) * cos(decl))
solarelev = asin((sin(lat)*sin(decl)) +
    (cos(lat) * cos(decl) * cos(hour)))
modifier = ((sin(slope) * cos(solarelev) *
    cos(azimuth - aspect))
```

```
        + (cos(slope) * sin(solarelev)))
correctedSW = directSW / sin(solarelev) * modifier
```

where `day` is the Julian day, `decl` is the solar declination for that day, `localtime` is the local solar time (0-24 hours), `hour` is the hour angle, `azimuth` is the solar azimuth, `solarelev` is the solar elevation, and `slope`, `aspect`, and `lat` are the slope, aspect (azimuth type clockwise from north), and degrees of latitude, respectively. This gives shortwave corrected for slope and aspect (`correctedSW`) from a value of direct beam shortwave (`directSW`). These calculations are not the most precise available, but are a good compromise between computational resources and accuracy. Possible improvements include, as mentioned, correcting for leap years, and a more accurate declination; these are discussed in Robinson (1966) and other meteorology texts.

These and other changes create deviations from the SWarm values. Other changes include a lower Linke turbidity appropriate for mountainous terrain in the winter, use of albedo in diffuse shortwave calculations, and terrain shading.

Due to these differences, it is difficult to estimate an overall percentage deviation from SWarm values. GSWarm tends to estimate more warming despite using the physical shortwave coefficient as mentioned above. This primarily may be due to GSWarm using a low Linke turbidity, as appropriate for mountainous terrain, but which would be difficult to incorporate in a single point for SWarm. However, this generalization of GSWarm predicting more warming is only a tendency – for example, SWarm will estimate extreme warming in areas that GSWarm will predict very little due to terrain shading. This prevents direct numerical comparison.

As SWarm has also not been independently validated, the absolute numeric value of GSWarm output likewise remains unvalidated. However, GSWarm uses physically-based and established concepts (e.g. the `r.sun` routine, horizon shading maps, and solar positions) which give weight to its use as a visual comparative method over different conditions and time.

GSWarm is publicly available and may be found at:
http://www.ucalgary.ca/asarc/gswarm

4.5 Discussion and summary

GSWarm and SWarm, although using essentially the same model, answer different types of questions. SWarm answers definite questions easily, such as: *What is the average warming on north slopes today?* GSWarm cannot easily answer these definite, numerical questions, but it can, at a glance, give visual answers for questions such as: *Which slopes along a route will warm substantially today?*

The development of GSWarm – in sum, a project to make possible the timely and intuitive presentation of complex data – depended on all decisions along the right-hand side of Figure 4.1. With a poor user input method, or too much run time needed over one area, or a difficult method of comparing results, GSWarm may not have been fit for operational forecasting, or other applications including education.

The decisions we made can certainly be improved upon; however, few papers in the avalanche field discuss the integration between the needs of the user – i.e. forecasters, recreationalists, guides, teachers, students, etc. – and the design of the system, despite the importance of the topic. Furthermore, these methods are often very specific to the resources available. For example, in meteorology, computers, analytical models, and physics have all established themselves as indispensable to both forecasting and general understanding. Many of the model run time problems have been solved in the meteorological area by using large supercomputers, which few – if any – avalanche forecasting operations have access to.

GSWarm fills a niche for users who think spatially, value visual images rather than numbers, want to skim over a lot of data quickly and process it intuitively rather than quantitatively, and have only a small amount of time to do all of this. Quite a few users felt this niche needed filling.

Other users felt that SWarm was just fine as a spreadsheet – why change it? And this variety is fine. GSWarm now provides terrain-based visual output; SWarm can continue to provide tangible numbers at well-defined points for those who prefer them.

Considering the user from the beginning can often make the difference between use or disuse for a computer forecasting system. This is not to say that GSWarm perfects this process; far from it. Rather, GSWarm will hopefully receive sufficient use to, in turn, provide discussion which will lead to an even differently designed presentation. Note that the key word here is *different* and not *better*: Different audiences simply require different presentations for communication to be effective. And, addressing each audience carefully and individually will make our tools that much more useful.

4.6 Acknowledgements

For feedback and direction we thank the Avalanche Control Section of Glacier National Park and the forecasters of the Canadian Avalanche Centre, including Greg Johnson and Ilya Storm. Cameron Ross also gave significant helpful user feedback.

For support, we thank the Natural Sciences and Engineering Research Council of Canada, HeliCat Canada, Parks Canada, the Canadian Avalanche Association, Mike Wiegele Helicopter Skiing, Teck Mining Company, Canada West Ski Areas Association, the Association of Canadian Mountain Guides, Backcountry Lodges of British Columbia, and the Canadian Ski Guides Association.

Chapter 5: Snow Surface Thermography

Although snow metamorphism occurs at a small scale, measuring the temperatures that drive metamorphism both on the surface and within the snowpack has been a tricky and coarse endeavor. The method presented here of using thermal imaging should greatly aid such study.

This chapter has been slightly modified from its original paper, entitled *Some fundamentals of handheld snow surface thermography*. It appeared as Shea and Jamieson (2011) in the journal The Cryosphere.

5.1 Abstract

This paper presents the concepts needed to perform snow surface thermography with a modern thermal imager. Snow-specific issues in the 7.5 to 13 μm spectrum such as ice emissivity, photographic angle, operator heating, and others receive detailed review and discussion. To illustrate the usefulness of this measurement technique, various applications are presented. These include detecting spatial temperature variation on snow pit walls and measuring the dependence of heat conduction on grain type.

5.2 Introduction

Many processes depend on thermal exchange at the surface of snow. For avalanche prediction, the persistent weak layers of surface hoar (Hachikubo and Akitaya, 1997) and near-surface facets (Morstad et al., 2007) form due to sustained thermal gradient. For hydrology, shortwave albedo – and thus radiation balance – changes as a direct result of surface grain type (Armstrong and Brun, 2008, pg 55), which follows from snow surface metamorphism via grain temperature and temperature gradient (McClung and Schaerer, 2006, pg 66).

In contrast to traditional contact thermometers, such as thermistors, which obtain a temperature by becoming the same temperature as their measured subject, remote thermal sensing provides a way to measure the thermal infrared emissions of a material surface without contact. By using a radiation sensor in the longwave range, and correcting for interference and the longwave emissivity of a material, the temperature of the material may be obtained. This idea has existed for quite some time; the first patent for an infrared thermometer was granted in 1899, and the Landsat satellite launch in 1978 included a spatial sensor in the thermal range (Lillesand et al., 2008).

Space-based thermal sensing can provide a spatial image, but it cannot provide a high spatial resolution measurement like a handheld infrared point thermometer can. Thermography, also known as infrared thermal photography, bridges this gap. This paper discusses sensing of snow via handheld thermography. Scales include distances from a few kilometers to less than one meter. The bandwidth varies by equipment; this paper discusses the thermal infrared spectrum of 7.5 to 13 μm, a common spectral span for handheld thermal imagers.

Hand-held thermography, in contrast to space-based sensing, brings a host of new concerns as well as new abilities. This paper demonstrates how this technique has been used already and how it may be used in the future. Space-based thermal sensing has received attention in the previous work discussed throughout this paper. To date, no other paper has provided a specific presentation and discussion of use of snow surface thermography with these spatial resolutions and distances, despite it offering a variety of research possibilities. Thus, this paper attempts to provide both a physical and practical basis for using handheld thermography for sensing snow.

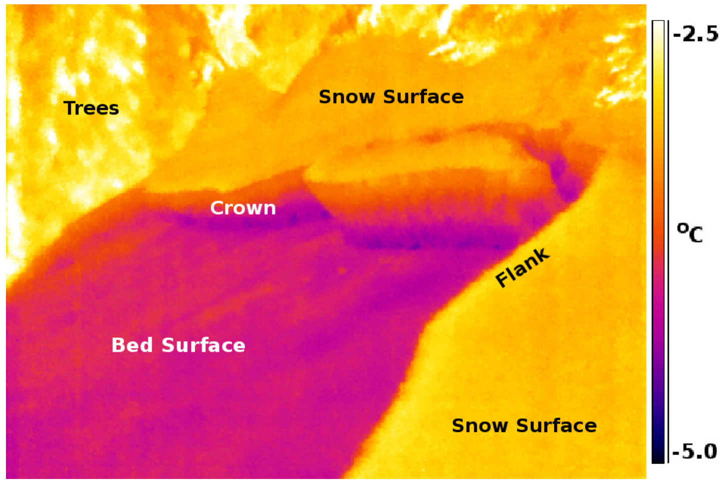


Figure 5.1: Bed surface and crown of a slab avalanche ($\sim$ 40 cm crown), approximately 5 minutes after release.

5.3 Motivation

The surface of snow can heat and cool at an astonishingly fast rate. As shown in the *Applications* section below, snow may be heated by an external heat source to 14 °C above the ambient temperature, and portions of the surface can then re-equalize with the ambient air, decreasing by 14 °C, within minutes.

The speed with which thermographs may be taken can record and measure many snow processes. Real-time video can capture spatial changes continuously, and even single thermographs can capture data that otherwise would be impossible to obtain point-by-point. Figure 5.1 shows but one example: a still-cold avalanche bed surface and warm crown fracture, minutes after the slab avalanche occurred. The measurement of the temperature below the entire avalanche failure layer at once would be impossible without handheld thermography.

Thermography not only provides a different view of known effects, but some applications – such as an unexpected instance of an inverse relation between conduction speed and snow density described in Sect. 5.8.2 – reveal situations which are currently neither well understood nor well documented. Thermography provides new ways to measure these processes.

5.4 Basic theory

A thermal imager measures the amount of thermal-band radiation that reaches its sensor. All materials emit radiation in a way that can be predicted ideally by the Stefan-Boltzmann law. The Stefan-Boltzmann law states that emissivity ϵ, temperature T in K, and watts output per square meter P are related by:

$$P = \epsilon\,\sigma\,T^4 \tag{5.1}$$

where σ is the Stefan-Boltzmann constant (5.67×10^{-8} J m^{-2} s^{-1} K^{-4}). The emissivity ϵ is material-specific and wavelength-specific, and it always lies between zero and one. An ϵ value of one indicates an ideal radiation emitter, also known as a blackbody.

The Stefan-Boltzmann law directly links emissions with temperature, and thus an observer may calculate the temperature of a material by measuring its radiation emission.

Another step in thermography is knowing how much emission is coming from one's subject of interest, and how much is coming from reflection and materials between the subject and the sensor. By Kirchoff's Law, the emissivity of a material in thermal equilibrium defines the fraction of energy of a given wavelength it emits ϵ, the fraction it absorbs A, and the fraction it reflects r, where $\epsilon = A$ and:

$$r = 1 - \epsilon \tag{5.2}$$

Kirchoff's Law indicates that materials with high emissivity absorb almost all incident energy in the thermal spectrum and reflect very little. This makes them easy to image via thermography since most of the measured emissions – at least at short ranges – are from the material of interest. Conversely, materials with low emissivity reflect more of the surrounding temperature – known as *downwelling radiation* – and larger corrections must be made.

Finally, these equations assume an ideal flat surface which is photographed directly along its surface normal line. Emittance is also a function of angle from surface normal. That is, even very diffuse emitters produce more radiation at an angle of 0 degrees from normal (directly facing the subject plane) than off to one side or another, falling off in theory via a cosine relationship with the angle from surface normal (Wolfe and Zissis, 1978, pg 1-6). When referring to this angle, which radiation exits along while leaving a surface, the term *exitant angle* is often used.

Unless pointed precisely along surface normal to a truly smooth surface, a thermal camera measures the radiation emitted along many different non-zero exitant angles throughout its frame. Thus, for thermography, the term *photographic angle* refers specifically to the exitant angle between the camera sensor and the center of the subject.

5.5 Equipment

Thermal imagers have significantly reduced in price over the last decade. Currently an imaging system that provides adequate resolution of at least 320 x 240 pixels, and a better-than 0.05 C between-pixel sensitivity within a single image can be purchased from multiple companies for under US $15,000. The experiments described in this paper use a FLIR B300. Such equipment is often also portable in a backpack with room for other field materials, and is capable of operating below -25 °C for thirty minutes – at which point the limitation is the comfort of the field operator.

Modern thermal imagers require a factory calibration in a near-blackbody setting. Yet camera drift seems minimal: after taking over 700 field images, the imager used for experiments in this paper still measured 0.0 °C for clean slush water through 1 m of inside air.

Many modern thermal imagers, including the one used for the applications in this paper, use a microbolometer containing many pixel-type sensors. Upon absorbing radiation particles, these pixels change some measurable property such as resistance. Material-level changes in measured pixel differences, then, are often called *raw voltages*. These voltages are then directly processed by the thermal camera into a measure of temperature, and a temperature value at this stage is known as a *brightness temperature*.

An uncorrected brightness temperature may be considerably different than the actual temperature of the subject. To estimate this actual temperature, thermographs need to be field calibrated – as is true for all materials – and additional corrections can be needed specifically for snow.

For most modern thermal imagers, basic corrective values such as emissivity and humidity may be entered directly into the thermal imager to use the correction specific to the camera. Some imagers also offers manipulation of integration times – a similar concept to exposure in visual photography – and further corrective factors. The individual algorithms for atmospheric and other corrections vary by equipment and are not discussed here.

The field calibration and snow-specific details, however, are addressed in-depth in the next sections. Each section presents its specific consideration, related previous work, and any new supplemental analysis together. Following that, the section entitled *Applications* contains a selection of possible uses for snow surface thermography, and a short summary concludes the paper.

5.6 Field calibration of images

The following subsections outline what information a user may need in order to properly calibrate a thermal image, including material emissivity, sources of external radiation and radiation interference, and distance to subject.

5.6.1 Emissivity

The high thermal emissivity ϵ of snow makes it easy to thermally image. Snow emissivity has been variably found to be 0.98 for frost (Wolfe and Zissis, 1978, pg 2-77), above 0.98 for small-grained snow under 1000 μm (Dozier and Warren, 1982), 0.96 for flat solid ice and water (Wolfe and Zissis, 1978, 2-77), and as low as 0.8 for old snow (Wolfe and Zissis, 1978). Measurements at specific wavelengths (e.g. 10 μm) have seen emissivities as high as 0.995 (Dozier and Warren, 1982) and 0.997 (Hori et al., 2006).

Inversely, work on the thermal reflectance of snow (Salisbury et al., 1994a) shows r to be less than 2.3% between the 4 to 14 μm wavelengths. This holds for many snow types: granular, fine, wet, and dry; newly fallen snow reflected less than 1% between 4 to 14 μm.

For non-bandwidth specific measurements, 0.99 and 0.98 seem to be good general ϵ values for use in dry seasonal snow thermography.

For most of the 7.5 to 13 μm spectrum, the emissivity is greater than this (Salisbury et al., 1994a; Dozier and Warren, 1982). However, the largest deviation that reaches 0.98 and below occurs at wavelengths near 11 μm and longer. Wien's displacement law places the peak wavelength of snow emission at 10.6 μm for snow at -0

°C, and 11.0 μm at -10 °C. Bandwidth-specific distributions may be found in Dozier and Warren (1982) and Salisbury et al. (1994a).

5.6.2 External natural influences

Although the high emissivity of snow reduces the amount of correction needed, external energy sources still have impact.

Solar emission provides very little thermal infrared to be either absorbed or reflected. Most of the solar heating of snow is linked with daily maximum shortwave input (Armstrong and Brun, 2008). The radiation from the sun irradiates the top of the atmosphere with less than 1 $W\ m^{-1}\ \mu m^{-1}$ for the 7 μm wavelength and longer (Wolfe and Zissis, 1978, 3-36), as compared with over 2000 $W\ m^{-1}\ \mu m^{-1}$ at a wavelength of 0.45 μm, or visible violet light.

Although the sun contributes very little electromagnetic radiation in the thermal wavelengths, water vapour in the atmosphere can affect the amount of thermal radiation reaching the ground. Other atmospheric constituents such as ozone (O_3), methane (CH_4), and carbon dioxide (CO_2) also have absorptive and emittive windows in the 7.5 to 13 μm spectrum (Wolfe and Zissis, 1978), but these have little influence when the camera and subject are in the troposphere, and photos are taken in alpine areas with little air pollution. If such materials were to be present in the atmosphere in significant amounts, atmospheric interference could be minimized by using short distance photography.

The larger effect of water vapour can be predicted by its emissivity curve, which is similar to that of ice (Dozier and Warren, 1982). Typically, water vapour becomes heated by solar shortwave and emits diffusely, including in the downward direction toward the snow surface. A thick cloud can easily emit as much thermal radiation toward the snow as the snow emits out toward space, making clouds a much greater effect in the thermal spectrum than the sun.

Confirmation for this comes easily. A diffusing and highly reflective material such as crumpled aluminum foil – thermal $\epsilon = 0.05$ (Wolfe and Zissis, 1978) – or gold plating – thermal $\epsilon = 0.06$ (Salisbury et al., 1994b) – may be used to measure the downwelling atmospheric radiation. Care should be taken to make a truly diffuse reflector that reflects nearly equally in all directions, and, when used, the reflecting material should not be allowed to heat up via shortwave absorption (e.g. sun exposure) during measurement nor shaded from the emitting source of interest.

Such diffuse reflectors were placed on the ground in thick cloud conditions with the emissivity set for ice (0.98). Five temperatures – four from different sides of the reflector at 45° photographic angles (see section 5.7.3 below for a full description of photographic angle) and one facing straight down at the reflector were taken and averaged to obtain one reflected temperature measurement. Four such experiments were carried out for air temperatures in clouds between +1 °C and -7 °C, with clouds up to 500 m above the surface. The reflected temperatures were always within 3 °C of the air temperature within the clouds, which were obtained from a remote weather station. This is in sharp contrast to five additional experiments under a clear sky, where the reflected temperature equaled the minimum equipment temperature, or -40 °C.

These experiments do not attempt to generalize the complex mechanism of downwelling atmospheric longwave radiation; extensive and more specific work such as that from Sugita and Brutsaert (1993) exists on the topic. Rather, as thermography temperature is directly calculated from detected emissions, these experiments simply imply that (a) thermal emissions from a thick cloud with the above temperatures are on the order of emissions from snow at the above temperatures, and (b) simple diffuse reflectors can give the operator an idea as to whether the atmosphere is contributing significant longwave radiation.

External natural influences can be corrected for in various ways. Thermal imagers allow the user to input a value for humidity for corrections. Some imagers allow the user to input reflected apparent temperature, which allows the imager to account for downwelling radiation.

However, imagers often only allow one value for humidity or reflected temperature per image. With a complex scene including ice lenses, dense snow, and other conditions packed into one area, it can be useful to the user to know whether external influences will appear as higher or lower brightness temperatures. For example, if an ice lens of interest consistently appears warmer than the surrounding snow in cloudy conditions, it may be of interest to photograph the ice lens in similar but clear conditions as well.

5.6.3 Distance to subject

For close subjects within a few meters, proper calibration is easier than at longer distances. For typical snow and air temperatures and all other things being equal, the difference in measured temperature between 0.99 and 0.98 emissivity is less than 0.1 °C error while operating at one meter distance or less because the effect of intervening water vapour in the air is small.

At a kilometer of distance, poorly calibrated values within a range of usual winter atmospheric temperatures (+10 to -10 °C) can cause around 4 °C of error between the thermal brightness seen at-camera and actual surface temperature. This is partly due to the falloff of radiation power density reaching the sensor – approximated by the inverse square law for large distances (Wolfe and Zissis, 1978, pg 1-32) – and partly due to an inability to capture the variance in water vapour temperature and density over large distances with single values for air temperature and humidity.

As is discussed in depth below, at shorter distances the primary source of error is the angle of photography creating a different brightness temperature than the one desired, as discussed in Sect. 5.7.3, and operator heating, as discussed in Sect. 5.7.4.

At longer photographic distances, variability in air humidity and temperature have the potential to generate the largest amount of error due to there simply being more between the subject and the sensor. Also, at longer photographic distances fine spatial features will blur more than in visual photography. This is due to the power density – and thus irradiance – falloff being dependent on the power of the emission itself, where longwave radiation has less power than visual light.

5.7 Snow-specific considerations

In addition to properly calibrating each thermograph with the information above, the user should be further aware of some snow-specific nuances. These include times when the snow is heated below its surface by shortwave radiation, the effect of different ice morphology on reflection properties, the different uses for different photographic angles, and heating of the snow by the camera operator.

5.7.1 Solid state greenhouse effect

Snow as a material is subject to a phenomenon known as the *solid state greenhouse effect* (Brandt and Warren, 1993), where shortwave radiation may heat the snow below the surface but not necessarily at the surface. Depending on the porosity of the uppermost snow layer, this may provide variable thermal imaging of the snow surface.

Maximum daytime temperatures in a snowpack being found *below* the surface during clear, sunny conditions is a subject of some debate, as it sometimes occurs and sometimes does not (Fierz, 2010; Morstad et al., 2007; Brandt and Warren, 1993) for not-yet-fully understood reasons. The general theory states that, for a greenhouse effect to occur, the conduction of the snow above the point of solar heating must be poorer than the surface as an emitter. Since snow is more transparent to shortwave radiation than it is to thermal radiation, one can see that snow may heat below its surface where it absorbs incoming shortwave radiation, and emit efficiently at its surface to release longwave radiation.

If the shortwave radiation-generated heat at depth cannot conduct through the ice lattice back to the surface efficiently, this would trap heat below the surface. In turn, due to continued longwave radiation emission, this allows a *skin* at the surface to get cooler than the underlying warmed snow layer, which then remains warmer than the snow underlying it in turn. Without the greenhouse effect, the expected state of seasonal snow is for the surface to be the warmest portion of the snow during the day and the coolest portion at night (McClung and Schaerer, 2006, pg 52).

5.7.2 Snow morphology

Grain Size and Type. The effect of grain size on emissivity has been studied for grains with radii 50 to 1000 μm (Dozier and Warren, 1982), and thermal emissivity differences are less than 0.005 for wavelengths of 7 to 15 μm. This is in sharp contrast to, say, 1.3 μm in the near infrared spectrum where albedo – i.e. total diffuse reflectivity – ranges from 0.2 to 0.6 depending on grain radius (Armstrong and Brun, 2008, pg 56).

Although this angle dependency is covered in more detail in Sect. 5.7.3, flat oriented and non-diffusely reflecting ice forms such as ice lenses need be photographed with some care. In such cases, one can imagine that unwanted reflected radiation may be much more intense at certain angles.

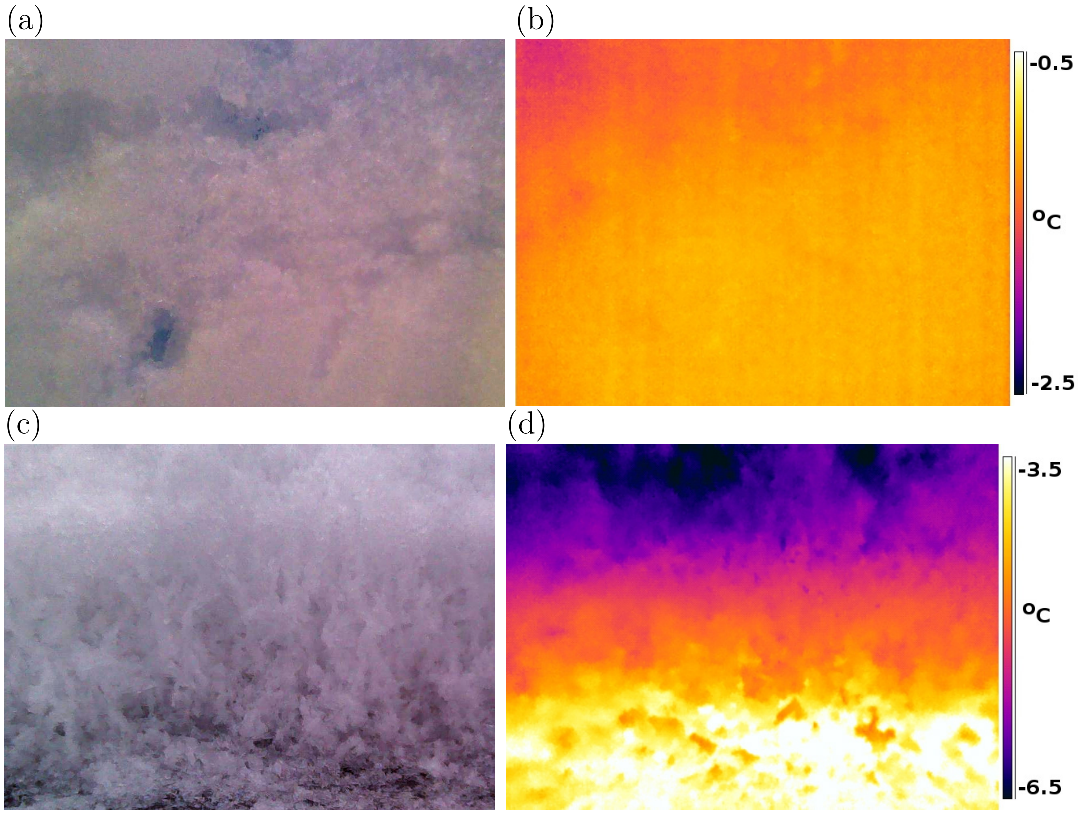


Figure 5.2: *Left*: Visual spectrum photographs of two areas with complex surface and grain qualities. *Right*: Thermographs depicting the same areas. Images (a) and (b) are of a visually complex crust with relatively uniform temperatures. Images (c) and (d) are of a visually complex layer of depth hoar with complex temperatures.

As snow becomes more morphologically similar to flat water, its emissivity decreases. Coarse refrozen granular crusts have been found to have emissivities close to that of water ($\epsilon = 0.96$) although finer grained crusts still mimic granular snow ($\epsilon = 0.98$), as found by Salisbury et al. (1994a). In the same work, the moisture content of granular snow (wet versus dry) appears to have little overall effect on emissivity, although wet snow displays lower reflectance by 0.005 in the 10 to 14 μm range.

Hori et al. (2006) found that emissivity decreases slightly with increasing snow grain coarseness. Fine grained snow showed an emissivity range of 0.984 to 0.997. Bare ice, on the lower extreme, ranged from 0.949 to 0.993.

Figure 5.2 photographically demonstrates the relative independence of thermal emission and grain size and type. Large variation in grain size and type can exist both with (Fig. 5.2c and 5.2d) and without (Fig. 5.2a and 5.2b) corresponding large thermal variation.

Grain Interactions. In laboratory experiments to measure infrared material interactions, Salisbury et al. (1994b) found that some fine powders naturally exhibited *clumping*, e.g. "bridging and void formation". This tendency created both an uneven surface and conditions for poor thermal conductivity. The combination of these two factors, they proposed, would create not only a steep temperature gradient between material surface and air due to the poor material conductivity, but also between different parts of the surface; specifically, the peaks of clumps and the valleys between them.

Though their experiments used quartz powder, this clumping phenomenon is a familiar one in some undisturbed new snow dendritic crystal surfaces. Experiments with similar types of powders with uneven surfaces of rapidly lessening density toward the surface (such as fresh snow) can deviate from Kirchoff's Law by as much as 6 percent (Salisbury et al., 1994b), perhaps affecting how the reflectance of clumped snow should be treated.

When combined with the solid state greenhouse effect detailed above, the surface of snow can be quite complex in its thermal layering. This subject and its pertinence to snow warrants further study. Figure 5.3 shows an instance of clumped new snow creating this type of layered thermal effect. Choosing the correct photographic angle, as discussed further below along with Fig. 5.3, may help the user measure the desired thermal layer.

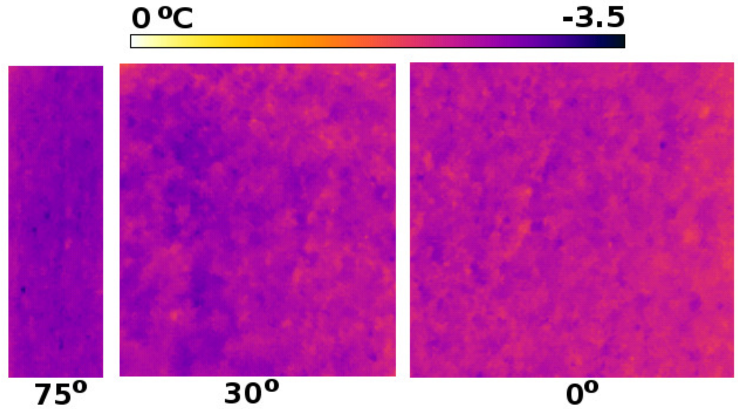


Figure 5.3: Thermographs of the same 20 x 20 cm area at a 75, 30, and 0 degree photographic angle.

5.7.3 Photographic angle

Different photographic angles measure different attributes of the snow surface. Dozier and Warren (1982), for space-based sensing purposes, found that apparent snow emissivity can vary by as much as 0.02 over the hemispheric angle range between 0° (along surface normal) and 75°. They found that apparent emissivity is at a maximum for 0° and decreases consistently to 75°. All such measurements found the emissivity $\epsilon > 0.96$. This translates into up to nearly a 2 °C difference between measured and real temperature at a 75 degree photographic angle, and less than a 0.5 °C difference for angles between 0 and 30 degrees, with most of the error occurring at wavelengths greater than 11 μm.

Conversely, from the work in clumping undisturbed powders described above in Sect. 5.7.2, using an infrared thermometer along the 70 degree photographic angle was recommended by Salisbury et al. (1994b) as an accurate measure of true *skin* surface temperature. In the case of snow, this means capturing the surface temperature of the grains actually at the very surface.

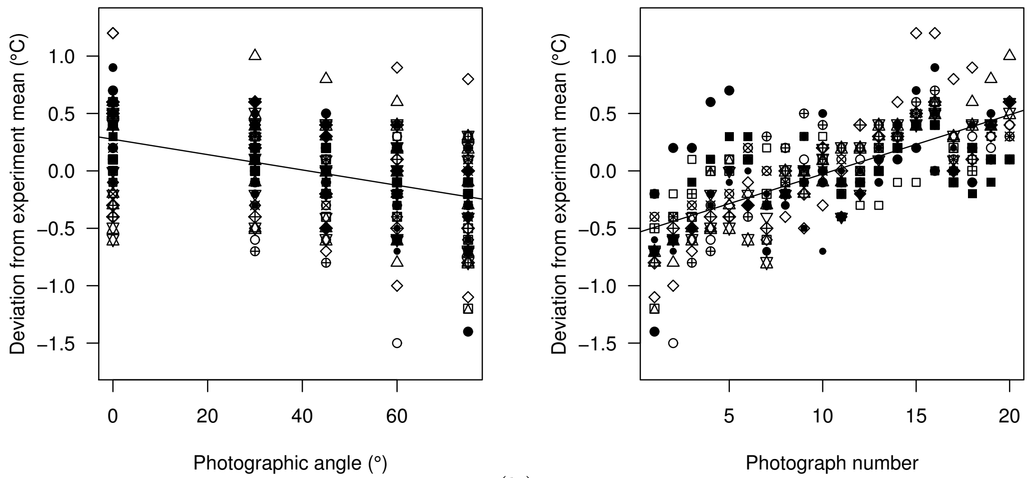
For fresh snow, these two factors – clumping and photographic angle variation –

may combine to complicate the separation between brightness temperature and the actual desired measurement. Figure 5.3 shows a series of photographs of clumping new dry snow, all of the same area at 0, 30 and 75 degree photographic angles. The warm underlayer is significantly more apparent at 0 degrees than at 75. One may see that the uppermost layer – the peaks of the clumps which are also the coldest portions of the snow – are the primary thermal features captured at a 75° photographic angle. In contrast, the 0° photographic angle captures the warmer – and lower – valley features between the clumps as well.

So on one hand, photographs at a shallow photographic angle, such as 75°, will experience less emission and perhaps measure a cooler temperature than desired. On the other hand, direct photographs at a photographic angle of 0 degrees may capture the warmer layer below the cooler porous surface snow. In short, the preferred angle of thermography will depend greatly on the application, especially given the possible clumping and the thin active thermal skin discussed above.

To assess the effect of photographic angle at small spatial scales, 15 experiments were performed taking thermographs at various angles to the surface. All photos were taken at a 70 cm distance from a 20 x 20 cm square area of natural snow surface, and each area was photographed at angles of 75, 60, 45, 30, and 0 degrees (five angles), from four different sides. This gave 20 thermographs per experiment and 300 photographs in total. An averaged area of approximately 10,000 pixels per thermograph gave an average temperature measurement for each experiment, side, and photographic angle. The photographs were taken in a different order for each experiment. The imaged areas varied, as did snow surface conditions (moist, dry, new, aged seasonal), time of experiment (from 0700 to 2200 hours, local time), and average snow surface temperature (between -0 and -10 °C).

To assess trends in the combined data, each experiment mean was re-centered at



(a) (b)

Figure 5.4: (a) Relationship between apparent surface temperature and photographic angle, with fit linear model. (b) Relationship between surface temperature and order in which the thermograph was taken, with fit linear model. Both graphs display each experiment with distinct symbols, and the mean of each experiment has been re-centered to 0 ° C to show cross-experiment trends.

0 °C. The resulting deviations from the mean were plotted in R (R Development Core Team, 2006). From this, the brightness temperatures were found to correlate with photographic angle (Pearson's r = -0.41, $p < 10^{-4}$) but had a stronger correlation with the order that the photographs were taken in, as discussed in Sect. 5.7.4 below.

A least-squares linear regression fitting to the data found the slope of the relationship between photographic angle and apparent temperature to be -0.0067 °C of temperature ($p < 10^{-3}$) per degree of photographic angle. This translates to 0.5 °C between a direct 0 degree angle photo and a very shallow 75 degree angle photo. This is approximately one-half to one-quarter of the difference found by Dozier and Warren (1982).

The residuals, when viewed on a Q-Q plot for normality, follow a reasonable but not precisely 45° linear relation. This implies that with a more extensive data set including an adequate examination of each specific condition (e.g. snow crystal type, size, and free moisture) a curved relationship might be extracted. Figure 5.4(a) shows the experiment data and the corresponding linear model.

5.7.4 Effect of operator heating

Measurements of snow temperature require that the observer be well insulated from the sample – by distance, clothing, etc. – in order to prevent *operator heating*.

Since human skin has an emissivity ϵ of 0.98, one sees by Eq. 5.1 that an unclothed human with 27 °C skin radiates 450 $W\,m^{-2}$, and a snow surface at -3 °C radiates 295 $W\,m^{-2}$. This yields around 155 radiative $W\,m^{-2}$ power for human skin to heat up a snow surface in the infrared spectrum, minus (a) the effect of any insulation such as clothing and (b) the effect of distance between the observer and the snow decreasing the power density.

The data from the photographic angle experiments described in Sect. 5.7.3 also revealed a strong correlation with the order in which the thermographs were taken (Pearson's r = 0.70, $p < 10^{-4}$). Least-squares linear regression on the data revealed a slope of 0.052 °C increase per photo taken ($p < 10^{-3}$). As an average experimental run of twenty angled images would take approximately ten minutes, and while taking photographs the operator was $\sim$ 1 m away from the sample, this implies an operator effect of 0.1 °C heating per minute while being within 1 m of a sample. Figure 5.4(b) shows the experiment data and the corresponding linear model.

With this relation, it takes an operator – wearing a down jacket and ski pants – approximately ten minutes to heat up the surface of a small snow sample 1 m away by 1 °C.

Additional cold lab experiments were performed to show that a bare hand within millimeters of snow will warm the snow surface nearly instantaneously. These experiments also showed it takes only around a minute for a hand in the same setup to warm a -15 °C snow surface by 6 to 10 °C in places, depending on the crystal morphology.

Such extremity of operator heating probably only holds true for these very close ranges between operator and sample. This topic warrants further study since (a) specific methods (such as insulation) to mitigate operator heating are as-yet untested, and (b) it very probably affects many snow pit temperature measurements. An example of the latter is seen below in Sect. 5.8.1, where extended operator heating and atmospheric heating penetrates at least to the depth of a short thermometer inserted into the snowpack.

5.8 Applications

Ideal snow surface thermography applications are those which need contactless measurement, benefit from an instantaneous spatial image, and have minimal influence from the considerations discussed above in Sects. 5.7.1 through 5.7.4. A selection of applications are presented below.

5.8.1 Thermal profiles

Hand-held thermography can be used to visualize the spatial variation of surface temperatures in a pit wall. Thermal profiles were created by overlapping multiple thermographs, with 6 to 10 thermographs per pit profile. The multiple thermographs per pit provide redundancy and show the relative stability of the thermal measurements over multiple photos. Point measurements from these photos came from the latest photo containing the point of interest, to mimic the thermometer being iteratively placed lower and lower along the pit wall over time.

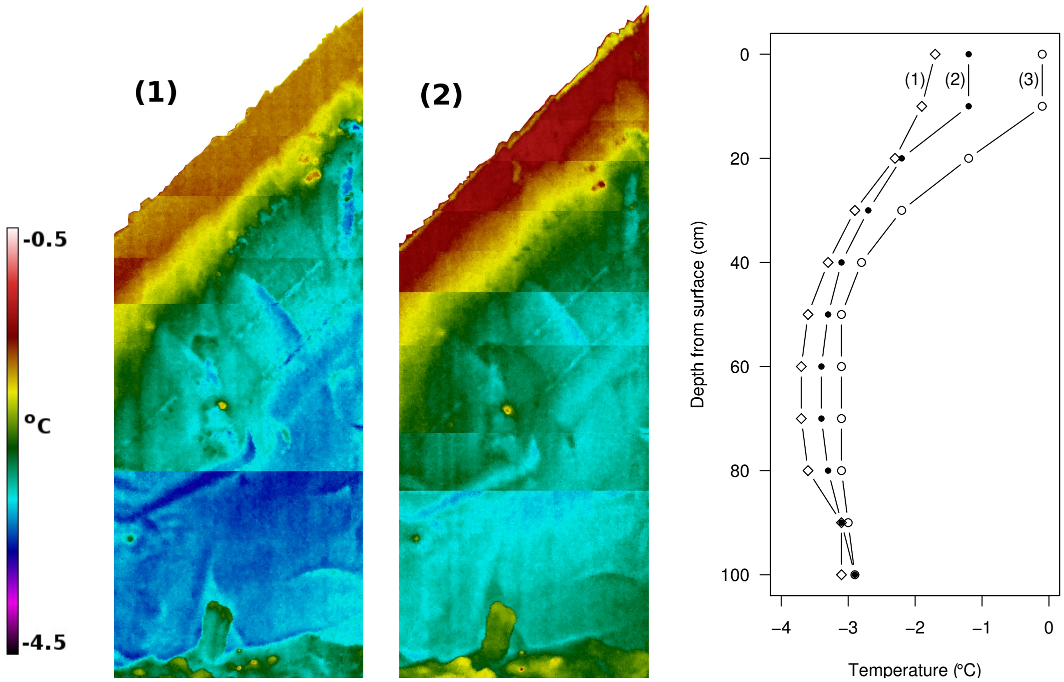


Figure 5.5: Two thermography temperature profiles of the same pit, in order of time photographed (1 then 2), and one temperature profile performed in the same pit (3, in graph). The effect of operator heating and atmospheric heating is visualized between (1) and (2), and the effect of atmospheric and operator heating at depth combined with spatial variation is visualized by adding point profiles from 1 and 2 to the graph containing profile 3.

To confirm the accuracy of these photographs, a standard temperature profile (Canadian Avalanche Association, 2007) was also taken with a calibrated point thermistor-type thermometer inserted $\sim$ 12 cm into the pit wall, in the classic method of temperature profiling (Fierz, 2010). Although thermograph measurements deviated from the hand-held thermometer by up to 2 °C, the relationship between traditional point and thermograph measurements have a Pearson's correlation of $r = 0.96$, $p < 10^{-4}$.

Figure 5.5 shows a three-stage progression of a profile obtained in this manner. It shows a thermograph from a freshly dug pit (1), a thermograph from the same pit $\sim$ 2 minutes later (2), and a standard point temperature profile taken in the same pit immediately after that (3). The thermographs are layered to put the values used in

(3) on top, that is, the first photos are under the later photos of the same area. In this way, the effect of operator heating and atmospheric changes on the surface may be seen over time, as well as the penetration of these effects into the snowpack. Note that the uneven pit wall areas, such as irregularities from shovel blade use, display photographic angle effects.

After the experiments were concluded (after approximately 30 elapsed minutes), $\sim$ 30 cm of surface snow was removed from a portion of the exposed pit wall, and a downward-looking thermograph confirmed horizontal heat penetration at depths 10 to 18 cm from the pit wall, with isolated patches of deeper penetration. However, one should note that the traditional point measurements of (3) occurred behind the snow surface at the length of the thermometer shaft, and thus slope-parallel spatial variation behind the pit wall may also play a role in those values.

5.8.2 Differing thermal conductivity

Snow conduction has been linked with snow density under the general theory that the denser the snow, the more bonds per volume, and thus the more paths for conduction (Armstrong and Brun, 2008, pg 36). However, recent research has shown that thermal conduction through snow is a spatially complex phenomenon, with conductive chains evolving generally but tortuously along a thermal gradient (Schneebeli and Sokratov, 2004). Other than that work, modelling, and bulk measurement, visualization and measurement of thermal conduction through snow is currently limited.

In a temperature-controlled lab at -15 °C, 2 cm thick snow samples were heated from behind using a 175 W infrared bulb. Samples of previously moist and subsequently refrozen rounded grains and polycrystals (RGlr and MFpc, respectively, as classified by Fierz et al. (2009)) were compared using time-lapse thermal video. The samples were extracted as they lay slope parallel by removing the snow above and

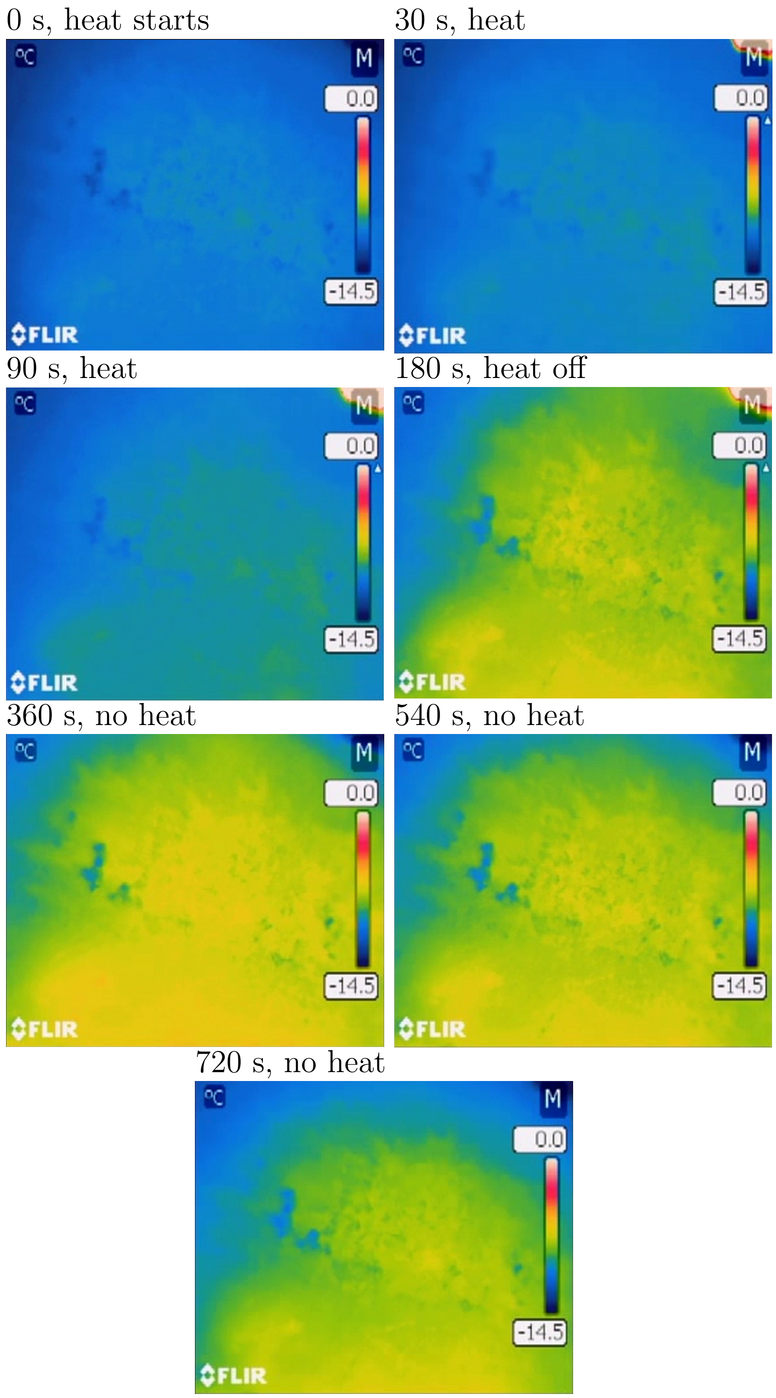


Figure 5.6: Seven frames of video from a 2 cm thick slice of a layer of rounded grains. The layer has crystals with extent to 0.5 mm and density of 390 $kg\, m^{-3}$. Heat was applied behind the sample from 0 to 180 seconds. Sample width is $\sim$ 10 cm.

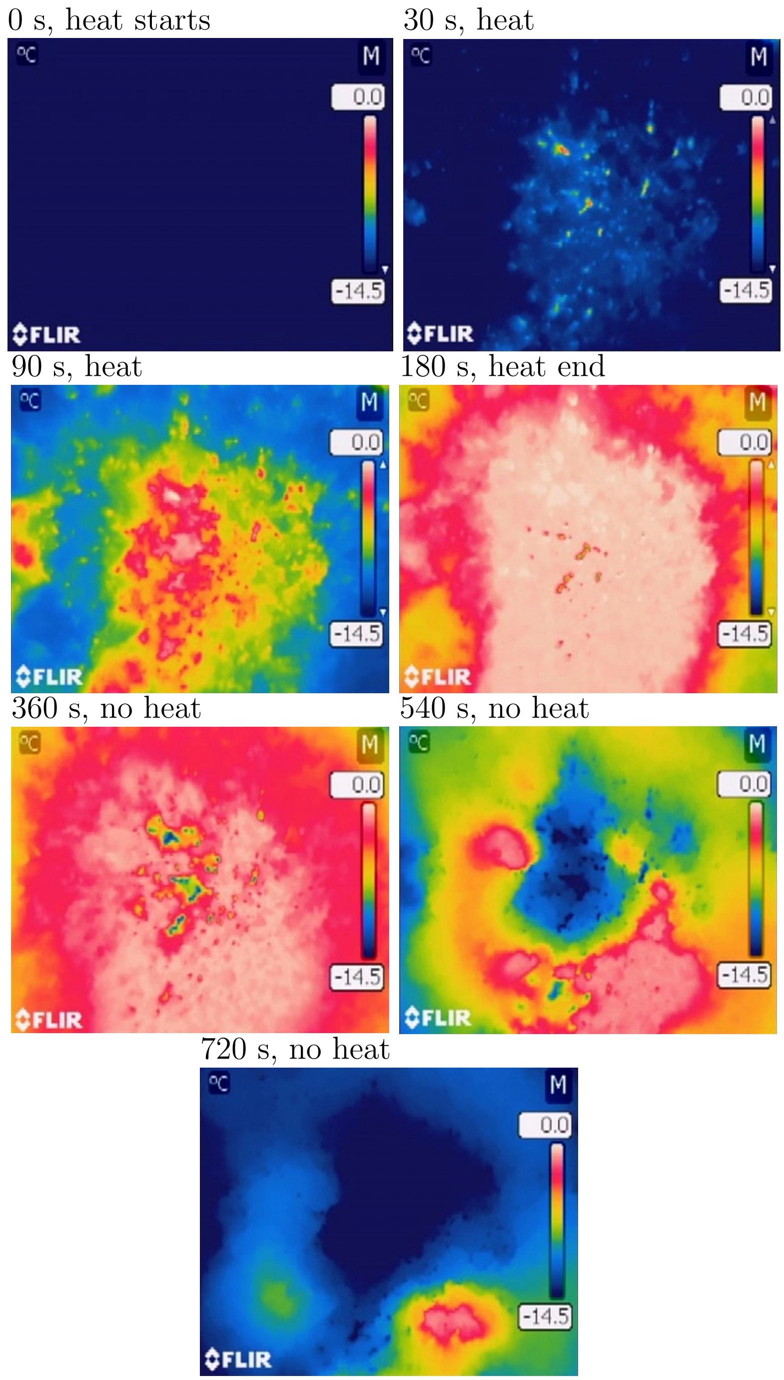


Figure 5.7: Seven frames of video from a 2 cm thick slice of a layer of polycrystals. The layer has crystals with extent to 6 mm and density of 270 $kg\, m^{-3}$. Heat was applied behind the sample from 0 to 180 seconds. Sample width is $\sim$ 10 cm.

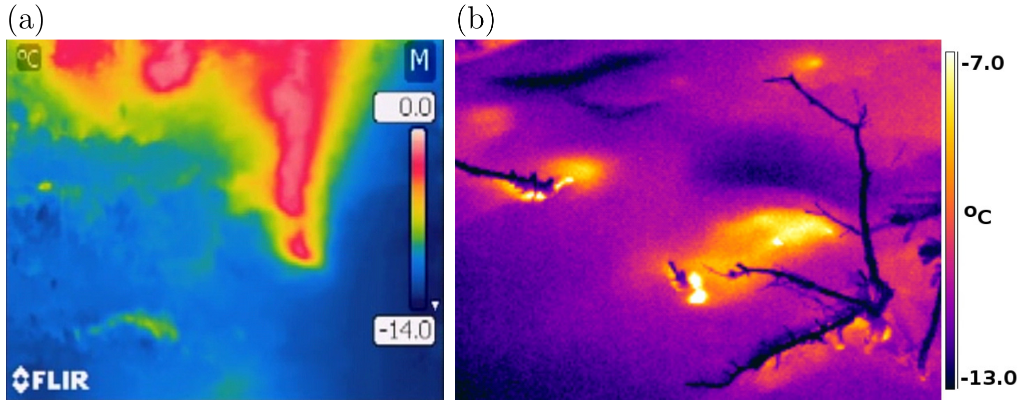


Figure 5.8: Two different types of heat penetration into snow. (a) A video frame of fingers of melt penetrating the surface of the snow after 20 minutes of heating. The fingers are up to 6 cm long. (b) Hibernating bushes are cold above the snow, but still transport heat from the ground to warm the snow from below. The image spans ~ 1 m.

below. These layers were then turned on end, heated for three minutes from behind, and allowed to cool for nine minutes each. The polycrystals measured to 6 mm in extent and had a density of 270 $kg\, m^{-3}$; the rounded grains measured to 0.5 mm in extent and had a density of 390 $kg\, m^{-3}$.

Figures 5.6 and 5.7 show frames from these time lapse videos. It is apparent from these videos that although the polycrystal layer had a much lower density, it was a faster conductor. Further, this particular layer of polycrystals appears much more spatially variable as a conductor, although perhaps this may be due to not being able to view the individual crystals in the sample of rounds.

Although emissivity may not vary much by crystal morphology, these videos show that conductivity does vary substantially. This form of visualization is quite new – only recently have advances in snow and radiation modelling (Kaempfer et al., 2007) enabled infrared reflection, refraction, and absorption to be modelled and visualized, and such methods operate on a per-photon level using lattice models of real snow.

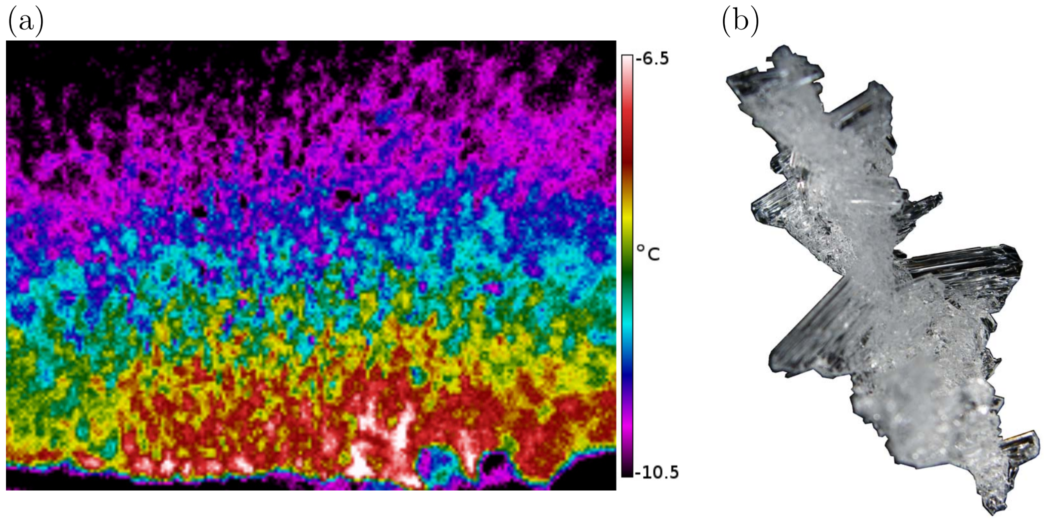


Figure 5.9: Thermal application in crystal metamorphosis. (a) An apparent phenomenon of *thermal chaining*, where chains of mature faceted crystals form aligned from the ground (bottom of image) to the snow surface (top of image). These crystals display physical connection and subsequent isothermal tendencies. They appear as isothermal fingers pointing from the base upwards. The image spans $\sim$ 30 cm in width. (b) A physical depth hoar chain extracted from a similar layer, length $\sim$ 40 mm

5.8.3 Additional applications

Thermal effects occur many places without our explicit knowledge. This section outlines additional applications at a high level, in order to show the breadth of application available. Beyond specific interest for avalanches, handheld thermography has wide application for the snow sciences generally, from providing confirmation data for spatial surface radiative balance models on glaciers, to detecting the effects of wind pumping.

Melt Tracking. The meltwater penetrating the snow in Figure 5.8a shows that the heat transfer from liquid water into the snow can be captured using thermography. This particular thermograph shows that the fingers of melt maintain a tight radius of heat around the wet area with minimal conduction beyond that.

Effect of Vegetation. Although thermography measures surface temperatures only, it can be useful in measuring thermal qualities caused by buried objects. In Figure 5.8b, a buried bush conducts heat upwards and creates a warmer area on the surface. Bushes often affect the internal snow temperature gradient around them and are associated with facet growth. Thermography would have application to these studies, both in measuring heat at the surface and, with excavation, studying the associated temperature gradients around the objects themselves.

Although shallowly buried bushes are easily seen on thermographs, the application of thermography to avalanche rescue seems limited. The bushes have been buried long enough for their heat to conduct to the surface to be seen, and it is this conduction time that would make thermography ineffective in detecting buried humans in a timely manner. Previous work at the Icelandic Meterological Office used barrels of warm water buried under the snow at various depths (0.5, 1.0 and 1.5 m) and thermal cameras to examine any possible clues of that heat being detected at the surface. More than two hours elapsed before the experimenters stopped timing and probed into the snow down to the barrels, confirming that the heat was detectable through the air in the probe holes but not on the snow surface (Jonsson and L.O., 1995).

Satellite Downscaling. As discussed in the introduction, some satellites providing space-based remote sensing measurements yield data in the thermal range, including Landsat and MODIS. These platforms average large areas – from 60 m to more than 1 km – in one pixel worth of data. Handheld thermography could aid in studying the conditions of interest for downscaling this space-based data. Also, as discussed in Section 5.7.3, previous studies have used hemispherical thermal-range sensors to address corrections needed for different angles of space-based data acquisition. Handheld thermography can also study photographic angle effects, as shown earlier.

Crystal Metamorphosis. Previous work in μCT tomography (Schneebeli and Sokratov, 2004) demonstrates that ice crystal metamorphosis under a strong temperature gradient will develop growth preferentially along the temperature gradient. Designing an equipment setup to allow a thermal camera – with a macro lens – to view and record the thermal metamorphosis on one surface of a μCT sample would be complex. However, examining the microstructure of crystals exposed to steep and prolonged temperature gradients in a natural environment may provide supporting data for that sought through tomography. Figure 5.9 shows that crystals can form chains along these natural temperature gradients from ground to surface, and that such chains display isothermal tendencies that we call *thermal chaining*.

5.9 Summary and conclusions

This paper has discussed the physical basis for thermography, summarized previous work in snow thermal emissivity, addressed handheld thermography concerns specific to snow, and presented a few of the possible applications. On one hand, snow is easy to thermally image due to its high emissivity; on the other hand, complications such as the solid state greenhouse effect, the need to correct for observer heating, the application-specific choice of photographic angle, and the influence of water vapour in the atmosphere all contribute to making thermography a careful endeavor.

Despite this, thermography is a powerfully spatial and visual measurement method. It can help measure and visualize everything from heat conduction to instantaneous spatial temperatures. It is our hope that this paper inspires further work using this technique.

5.10 Acknowledgements

For logistical support, we are very grateful to the Avalanche Control Section of Glacier National Park. For careful field work, we thank Mike Smith, Katherine Johnston, David Tracz, and Sascha Bellaire.

For financial support, we thank the Natural Sciences and Engineering Research Council of Canada, HeliCat Canada, the Canadian Avalanche Association, Mike Wiegele Helicopter Skiing, Teck Mining Company, Canada West Ski Areas Association, the Association of Canadian Mountain Guides, Backcountry Lodges of British Columbia, and the Canadian Ski Guides Association.

Also thank you to Thomas Kaempfer, then of the USA Cold Regions Research Engineering Laboratory, for the initial idea that thermography might be used to measure snow surface processes, and Arni Jonsson who provided a reference for previous attempts to use thermography in avalanche rescue.

Chapter 6: Summary and Recommendations for Future Work

This thesis has presented four new methods for visualizing the spatial variation of different snow attributes. These methods were: (a) a method to efficiently and repeatedly collect spatial data across slopes (Chapter 2), (b) a method to relate Google Earth photography to growth of a surface hoar layer (Chapter 3), (c) a method to view modelled snow warming GIS data over the internet for use in hypothesis testing (Chapter 4), and (d) a method to observe a spatial field of surface temperatures on observational snow pit walls and the snow surface (Chapter 5).

The work in this thesis has had time to gain feedback from the avalanche community at large to assess how well it has achieved its goals of providing useful spatial visualization and integration of technology in operational and research use. The following sections outline notes from using each of the methods and its success in achieving the objectives.

6.1 Notes from the field: Star

Star was used actively in the field for two field seasons. Overall, it proved to be efficient, usable, and re-usable for repeat measurements and hence to visualize and measure properties over time and terrain. It successfully provided a way to spatially visualize the variation of surface hoar size in Chapter 3.

For observers who had laid out a Star before, it took approximately two hours total to lay out a Star via ski tracks in an area 40 to 100 m on a side and then sample 48 surface sample points. For a sample area with lines and sample points already laid out, it took one observer one hour to repeat the 48 points, or two observers one hour to repeat the points and perform two measurements at each point such as crystal size and surface temperature.

Star was found to work best on areas between 40 $\times$ 40 and 100 $\times$ 100 m. Theoretically it is scalable to any distance, but having the transects be straight and the general area be square becomes hard when one drapes the sample layout over, say, a mountain. To ensure randomness in the field, a spacing book was designed with seven different random spacings for squares 40 to 100 m on a side. The book was brought into the field, and when a Star sample was laid out, the spacing was taken verbatim from the book. Not only did this ensure random placement of the points along each transect, but it also provided a way of keeping track of the spacing used by recording a reference number for a particular layout rather than all of the distances between all of the points.

Due to the observer being able to visually confirm following a straight line in the field, and by using a pacing booklet with random sample point layout distances, the elimination of observer bias was the key to being able to spatially visualize conditions. When sampling for the data in Chapter 3, the observer was required to push through tight stands of trees, over awkward rolls, and so on. The sample points obtained in these less-than-desirable areas gave a more accurate representation of variation than otherwise would be obtained. Due to the lack of centralized point concentration in Star, the ease-of-sampling bias - possibly favouring open areas - was avoided.

A few steps were necessary in the field. The efficiency of Star is built upon pacing rather than measuring. Therefore, an inefficient measuring method would reduce the efficiency of the overall method. Often, the measuring method of choice was to put duct tape pieces on one's skis such that when the very toe edge of the leading foot's boot slid forward to just reach the edge of the duct tape, that leading toe was exactly 1 m in front of the following toe. For smaller increments, smaller duct tape markers

Figure 6.1: The Google-maps based analysis application for Star. Here, one track has been uploaded and the idealized Star overlay fit to the track. The user has the ability to zoom in many more steps to better view the track and layout.

and labels could be added to skis. Another method used probes, which are long and relatively unwieldy but which are marked by the centimetre.

To lay out an entire Star using the ski-and-duct-tape method, an efficient method was found to lay out the outer boundary as a box first, and thus be able to sight 90 degree angles at each corner using a compass. From there, one third along each boundary line could be marked, and then the sampling lines actually running by the sampled points simply had to connect those thirds-marks on the boundaries in a straight manner. Star held intuitive speed advantages to the other three methods compared in Chapter 2.

In dry snow, the location for repeat measurements would often be preserved via evidence of the previous measurement, such as the dent of a crystal screen from sampling crystals on the snow surface. These small measurement supports allowed repeat measurements to be taken within a few centimetres of the first sample. In warmer temperatures where the surface changes rapidly, markers such as small pieces of tree needed to be places on the non-sampled side of the ski tracks to remind the observer of the location of the point.

Global positioning system (GPS) tracks would be taken at each sample; repeat samples offered additional tracks over the exact same lines that could be augmented with terrain data at each point for further precision.

Another step with Star was translating the GPS track to actual latitude and longitudes of each sample point for spatial and other analyses. To this end, I wrote a custom map application that provided the ability to upload a GPS track and graphically lay out an idealized Star along and over the track lines. The ideal Star layout could then be stretched and turned to fit the actual GPS track and recorded terrain data. Figure 6.1 shows a screenshot of this custom application being used to obtain the location points for analysis.

Despite the need for a pacing book and custom computer program, Star successfully fulfilled the goals of (a) being a useable, applied tool for spatial visualization of snow processes, as evidenced by its use in Chapter 3, and (b) bringing technology to the field in a practical manner, as evidenced by the success of the pacing layout method and its reduction of observer bias by de-centralizing the layout, as well as the statistical comparison in Chapter 2.

6.2 Notes from the field: Greyscale

Although the manuscript explored the link between surface hoar size and distance from trees in terrain imagery, this found an explanation only in surface temperatures, from which the more general link to sky view was proposed. Thus, although it was useful for Google Earth imagery, we found that temperatures and greyscale correlation alone provided an inadequate conceptual link to avalanche forecasters and observers

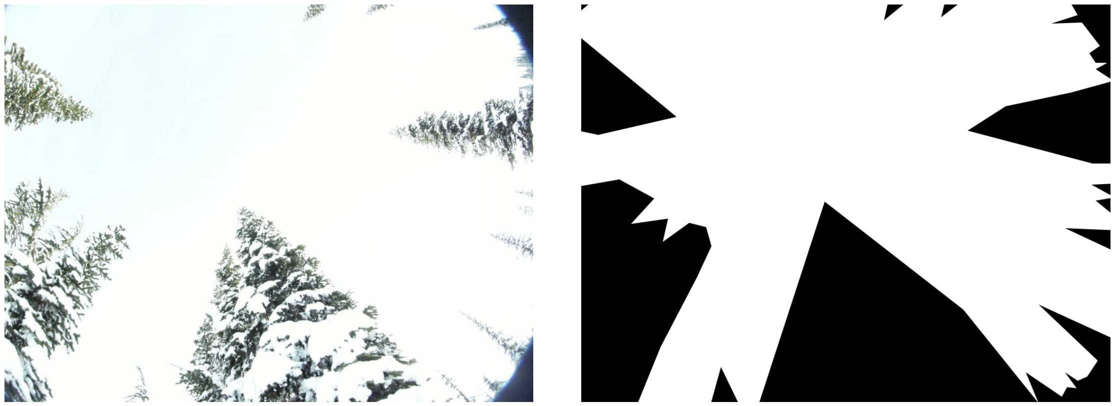


Figure 6.2: Example wide angle lens photograph (left) of skyview for one of the 48 points, and (right) the corresponding white and black mask matrix

for use in the field.

So, we collected more data at the end of February, 2009. We repeated the day/night size and temperature measurements using the same methods described in Section 3.4, and in the same area as the areas used to build the original greyscale relation. To connect the concept of sky view to the size of surface hoar, we took 175-degree wide angle photos of slope-perpendicular skyview at each of the 48 points from approximately one half metre above the snow surface.

We masked each of the 48 photos by hand into areas of sky (white) or no sky (black). The hand method was chosen because the snow covering the trees would at times be a similar color to a cloudy sky and was therefore difficult to mask automatically. Example original photos and masked photos can be seen in Figure 6.2. From a results perspective, we found good correlation between open skyview percentage (white) and both night surface hoar size and night surface temperature (Pearsons 0.52 and -0.46 respectively, $p < 0.001$), confirming that generally, the more skyview, the colder surface temperatures are at night, and the bigger the surface hoar grows as a result.

From a spatial visualization perspective, this additional data provided a valuable link to the greyscale findings for the field. Rather than, as an observer, having to abstract the findings based on distance from trees, the photos in Figure 6.2 closely mimicked what one would see when looking up while travelling. This linked the spatial finding from Chapter 3 to the needed visualization to make it practical in the field.

Further research since the publication of this manuscript (Lutz and Birkeland, 2011) has confirmed this longwave skyview tree effect on surface hoar via more physical simulations of skyview and longwave interactions. Consideration of skyview and openings in trees has helped the way our research group works in the field, sometimes observing audible weak layer collapse in open areas in the middle of forest openings, but less frequently at the edge of openings, and not in closed forest. When unstable surface hoar layers are thought to exist, the central part of sloping forest openings can be avoided.

Hence, Chapter 3 provided an example of when research-oriented results were enhanced by additional results purely for practical spatial visualization. Overall, through presentation at conferences to its users, and use during the travel of our own research group, this method has fulfilled the role of being a useable applied spatial visualization method for surface hoar.

6.3 Notes from the field: GSWarm

GSWarm was used operationally in a limited capacity during the 2009-2010 season, and released to the public and with more areas of operational interest during the 2010-2011 season.

The main comment on useful visualization attributes was in praise of the cartesian-type spread of thumbnail images displaying albedo and cloud cover variables. This is important because it identifies a distinct, non-obvious type of spatial visualization. The spread was designed because a large forecast area may have different types of cloud cover, and past history of the snow within the same area. Thus, when interpreting the small multiples of warming maps laid out by albedo and cloud cover, the forecaster can mentally estimate the extremes of albedo and cloud cover within the area, and thus also visualize the possible extremes of warming over the large area. Hence, GSWarm not only filled the role of spatially extrapolating warming over a single map, but also provided visualization of spatial extremes at a glance.

The main comment on possible improvements was that more terrain landmarks would be useful. Some landmarks such as roads, rail, a nearby town, and a commonly visited backcountry area was added to the single area used in the 2009-2010 season, but this was still often too sparse for some users. In particular, the visualization role of the GSWarm shaded terrain overlay appeared to be different than a map. On a map, a user will have some tolerance with spending time figuring out where things are, but to assess warming data for a particular area that time should be cut to a minimum for visualization of extremes as above.

Thus, although the temperature data maps stayed in full resolution for the 2010-2011 season when three new areas were added, some landmarks of *areas of interest to warming* were delineated to allow quick comparisons between days for the same slope of concern. An example of this type of landmark is shown for the Kicking Horse Canyon area in Figure 6.3.

Of interest, it seems that this model is used when forecasters deem warming to be already a problem, rather than using the model to determine when warming may become a problem. GSWarm's (anonymous) uses were tracked through the 2010-2011 season. The model received a handful of uses in January and February, but it received over 90 uses in March, a month well known informally for its warming problems with

Figure 6.3: The Kicking Horse Canyon middle-size terrain overlay. The pink area is not a terrain, road, or city landmark but rather an area that is of particular interest for warming concerns. Highlighting it allows easy comparison of a particular, area-specific concern across different temperature maps.

regards to snow stability.

The use of GSWarm for spatial visualization of warming in operational forecasting was successful overall. The use of this program showed that an important form of spatial visualization does not only have to occur via a single map, but can also occur through hypothesis testing. Furthermore, it shows that traditional spatial visualization landmarks – e.g. roads and terrain – help but do not entirely support fast landmark interpretation on maps. Therefore, it was useful to add landmarks specific to warming.

6.4 Notes from the field: Thermal

By far, the method that avalanche practitioners, educators, and forecasters find most engaging and interesting is the thermal imaging work from Chapter 5.

As this thesis is a study of methods, so to speak, it is quite interesting as to why.

First, it is important to note that thermal imaging is not (yet) used operationally. Hence, the non-research audiences are engaged with the *research results*. Most often, comments from the readers point out that thermal images are colourful. Which, as the GSWarm presentation method in Chapter 4 is also colourful, cannot be the entire answer.

To speculate, the reason may be that it can potentially replace a traditional method of obtaining point measurements with directly obtaining an image – or array of measurements – that is real. This seems similar as to why the much more general concept of observing spatial variation has also caught on with non-research avalanche audiences; the most useful observations provides colourful data images of what observers cannot easily see, yet desire to.

As a result, observers seem more interested in the *relative* temperatures than the absolute temperatures in the images. This points to the success of thermal images as a tool for spatial visualization. For example, in Figure 5.1, many viewers expressed surprise that the bed surface of the imaged avalanche was colder than the surface, but few asked what the temperature was. This translates well to exposed snow observation pits, where the relative difference in temperatures drives vapour pressure differences and therefore contributes to crystal morphology and strength, and which is discussed in Section 6.5 below. As the sensitivity – e.g. between-pixel accuracy – of thermal imagers are better than their absolute temperature accuracy, the application is apt.

Overall, the thermal camera was very successful in providing a way to spatially visualize snow temperatures for many different applications, as discussed in Chapter 5 and specifically in Section 5.8, Applications.

6.5 Future work

Each of the methods presented in this thesis could be usefully expanded by future work.

Much could be gained by designing a set of spatial observation methods that lend themselves to different avalanche conditions and experiment types; Star and the others mentioned in Chapter 2 are but a beginning. New research has shown that multiple, continuous observations while travelling can give better information about avalanche conditions than a single traditional snow pit test at a single location (Jamieson et al., 2009; Bakermans et al., 2010). Star will probably remain a pure research method due to the need for a random spacing book and a mapping application to extract the location data. However, bridging the gap between pure research methods and potential future spatial observation methods to help recreationalists and guides place these newer, more informal observations and resulting visualization of variability over terrain could greatly impact the way observers assess snow.

The use of Google Earth in Chapter 3 shows that the use of remote sensing for avalanche path mapping, terrain complexity mapping, conditions mapping similar to the warming from Chapter 4, and many other directions, are promising. Furthermore, Google Earth has become well-used for route planning and run choices, so the integration of Google Earth photography into whatever future directions may result could provide a more seamless link between the research result and the use of that result by those who would use it practically.

The GSWarm warming model graphical presentation from Chapter 4 has promise as a template for future presentation of temporally complex GIS data. It would be a research topic itself – and an important one – to determine exactly what types of data presentation are most useful for forecasters and recreational users. As discussed in Chapter 4, many models exist, yet it is not clear how many are actually used in an applied setting.

The thermal imaging method from Chapter 5 needs a substantial amount of practial field use in the hands of researchers in order to make it an operationally viable method. Yet this method, in particular, presents the most exciting future opportunities for studying and observing natural snow metamorphism and surface variation with regards to weak layer formation, as begun by Shea et al. (2011). The visualization provided by a thermal imager could provide future avenues for examining weak layer evolution not easily explained by existing means. This type of evolution could be important in forecasting deep slab avalanches, and slab avalanches subsequent to fast atmospheric warming or cooling weather events.

References

Adams, E., McKittrick, L., Slaughter, A., Staron, P., Shertzer, R., Miller, D., Leonard, T., McCabe, D., Henninger, I., Catharine, D., Cooperstein, M., and Laveck, K. (2010). Modeling variation of surface hoar and radiation recrystallization across a slope. *In Proceedings of the International Snow Science Workshop, 27 September to 2 October 2009 Davos, Switzerland, Jürg Schweizer and Alec van Herwijnen, eds.*, pages 97–101.

Armstrong, R. L. and Brun, E. (2008). *Snow and Climate: Physical Processes, Surface Energy Exchange and Modeling.* Cambridge University Press, Cambridge, United Kingdom.

Atkins, R. (1992). Computer graphics applications in avalanche forecasting. *In Proceedings of the International Snow Science Workshop, 4–8 October 1992 at Breckenridge, CO, USA*, pages 116–125.

Bakermans, L. and Jamieson, B. (2009). SWarm: A simple regression model to estimate near-surface snowpack warming for back-country avalanche forecasting. *Cold Regions Science and Technology*, 59(2-3):133–142.

Bakermans, L., Jamieson, B., Schweizer, J., and Haegeli, P. (2010). Using stability tests and regional avalanche danger to estimate the local avalanche danger. *Annals of Glaciology*, 51:176–186.

Bartelt, P. and Lehning, M. (2002). A physical SNOWPACK model for the Swiss avalanche warning Part I: numerical model. *Cold Regions Science and Technology*, 35:123–145.

Bavay, M., Egger, T., Zwaaftink, C., Clifton, A., Perot, C., and Roche, J. (2010). The MeteoIO Library. Joint Project of the Hydrosys and Swiss Multi-Science Computing Grid projects, and WSL Institute for Snow and Avalanche Research SLF.

Bellaire, S. and Schweizer, J. (2008). Deriving spatial stability variations from penetration resistance measurements. *In Campbell, C., ed. Proceedings of the International Snow Science Workshop, Whistler, Canada, 21-27 September*, pages 188–198.

Birkeland, K., Kronholm, K., and Logan, S. (2004). A comparison of the spatial structure of the penetration resistance of snow layers in two different snow climates. *In Proceedings ISSMA-2004, International Symposium on Snow Monitoring and Avalanches, Manali (HP), India, 12-16 April*, pages 3–11.

Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modelling – a review. *Hydrological Processes*, 9:251–290.

Brabec, B., Meister, R., Stockli, U., Stoffel, A., and Stucki, T. (2001). RAIFoS: Regional Avalanche Information and Forecasting System. *Cold Regions Science and Technology*, 33:303–311.

Brandt, R. E. and Warren, S. G. (1993). Solar-heating rates and temperature profiles in antarctic snow and ice. *Journal of Glaciology*, 39(131):99–110.

Brown, M., Grimmond, S., and Ratti, C. (2001). Comparison of methodologies for computing sky view factor in urban environments. *International Society of Environmental Hydraulics*. Internal Report Los Alamos National Laboratory, Los Alamos, NM. LA-UR-01-4107.

Campbell, C. and Jamieson, B. (2007). Spatial variability of slab stability and fracture characteristics within avalanche start zones. *Cold Regions Science and Technology*, 47:134–147.

Campbell, C., Jamieson, B., and Hägeli, P. (2004). Small-scale mapping of snow stability: If not, why not. *Avalanche News, Canadian Avalanche Association, Revelstoke, BC, Canada*, 71:45–49.

Canadian Avalanche Association (2007). *Observational Guidelines and Recording Standards for Weather, Snowpack and Avalanches.* Canadian Avalanche Association, Revelstoke, British Columbia, Canada.

Canadian Avalanche Association (1991). Industry Information Exchange (InfoEx). Online Information at: http://www.avalanche.ca/caa/industry-services/infoex. Operational Years 1991–present.

Cheng, C. and Shiu, C. (2002). Frost formation and frost crystal growth on a cold plate in atmospheric air flow. *International Journal of Heat and Mass Transfer*, 45:4289–4303.

Cline, D., Armstrong, R., Davis, R., Elder, K., and Liston, G. (2001). NASA Cold Land Processes Field Experiment Plan. Online at http://www.nohrsc.nws.gov/ cline/clpx.html.

Colbeck, S. (1988). On the micrometeorology of surface hoar growth on snow in mountainous area. *Boundary Layer Meteorology*, 44(1):1–12.

Colbeck, S. (1991). The layered character of snow covers. *Reviews of Geophysics*, 29(1):81–96.

Colbeck, S., Jamieson, B., and Crowe, S. (2008). An attempt to describe the mechanism of surface hoar growth from valley clouds. *Cold Regions Science and Technology*, 54:83–88.

Colorado Avalanche Information Center (2010). U.S. fatalities by season, 1950/51 to 2009/10. Colorado Avalanche Information Center Statistics, 325 Broadway Street, Boulder, Colorado, USA.

Cookler, L. and Orton, B. (2004). Developing a GIS avalanche forecasting model using real-time weather telemetry information for the south side of Mount Hood. *In Proceedings of the International Snow Science Workshop, 19–24 September 2004 at Jackson Hole, WY, USA.*

Cooperstein, M. (2008). The effects of slope aspect on the formation of surface hoar and diurnally recrystallized near-surface faceted crystals. MSc Thesis, Department of Earth Sciences, University of Montana, Montana.

Cordy, P., McClung, D., Hawkins, C., Tweedy, J., and Weick, T. (2009). Computer assisted avalanche prediction using electronic weather sensor data. *Cold Regions Science and Technology*, 59:227–233.

Cressie, N. (1993). *Statistics for Spatial Data, Revised Edition*. John Wiley and Sons, Inc.

Davis, S. (Summer 2010). Playing Battleship: Doing battle with buried surface hoar, Calls for some innovations in strategy. *The Journal of Canada's Avalanche Community*, 93:48–49.

Deems, J., Fassnacht, S., and Elder, K. (2006). Fractal distribution of snow depth from LiDAR data. *Journal of Hydrometeorology*, 7(2):285–297.

Dozier, J. and Warren, S. G. (1982). Effect of viewing angle on the infrared brightness temperature of snow. *Water Resources Research*, 18(5):1424–1434.

Ellis, C., Pomeroy, J., Essery, R., and Link, T. (2011). Effects of needleleaf forest cover on radiation and snowmelt dynamics in the Canadian Rocky Mountains. *Canadian Journal of Forest Research*, 41:608–620.

Feick, S., Kronholm, K., and Schweizer, J. (2007). Field observations on spatial variability of surface hoar at the basin scale. *Journal of Geophysical Research*, 112.

Fierz, C. (2010). *Encyclopedia of Snow, Ice and Glaciers*. Springer Publishing, New York. Singh, Vijay P. and Singh, Pratap and Haritashya, Umesh K. (eds.).

Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., McClung, D., Nishimura, K., Satyawali, P., and Sokratov, S. (2009). *The International Classification for Seasonal Snow on the Ground*. IHP-VII Technical Documents in Hydrology N. 83, IACS Contribution N. 1, UNESCO-IHP, Paris.

Floyer, J. and McClung, D. (2002). Numerical avalanche prediction in Bear Pass, British Columbia, Canada. *In Proceedings of the International Snow Science Workshop, 29 September to 4 October 2002 at Penticton, BC, Canada.*

Föhn, P. (2001). Simulation of surface-hoar layers for snow-cover models. *Annals of Glaciology*, 32:19–26.

Gassner, M. and Brabec, B. (2002). Nearest neighbour models for local and regional avalanche forecasting. *Natural Hazards and Earth System Sciences*, 2:247–253.

Giraud, G., Martin, E., Brun, E., and Navarre, J. (2002). CROCUS-MERA-PC Software: A tool for local simulations of snow cover stratigraphy and avalanche risks. *In Proceedings of the International Snow Science Workshop, 29 September to 4 October 2002 at Penticton, BC, Canada.*

Government of Canada (2007). Geobase orthoimage 2005-2010. Online at http://www.geobase.ca. Government of Canada, Natural Resources Canada, Centre for Topographic Information - Sherbrooke (CTI-S).

Gravner, J. and Griffeath, D. (2009). Modeling snow-crystal growth: A three-dimensional mesoscopic approach. *Physical Review*, 79:011601–1–011610–18.

Gray, D. and Male, D. (1981). *Handbook of Snow: Principles, Processes, Management and Use*. The Blackburn Press, Caldwell, New Jersey, USA.

Gross, J. (2006). Nortest: Tests for Normality. R package version 1.0.

Hachikubo, A. (2001). Numerical modelling of sublimation on snow and comparison with field measurements. *Annals of Glaciology*, 32:37–42.

Hachikubo, A. and Akitaya, E. (1997). Effect of wind on surface hoar growth on snow. *Journal of Geophysical Research*, 102(D4):4367 – 4373.

Haegeli, P. (2004). *Scale analysis of avalanche activity on persistent snowpack weaknesses with respect to large-scale backcountry avalanche forecasting*. PhD, University of British Columbia.

Hägeli, P. and McClung, D. (2000). A new perspective on computer-aided avalanche forecasting: Scale and scale issues. *In Proceedings of the International Snow Science Workshop, 2–6 October 2000 at Big Sky, MT, USA.*, pages 66–73.

Hägeli, P. and McClung, D. (2007). Expanding the snow-climate classification with avalanche-relevant information: initial description of avalanche winter regimes for southwestern Canada. *Journal of Glaciology*, 53(181):266–276.

Heierli, J., Gumbsch, P., and Zaiser, M. (2008). Anticrack nucleation as triggering mechanism for snow slab avalanches. *Science*, 321(5886):240–243.

Hofierka, J. (1997). Direct solar radiation modelling within an open gis environment. *Proceedings of JEC-GI'97 conference in Vienna, Austria, IOS Press Amsterdam*, pages 575–584.

Höller, P. (1998). Tentative investigations on surface hoar in mountain forests. *Annals of Glaciology*, 26:31–34.

Hori, M., Aoki, T., Tanikawa, T., Motoyoshi, H., Hachikubo, A., Sugiura, K., Yasunari, T., Eide, H., Storvold, R., Nakajima, Y., and Takahashi, F. (2006). In-situ measured spectral directional emissivity of snow and ice in the 8–14 μm atmospheric window. *Remote Sensing of Environment*, 100:486–502.

Jamieson, B. and Geldsetzer, T. (1996). *Avalanche Accidents in Canada, Volume 4, 1984–1996*. Canadian Avalanche Association, Revelstoke, British Columbia, Canada.

Jamieson, B., Haegeli, P., and Gauthier, D. (2010). *Avalanche Accidents in Canada, Volume 5, 1996-2007*. Canadian Avalanche Association, Revelstoke, BC, Canada.

Jamieson, B., Haegeli, P., and Schweizer, J. (2009). Field observations for estimating the local avalanche danger in the Columbia Mountains of Canada. *Cold Regions Science and Technology*, 58:84–91.

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

Jonsson, A. and L.O., S. (1995). Search for avalanche victims with a heat camera, experiment at blue mountain area. Unpublished letter to the Director of the Civil Defense in Iceland, June 5th.

Kaempfer, T., Hopkins, M., and Perovich, D. (2007). A three-dimensional microstructure-based photon-tracking model of radiative transfer in snow. *Journal of Geophysical Research*, 112(D24113). doi:10.1029/2006JD008239.

Kaempfer, T. and Plapp, M. (2009). Phase-field modeling of dry snow metamorphism. *Physical Review*, 79(3):031502, 1–17.

Kronholm, K. (2004). *Spatial Variabillity of Snow Mechanical Properties with Regard to Avalanche Formation*. PhD, Department of Mathematics and Science, Department of Geography, University of Zurich, Zurich.

Kronholm, K. and Birkeland, K. (2005). Integrating spatial patterns into a snow avalanche cellular automata model. *Geophysical Research Letters*, 32(19).

Kronholm, K. and Birkeland, K. (2007). Reliability of sampling designs for spatial snow surveys. *Computers and Geosciences*, 33:1097–1110.

Kronholm, K., Schneebeli, M., and Schweizer, J. (2004). Spatial variability of micropenetration resistance in snow layers on a small slope. *Annals of Glaciology*, 38:202–208.

Lang, R., Leo, B., and Brown, R. (1984). Observations on the growth process and strength characteristics of surface hoar. *Proceedings of the International Snow Science Workshop*, pages 188–195.

Lang, R. M. (1985). Studies on surface hoar: Formation and physical properties.

MSc Thesis in Engineering Mechanics, Department of Engineering, Montana State University, Bozeman, Montana.

Lehning, M., Bartelt, P., Brown, B., and Fierz, C. (2002). A physical SNOWPACK model for the Swiss avalanche warning Part III: meteorological forcing, thin layer formation and evaluation. *Cold Regions Science and Technology*, 35:169–184.

Libbrecht, K. (2005). The physics of snow crystals. *Reports on Progress in Physics*, 68:855–895.

Lillesand, T., Kiefer, R., and Chipman, J. (2008). *Remote Sensing and Image Interpretation, Sixth Edition.* John Wiley and Sons.

Lutz, E. and Birkeland, K. (2011). Spatial patterns of surface hoar properties and incoming radiation on an inclined forest opening. *Journal of Glaciology*, 52(202):355–366.

McClung, D. and Schaerer, P. (2006). *The Avalanche Handbook, Third Edition.* The Mountaineers Books, Seattle, WA, USA.

McCollister, C., Birkeland, K., Hansen, K., and Aspinall, R. (2003). Exploring multi-scale spatial patterns in historical avalanche data, Jackson Hole Mountain Resort, Wyoming. *Cold Regions Science and Technology*, 30:299–313.

Merindol, L., G., G., and Giraud, G. (2002). A French local tool for avalanche hazard mapping: Astral, current state and new developments. *In Proceedings of the International Snow Science Workshop, 29 September to 4 October 2002 at Penticton, BC, Canada.*

Morstad, B., Adams, E., and McKittrick, L. (2007). Experimental and analytical study of radiation-recrystallized near-surface facets in snow. *Cold Regions Science and Technology*, 47:90–101.

National Weather Service (2010). Weather fatalities, 2001–2010. United States National Weather Service, 1325 East West Highway, Silver Spring, Maryland, USA.

Neteler, M. and Mitasova, H. (2008). *Open Source GIS: A GRASS Approach*. Springer Science and Business Media, LLC.

O'Sullivan, D. and Unwin, D. (2003). *Geographic Information Analysis*. John Wiley and Sons.

Pebesma, E. J. and Wesseling, C. G. (1998). Gstat: a program for geostatistical modelling, prediction and simulation. *Computers and Geosciences*, 24(1):17–31.

Purves, R., Morrison, K., Moss, G., and Wright, B. (2002). Cornice - development of a nearest neighbours model applied in backcountry avalanche forecasting in Scotland. *In Proceedings of the International Snow Science Workshop, 29 September to 4 October 2002 at Penticton, BC, Canada.*

R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Robinson, N. (1966). *Solar Radiation*. American Elsevier Publishing Company, New York, New York.

Rosenthal, W., Elder, K., and Davis, R. (2002). Operational decision tree avalanche forecasting. *In Proceedings of the International Snow Science Workshop, 29 September to 4 October 2002 at Penticton, BC, Canada.*

Salisbury, J. W., D'Aria, D. M., and Wald, A. (1994a). Measurements of thermal infrared spectral reflectance of frost, snow, and ice. *Journal of Geophysical Research*, 99(B12):24,235–24,240.

Salisbury, J. W., Wald, A., and D'Aria, D. M. (1994b). Thermal infrared remote sensing and Kirchoff's law: 1. Laboratory measurements. *Journal of Geophysical Research*, 99(B6):11,897–11,911.

Schirmer, M., Schweizer, J., and Lehning, M. (2010). Statistical evaluation of local to regional snowpack stability using simulated snow-cover data (In Press). *Cold Regions Science and Technology*. DOI: 10.1016/j.coldregions.2010.04.012.

Schneebeli, M. and Sokratov, S. (2004). Tomography of temperature gradient metamorphism of snow and associated changes in heat conductivity. *Hydrological Processes*, 18:3655–3665.

Schweizer, J., Kronholm, K., Jamieson, B., and Birkeland, K. (2008). Review of spatial variability of snowpack properties and its importance for avalanche formation. *Cold Regions Science and Technology*, 51(2–3):253–272.

Shea, C. and Jamieson, B. (2009). Predicting surface hoar spatial variability in sparse forests using shading in satellite imagery. *In Schweizer, J. and A. van Herwijnen, eds. Proceedings of the International Snow Science Workshop, Davos, Switzerland, 27 September - 2 October*, pages 103–106.

Shea, C. and Jamieson, B. (2010a). GSWarm: An example of making a GIS model for everyday use. *International Snow Science Workshop, Squaw Valley, California*, pages 523–529.

Shea, C. and Jamieson, B. (2010b). Some fundamentals of handeld snow surface thermography. *The Cryosphere Discussions*, 4:1–30. doi:10.5194/tcd-4-1-2010.

Shea, C. and Jamieson, B. (2010c). Spatial distribution of surface hoar crystals in sparse forests. *Natural Hazards and Earth Systems Science*, 10(6):1317–1330. doi:10.5194/nhess-10-1317-2010.

Shea, C. and Jamieson, B. (2010d). Star: An efficient snow point-sampling method. *Annals of Glaciology*, 51(54):64–72.

Shea, C. and Jamieson, B. (2011). Some fundamentals of handeld snow surface thermography. *The Cryosphere*, 5:55–66.

Shea, C., Jamieson, B., and Birkeland, K. (2011). Use of a thermal imager for snow pit temperatures. *The Cryosphere Discussions*, 5:1–34. doi:10.5194/tcd-5-1-2011.

Sicart, J., Pomeroy, J., Essery, R., Hardy, J., Link, T., and Marks, D. (2004). A sensitivity study of daytime net radiation during snowmelt to forest canopy and atmospheric conditions. *Journal of Hydrometeorology*, 5:774–784.

Statham, G., Haegeli, P., Birkeland, K., Greene, E., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J. (2010). The North American public avalanche danger scale. *Proceedings of the International Snow Science Workshop, Squaw Valley, California, USA*.

Sugita, M. and Brutsaert, W. (1993). Cloud effect in the estimation of instantaneous downward longwave radiation. *Water Resources Research*, 29(3):599–605.

Tufte, E. R. (1990). *Envisioning Information*. Graphics Press, Cheshire, Conneticut, USA.

Wolfe, W. L. and Zissis, G. J. (1978). *The Infrared Handbook*. The Infrared Information and Analysis Center (IRIA) Center, Environmental Research Institute of

Michigan. Prepared for The Office of Naval Research, Department of the Navy, Washington, DC and Arlington, VA.

Zeidler, A. and Jamieson, B. (2004). Computer assisted avalanche forecasting: skier-triggered avalanches. *Proceedings of the Western Snow Conference.*

Zeidler, A., Jamieson, B., Chalmers, T., and Johnson, G. (2006). SAWLEM: Slab and weak layer evolution model. *In Proceedings of the International Snow Science Workshop, 1–6 October 2006 at Telluride, CO, USA.*