# Mixed-Effects Regression for Snow Distribution Modeling in the Central Yukon

ANDREW KASURAK, RICHARD KELLY, ALEXANDER BRENNING [1]

## ABSTRACT

To date, remote sensing estimates of snow water equivalent (SWE) in mountainous areas are very uncertain. To test passive microwave algorithm estimations of SWE, a validation data set must exist for a broad geographic area. This study aims to build a data set through field measurements and statistical techniques, as part of the IPY observations theme to help develop an improved algorithm. Field measurements are performed at, GIS based, pre-selected sites in the Central Yukon. At each location a transect was taken, with sites measuring snow depth, density, and structure. A mixed-effects multiple regression was chosen to analyze and then predict these field measurements over the study area. This modelling strategy is best capable of handling the selective and hierarchical structure of the field campaign. A regression model was developed to predict snow depth from elevation derived variables, and transformed Landsat data. The final model is: Snow depth = horizontal curvature + cos(aspect) + log10(elevation range, 270m) + tassel cap (Landsat imagery) + interaction of elevation and landcover. This model is used to extrapolate field measurements over the study area. A second, simpler regression links snow depth with density giving the desired SWE measurements. This instantaneous snow map will allow for passive microwave (AMSR-E) remote sensing calibration work in a generally inaccessible area. The Root Mean Squared Error (RMSE) of this snow depth estimation is 12.58 cm over a domain of 200 x 200 km.

**Keywords**: snow, snow water equivalent, SWE, SD, mixed effects, regression, modeling, snow distribution, remote sensing, stepwise regression

## INTRODUCTION

The measurement of snow water equivalent over a region allows a number of important tasks to be undertaken such as water management for flood forecasting, and climate change studies. Estimation by remote sensing is ideal, as field sampling the spatial distribution of snow over large areas is time consuming and expensive (Erxleben et al., 2002). In northern regions, passive microwave observations are a good approach, offering frequent repeat, wide area coverage at a scale suitable for regional management. Winter clouds and darkness do not interfere with measurements, providing information when it is most useful. For remote sensing measurements to be converted into geophysical variables, a calibration data set must be obtained. Such data and validations exist for prairie environments, as well as modifications for tundra and boreal forests. It is known that the measurements must be adjusted for the presence of forest, and small lakes.

In mountainous areas, studies of snow distribution do not have sufficient extent to calibrate and validate an instrument which has a best case ground resolution of 6□4 km. Typical studies are done at the basin scale Watson et al. (2006). The relationships developed between terrain attributes

[1]University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1.

and snow depth (SD) or snow water equivalent (SWE) at the basin scale may not be transferable to larger regions with different topographic and climatic characteristics.

This study aims to generate a snow map of SD, and snow water equivalent through a second regression relationship, of large enough spatial extent to calibrate the Advanced Microwave Scanning Radiometer EOS (AMSR-E) sensor in mountainous regions for which no current data set exists.

To meet this objective, a multi-level spatial sampling design is developed, and a suitable linear mixed-effects model for analysis and prediction of SD and SWE is constructed from the gathered data. The sampling scheme is designed to achieve a good geographical coverage and reflect the variety of environmental conditions such as topography and land cover. It is further more adapted to the logistic constraints and to the presence of local scatter and spatial autocorrelation.

**Snow distribution**

There are two main factors that control how much snow is in a given spot on the ground: the amount of snow that falls, and the way it is (re)distributed. This study aims to understand the distribution of snow, assuming constant snowfall over the region. In a variety of similar regression type studies, the authors Anderton et al. (2004); Carroll and Cressie (1997); Elder et al. (1991); Elder et al. (1998); Erickson et al. (2005); Erxleben et al. (2002); Lapena & Martz (1996); Leydecker et al. (2001); Lopez-Moreno and Nogues-Bravo (2006); Luce et al. (1999); Molotch et al. (2005); Plattner et al. (2006); Stahli et al. (2002); Winstral et al. (2002); Trujillo et al. (2009) examine the regression relationship between snow depth (SD) or SWE and a variety of terrain factors including: elevation, slope, aspect, incoming solar radiation, land cover, wind (via shelter, or drifting). The produced regression equations are typically linear, with R squared of 0.98 to 0.78 and high variability. Sample spacing also ranges widely from 2m to 250m between points. Autocorrelation of SD (when reported) ranges from 18-30m, and residual autocorrelation generally has a range of 250m. SWE is frequently derived from a model of snow depth and a fixed or very simple relationship between depth and density.

## STUDY AREA

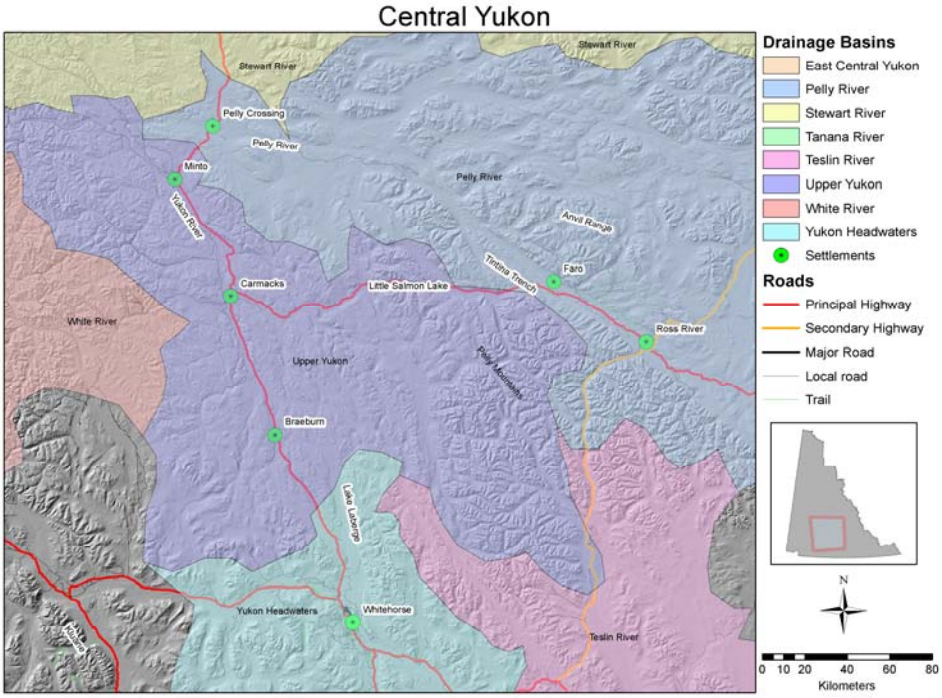



Figure 1: Field location in the central Yukon

The study area is located in the upper portion of the Yukon river watershed (Fig. 1) incorporating portions of the Yukon headwaters, Pelly river, and upper Yukon river drainage basins. The Yukon is mountainous with elevations ranging from sea level, to Canada's highest peak, Mt. Logan at 5959 m. In the study area, elevation ranges from 504m to 2222 m. The Tintina trench, a large rift valley 5–19 km wide with an average elevation of 600m a.s.l., runs NW-SE bordering the Pelly mountain ranges on its western flank.

The study area encompasses the region from Whitehorse in the south to Carmacks, approximately 200 km north along highway 6 and the Yukon river, then to the town of Ross river, 200 km east of Carmacks on Robert Campbell highway (highway 4), along the Ross and Pelly rivers The study area includes two large lakes, Lake Laberge, aligned N-S, and Little Salmon Lake, E-W, as well as numerous small lakes. Permafrost is sporadic for the majority of the study area, becoming discontinuous on the north and west edges.

Vegetation is dominated by white and black spruce on dry and wet soils respectively. Lodgepole pine is common as regrowth and on very dry areas. Aspen dominates burn regrowth and south facing slopes, birch and dwarf willow are also common. The tree line occurs around 1500 m. Details on the distributions of topographic variables are described in table 2, and more information on the region in Brabets et al. (2000). Large areas have been irregularly subject to forest fire in the last 50 years, and the burn zones are still easily identified in the field.

In climatological terms, the snow season for 2008 was an above average SWE (111–130% normal SWE) year in the study area, according to the Yukon snow bulletin (Janowicz, 2008), which observes snow pack conditions at 56 snow pillow and transect sites spread over the central Yukon. The snow season for 2009 was also above average (130–150% normal SWE). Mean annual temperature is -3 degrees celsius, with a cold and semi-arid climate.

**Table 1: Climatic SWE conditions (mm SWE) (Janowicz, 2008). Mean data based on records of 5–50 years.**

| Basin | Mean | 1-apr-08 | 1-apr-09 |
|---|---|---|---|
| Upper Yukon | 200 | 250 | 275 |
| Whitehorse | 125 | 130 | 200 |
| Pelly River | 150 | 175 | 210 |

### Fieldwork

Field measurements were taken between March 21st and 31st, 2008. Snow depth (SD) was measured with the Magna-Probe, a backpack based snow depth sensor with GPS receiver. The precision of the depth measurement is sub-centimeter. Depths were taken parallel to gravity, as per Cline et al. (2001), and Magna-Probe operating instructions. Measurements were located by hand-held GPS, and were taken as close as possible to the GPS location to reduce operator selection bias against difficult terrain or tree wells. If the probe struck rock or wood before ground, a new point was taken as close as possible. Snow density was sampled with an ESC-30 snow corer. Three measurements were taken in an equilateral triangle of edge length 1m located at the center of a study site. Cored snow was measured on a spring scale. An underestimation of density may be present during warmer periods of some days as some snow would adhere to the inside of the tube.

A snow pit was dug at least once per transect. Density, temperature and grain size were measured for each layer. This data will be required for passive microwave analysis.

## METHODS

### Stratified hierarchical sampling design

To achieve the goal of a wide area map of snow depth or SWE, the sampling strategy must be one that encompasses a large area. Logistical limitations prohibit intensive sampling over the whole area, so a subset was chosen close to winter roads. Sites were chosen within a stratification of driving distance from base camps.

Watson et al. (2006) showed that a stratified sample gave more efficient estimates of model parameters than a simple random sample given the constraints associated with acquiring measurements of SD and SWE. A simplified version of this was implemented in this work.

Watson et al. (2006) also make a number of observations on the inadequacy of typical sampling procedures. They suggest that due to spatial autocorrelation, the *in-situ* snow variation reported by many studies is biased downwards, as sample locations are closely spaced. Multiple sampling at short range allows for larger than point support, and thus more comparability to model data, which will likely have areal support from included DEM data. Mixed-effects models furthermore allow us to estimate variance components corresponding to different levels of the sampling design, and thus different spatial scales.

The sampling design in this study consisted of four hierarchical levels: year, transect, site, and measurement, and covered different topographic and landcover units through stratified transect selection. Through GIS analysis, the accessible portion of the study area was stratified based on terrain attributes. Then a set of locations for transects were randomly selected from these stratifications.

**Table 2: Comparison of field sites and study area terrain attributes. As distributions are non-normal median and interquartile range are reported. Variables are described in the section Terrain attributes.**

| Variable | median.field | IQR.field | median.area | IQR.area |
|---|---|---|---|---|
| elev | 706.00 | 251.00 | 1030.00 | 496.00 |
| slope | 0.04 | 0.11 | 0.16 | 0.22 |
| hcurv | 0.00 | 0.00 | 0.00 | 0.00 |
| vcurv | -0.00 | 0.00 | -0.00 | 0.00 |
| elev_stdev | 7.76 | 13.91 | 22.57 | 27.27 |
| elev_rng | 32.00 | 49.00 | 84.00 | 101.00 |
| ndvi | 187.00 | 73.00 | 202.00 | 50.00 |
| ts_bright | 115.00 | 56.00 | 106.00 | 52.00 |
| ts_green | 187.00 | 53.00 | 200.00 | 34.00 |
| ts_wet | 137.00 | 65.00 | 146.00 | 63.00 |
| solrad | 136203.52 | 8495.00 | 74123.92 | 15990.64 |
| aspect_sin | -0.18 | 1.50 | -0.04 | 1.48 |
| aspect_cos | 0.22 | 1.31 | 0.07 | 1.34 |

For each transect location, a traversable location was chosen in its vicinity (subject to property and terrain accessibility), and a transect was walked from this point. Each transect was made up of a number of intensive study sites arranged in a line, separated by 125m. This distance was chosen as the distance at which snow depth measurements could be considered to be independent in a similar land cover (Pomeroy & Gray, 1995).

Each intensive study site was laid out as an equilateral triangle, with a distance from the center to a vertex of 16m. At each vertex 3 SD samples were taken in a 1m equilateral triangle. From the vertex back to the center an additional 4 equally spaced measurements were taken, as described by Watson et al. (2006), modified by adding measurements along the lines from center to vertex. This alteration was done to facilitate rapid movement through difficult terrain. The goal of the measurement design was to construct a within-site variogram representing the spatial autocorrelation of measurements.

At the center of every other sample site, a measurement of SWE was taken either by snow pit, or density core (ESC-30 snow tube). One pit was dug per transect, and the remainder were density measurements. These allow the SD measurements to be converted to SWE measurements, as density was expected to change conservatively over space as compared with SD (Elder et al., 1998).

Overall, 37 transects with 212 sites were visited in the field in the two years of the study, containing a total of 7847 observations of snow depth, and 324 observations of snow density.

**Terrain attributes**

To summarize the characteristics of the study sites, a number of terrain attributes are calculated. A 90m digital elevation model (DEM) was acquired from Geomatics Yukon. From this, slope, elevation, aspect, and curvature (plan, profile, 3D) were derived in the System for Automated Geoscientific Analyses (SAGA) GIS program. Aspect was transformed to both sine and cosine to facilitate linear regression analysis. Positional information from the Magna-Probe allowed the ancillary information to be tied to the field data (via SAGA GIS and RSAGA package for R). The following description explains the terrain variables considered for inclusion in the model.

- Elevation (*elev*): Height above sea level in meters.
- Elevation Range (*elev_rng*): The range of elevation change (max-min) within a radius of 3 pixels (270 m). This variable is similar to slope, but has a larger support.
- Elevation Standard Deviation (*elev_stdev*): The standard deviation of all pixels surrounding the location within a radius of 270m. A measure of roughness.
- Slope (*slope*): Instantaneous slope at the location.

- Aspect: Sin, Cos, Factor (*aspect_sin, aspect_cos, aspect_fac*): Aspect of slope, measured in degrees from north, clockwise, decomposed into sine and cosine components, so as to remove the circularity of the number and allow for linear regression. Flat areas are assigned a value of -1 by the geoprocessor. To allow for proper incorporation of these values into the regression, aspect was classified into eight 45 degree groups (N, NE, E, SE, S, SW, W, NW), and one additional group for flat areas. Both sin/cos and factorial transformations were considered for inclusion.
- Curvature: V, H (*vcurv, hcurv*): Horizontal (plan), and vertical (profile) curvature.
- Incoming radiation (*solrad*): Net incoming radiation from 1 January to 31 March, assuming clear sky. Calculated as the sum of short and longwave radiation, taking into account the slope of the ground, and elevation. As described in Wilson & Gallant (2000), and implemented in SAGA GIS.
- Position: (*Lat, Lon; x, y*): Measured in the WGS84 coordinate system, NAD83 datum. Reported in decimal degrees. Transformed to Albers equal area conic, units (x, y) in meters.

3.3 Landsat landcover categorization and continuous vegetation fields

Spectral information from the Landsat satellite will provide landcover information at 30m resolution. These will be used as additional variables for the model.

**Feature extraction methods**

For this study, the detection of land cover, detailing coniferous, and deciduous forest, open water, burn areas was desired. The normalized difference vegetation index (NDVI) detects vegetation health, and can be related to leaf area index. The tasseled cap transformation is a principle-component decomposition with pre-determined coefficients (Jensen, 2007). The three primary components are brightness (*ts_bright*), greenness (*ts_green*) and wetness(*ts_wet*), respectively. Brightness is useful for detecting soil and urban areas. Greenness is similar to NDVI, and wetness detects water and vegetation. Both of these transformations work on a per-pixel basis, rather than with global statistics, reducing the impact of changing mean brightness levels or seasonality between different scenes (Jensen, 2007). *Classifications*: A supervised classification of the scenes was undertaken (*lc_sat*). Each scene was individually processed. For each land cover class of: Deciduous, Coniferous, Water, Burn, Ice & Cloud, pure pixels were selected, at least 100 each. Then a minimum distance classification procedure was run on the whole image, and the results mosaicked together.

**Available imagery**

Landsat imagery was acquired from the USGS Earth Explorer data portal, and the Yukon Geomatics data portal. When different years had to be selected for cloud free imagery, the same month was sought. All Imagery was from 1999–2001, and may thus be inaccurate in areas where forest fires have since occurred.

**Spatial analysis and prediction**

The first step in performing a regression was to choose the predicted variable. Two choices were available: SD and SWE. SWE is derived by combining SD and snow density, which was measured with a lower frequency than SD, causing SWE to be smoothed. However, since density was not uniform over the whole study area additional information might gained by including density as SWE prior to regression.

*Linear and non-linear regression*: A regular regression model was inadvisable, as there will be significant spatial autocorrelation between samples in the same site, and sites in the same transect due to the nested structure. This violates the assumption in linear regression of independent samples.

**Exploratory data analysis**

A scatterplot of elevation and snow depth revealed a local minima in trend at 640m a.s.l.. In addition, the response of elevation on water-covered surfaces was drastically different, presenting no strong upward trend. Above 1350m, the tree line in the field data, snow depth no longer increases with elevation, likely due to wind redistribution. For this reason, a new variable

(isforest) is created to categorize areas as water or above tree line, and its interaction with elevation.

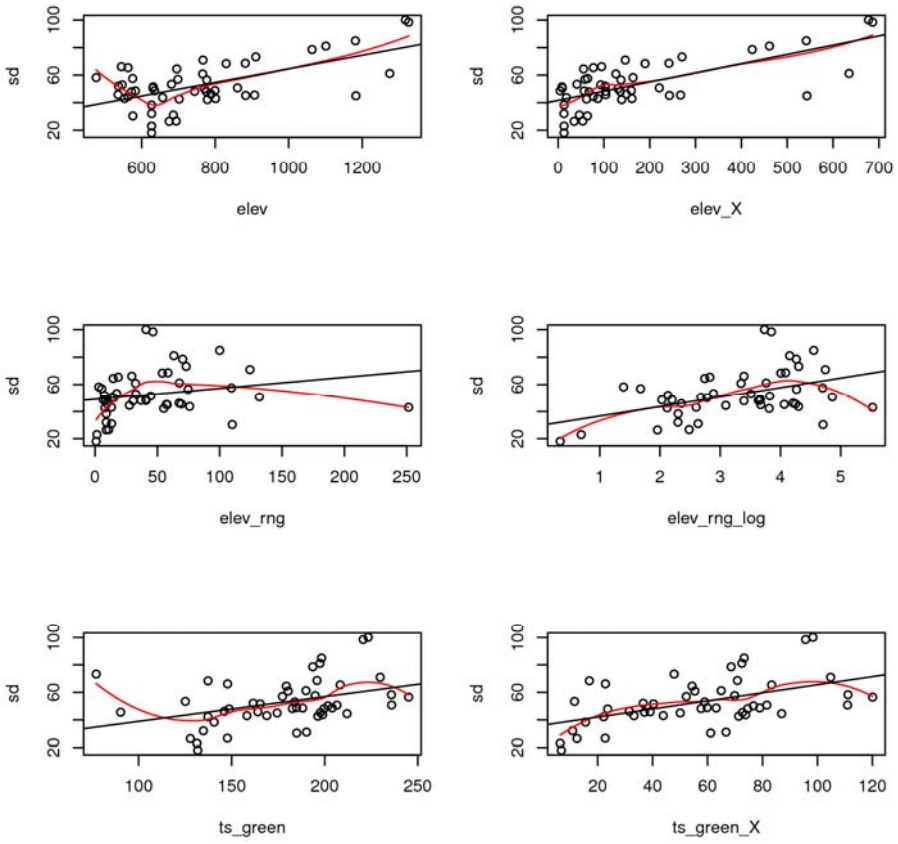



Figure 2: Transformations of select predictors (grouped at the site level, n=37), comparing observed (raw) data (left column) to transformed data (right column). The red line is a loess smoother fitting the data, and the black line represents a simple linear fit.

Scatterplots (Figure 2) of predictor variables versus snow depth, grouped at the site level, were examined for non-linearities. A number of potential centerings to enhance the linear relationships were noted:

- Elevation (elev_X): center at 640.
- Slope, range, stdev (*_log): add $log(x+1)$
- Aspect (aspect_(sin,cos)_X): center at 0
- All tassel (ts_*_X): center at 125
- Solrad (solrad_X): center at 136000.

An improvement in the linearity of the terms is evident in figure 2, especially the tasseled cap values. Non-linearity still exists in the remainder of the terms which cannot be removed by centering or simple transformations.

**Mixed Effects Regression (MER)** includes a structure to the normal regression, allowing grouping factors to have separate variance structures, and within-group correlations. This accommodates the non-independence of spatial sampling. For this study, a grouping factor based on the sampling scheme was the simplest approach.

$$y_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i, i = 1, \cdots, N \quad (1)$$

A MER model has two components shown in equation 1: fixed effects and random effects (Demidenko, 2005). Fixed effects (β) are the standard regression covariates. Random effects () are population parameters which affect the distribution of the variable through the covariance matrix, , but which will not necessarily be used calculate a direct relationship. In a multi-level model, they must also be used specify a grouping structure (Pinheiro & Bates, 2000). They must have a mean of 0, and **D**, and is the design matrix. The *i*'s are the observations, of which there are *N*. The variance parameters, and **D** are not known, and are part of the model fitting of the standard portion of the regression, β.

Using the sampling structure as a random effect partitions the resulting lack of fit variance into the sampling levels. While this information is unavailable to prediction (it is not known which transect a new point belongs to), it aids in the understanding of the snow depth variability. Testing a MER model can be undertaken on either the fixed or random terms. Fixed effects may be tested with t or F tests (conditional on the estimated variance), as likelihood ratios tend to be anticonservative (Pinheiro and Bates, 2000), although correction is possible through empirical simulation. Random effects specifications may be compared between models if they are fitted with the same fixed effects specification (Pinheiro and Bates, 2000).

**Mixed Effects Regression**

As the data set was multi-level, and its measurements were correlated within this hierarchy, weXXX use the mixed effects regression (MER) model as described by Pinheiro & Bates (2000)., who suggest a model building strategy that first builds a no-fixed effects model with the random effects specified. Then, in a forward-stepwise manner, candidate fixed effects are added and tested for significance using the t test. The next candidate fixed effect is chosen by graphical interpretation of structure in a plot of estimated random effects vs. covariate. Included random effects may be dropped if a fixed effect accounts for the intergroup variation.

**Random Effects**

The random effects must then be defined. Year (as López-Moreno and Stähli (2008)), transect, and site are the variables describing the nesting structure, and thus describe the random sample of the population. The direct contribution of transect or site on SD is not required, but rather the effect of the structure on the model fitting is necessary to allow for the non-independence of samples. With a more extensive field set, additional effects might also be described as random as well as fixed, such as elevation or landcover. However, without enough samples, this would reduce the ability of the fitting algorithms to find a solution. The random effects are configured to only adjust the intercept, not slope for the different locations. This was the simplest formulation. Random effects are chosen as measurements nested within sites, nested within transects, nested within years. This will control how parameters are estimated from observations, bringing the estimation in line with the sampling strategy.

A parsimonious model was desired, so as many candidates for fixed effects as possible will be discarded. Optimally those fixed effects which are highly correlated with other fixed effects, or which contribute the least to the model fit will be discarded. The list of independent variables includes the transformed and centered variables as well as the untransformed variables so as to test the value of the transformation.

**Regressor selection**: Stepwise selection is the tool best suited to reducing a large number of possible regressors, as it ensures impartial selection (of terms included in the model space). In accordance with Pinheiro and Bates (2000), a forward stepwise search was conducted. Rather than the suggested visual inspection of candidate regressors plotted against random effects, a quantitative statistic was generated and compared. This allows for bias free comparison, as well as allowing more of the parameter space to be explored. Pinheiro and Bates (2000) recommend against using log likelihood tests and ANOVA to compare models differing in fixed effects, as the

statistics will be unpredictably under or over conservative. Although an adjustment method is described, an alternate statistic of goodness-of-fit is simpler. The statistic used was the sum of the standard deviations of all of the random effects. A reduction in this value indicates that the model has explained more of the variance. Candidate regressors included all terms, as well as their listed transformations, in addition, a term representing the interaction of elevation and water was included.

Stepwise selection was performed inside a transect-level cross-validation (CV) framework. For each transect the model was built with that one transect excluded. The final model formulation was chosen as all terms which were included in at least 50% of the cross-validation subsets. This helps adjust for model shrinkage.

After the stepwise selection, included terms in the CV selected model which have a p-value greater than 0.5 in a marginal F-test are discarded, as are terms which are strongly correlated with other included terms. This p-value should not be considered absolute as multiple-testing has occurred, however certain terms may contribute very little. Equation 2 contains the results of stepwise selection procedure to reduce the full model with all terms to a sub-model including only terms with a detectable and significant contribution. Snow depth was calculated as the sum of the included terms multiplied by their respected regression coefficients.

$$sd = hcurv + aspect\ cos + elev\ rng\ log + ts\ bright\ X + ts\ green\ X + elev\ X : isforest \qquad (2)$$

**Weighting:** The model may have a weights term added, to adjust for the within-group heteroscedasticity structure. The standardized residuals were plotted against fitted values and the included regression variables. By observing the plot a heteroscedastic structure can be seen, with a characteristic wedge shape increasing from left to right. Having examined all candidate regressors, it was seen that lc_sat, ndvi, elev_rng, ts_green_X, and fitted values bear closer inspection.

An adjusted model for each of these possibilities for weighting was fit, and the resulting residuals were plotted again and assessed graphically (not presented). The classified landcover, and a power adjustment of fitted terms provide the best overall adjustment.

The two ratings were combined, and found to be better in terms of residual versus fitted scatterplots. This result indicates that variance was significantly different in different landcovers, and also increased with the fitted value, non-linearly.

**Correlation structure:** The model may have a within-group correlation structure specified. This structure will account for the spatial autocorrelation (SAC) of the measurements made within a site, and the sites within a transect. The within-site SAC were modelled using a variogram (Figure 3). Examining the variograms made for kriging in figure 3, the within site autocorrelation can be modelled with a spherical structure, with approximate values for range of 15 m, sill of 60 , and a nugget of 35 .

The model was adjusted for this correlation structure, giving a value for the fitted spherical semivariogram of: range 14.11, and a nugget factor (ratio of *nugget*/*sill*) of 0.372

**Final model**: Lastly, the variance adjustments and SAC corrections are combined and the model is refit under restricted maximum likelihood (REML).

## RESULTS

### Fieldwork and design

By examining the values in table 2 and the first column of scatterplots in figure 2 the success of the transect location selection can be assessed. In Elevation (elev), a downward bias is present, as higher elevation sites were not accessible. This is further reflected in the slope, range, and stdev variables as these are correlated with elevation. The sampling of vegetation was satisfactory.

### Exploratory data analysis

Due to the alignment of the survey along highways, latitude and longitude are dropped as predictors, as they are too strongly correlated with elevation. This gives very poor coverage of the

area, and so a potential trend of depth with location cannot be separated from the other attributes measured in each region. For instance, all of the very high elevation sites were in the NE of the study area, which was not representative of the study area's elevation, merely access. This would bias a relationship between position and depth, as high areas had more snow.

**Landsat classifications:** All regeneration should proceed at a similar pace for sites observed to be burned, however it is acknowledged some inaccuracy will be present from this.

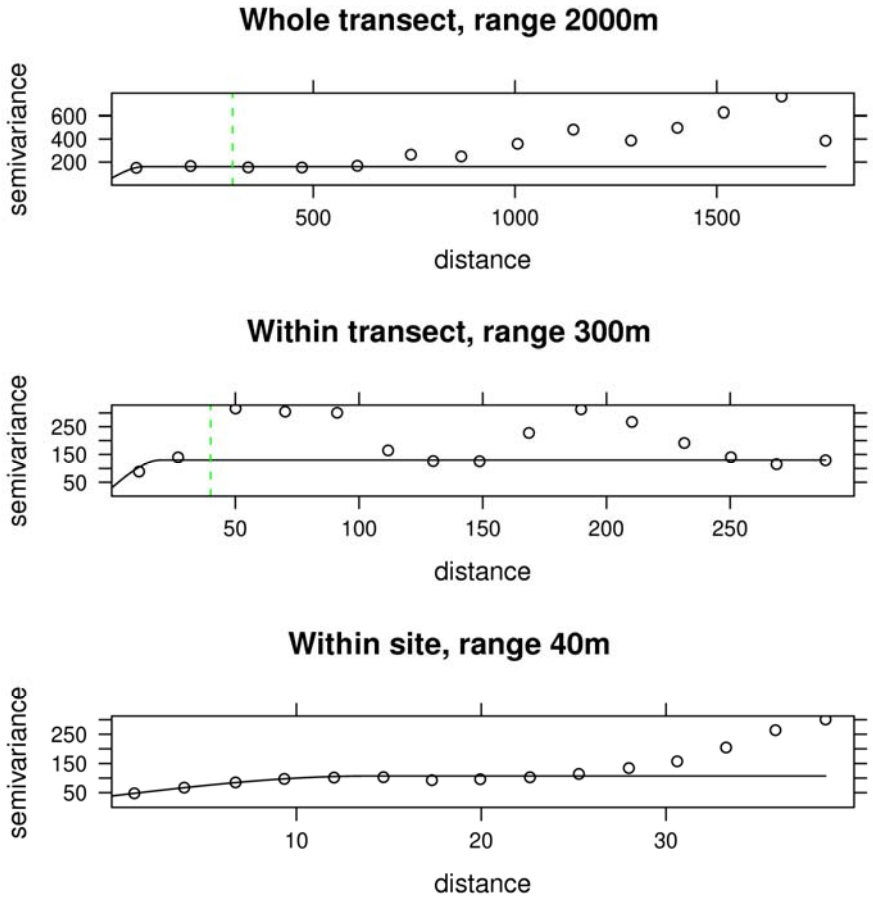



Figure 3: Variograms of snow depth at transect and site level. Semivariance is in.

In Figure 3 a within site scale SAC is visible having range of about 15m, accounting for about half of the semivariance, the rest being nugget. At the transect scale a sill was evident at the within-site scale of <30m, and it again occurred once enough distance has passed to include multiple whole sites (separated by 125m). At the whole transect scale, a sill was very evident until about 1000m where a scale-break was evident, probably due to the inclusion of multiple transects.

**Model description**

The terms and parameters of the snow depth model are presented in Table 3.

**Table 3: Final model for snow depth (cm)**

| | Term | Value |
|---|---|---|
| Autocorrelation parameters | range (m) | 14.000 |
| | nugget | 0.370 |
| Weighting parameters | conif | 1.000 |
| | bare | 0.830 |
| | water | 0.440 |
| | decid | 0.880 |
| | power | 0.660 |
| Fixed effect coefficients | Intercept (cm) | 39.000 |
| | hcurv (XXX) | 1059.000 |
| | aspect_cos_X (cos(degrees)) | 1.700 |
| | elev_rng_log (log(m)) | 1.000 |
| | ts_bright_X (DN) | -0.073 |
| | ts_green_X (DN) | 0.074 |
| | elevX_lctree (m) | 0.058 |
| | elevX_nlctree (cm) | 0.014 |
| Predictive performance | SDyear (cm) | 8.890 |
| | SDtransect (cm) | 7.440 |
| | SDsite (cm) | 5.120 |
| | SDres (cm) | 0.730 |
| | rmse_fit1 | 13.000 |
| | rsq_fit1 | 0.800 |
| | mbe1 | 0.780 |

The included terms have an understandable physical basis for interpretation. Horizontal curvature implies a sheltering effect from wind. Centered cosine of aspect relates to the east-westness of a slope, having potential implications for incoming radiation or wind shelter. Elevation range is an index of local terrain roughness, which may act to shelter pockets of snow, as well as a correlation with higher elevations. Tassel brightness indicates areas of bare ground, which are likely to be wind scoured. Tassel greenness indicates higher amounts of vegetation, which shelter and trap snow. SD increases with elevation (excepting when combined with lc_forest), providing the largest contribution. This is a similar finding to many other regression studies.

**Model diagnostics**

*Cross-validation*: The model was assessed by performing a leave-one (transect)-out CV. No term selection was performed during this step, rather the range of values assigned to the included terms was scrutinized. Table 4 contains a summary of the change in model statistics. The model was found to be reasonably stable, with the variation of the most influential terms below 10%, and the change in RMSE between fits having a coefficient of variation of 1%. It is likely that the low number of transects containing water features would have significantly contributed to the range of variation.

**Table 4. Cross-validation of the final model, no selection of fixed effects. RMSE_fit1 is the root mean squared error between observed and fitted values considered knowing (level of prediction) only the year.**

| | Term | units | mean | stdev |
|---|---|---|---|---|
| Autocorrelation parameters | range | m | 14.32 | 0.60 |
| | nugget | % of sill | 37 | 0.01 |
| Weighting parameters | conif | Variation, standardized to 100 | 100 | 0.00 |
| | bare | % of conif. | 87 | 0.26 |
| | water | % of conif. | 48 | 0.27 |
| | decid | % of conif. | 91 | 0.18 |
| | power | NA | 0.65 | 0.03 |
| Fixed effect coefficients | Intercept | cm | 38.71 | 0.57 |
| | hcurv | NA | 1102.27 | 300.29 |
| | aspect_cos_X | cos(degrees) | 1.70 | 0.23 |
| | elev_rng_log | log10(m) | 1.04 | 0.12 |
| | ts_bright_X | DN | -0.07 | 0.00 |
| | ts_green_X | DN | 0.07 | 0.01 |
| | elevX_lctree | m | 0.06 | 0.00 |
| | elevX_nlctree | m | 0.01 | 0.00 |
| Predictive performance | rmse_fit1 | cm | 12.58 | 0.25 |

**Random effects structure, GLS:** To demonstrate the utility of the random effects structure, the adjusted model was compared against a generalized least squares (GLS) regression model (Pinheiro and Bates, 2000). The GLS model was constructed with the same weighting and correlation structure, but without any random effects. Examining the GLS results, it was observed that the GLS model estimates a much larger range and smaller nugget than the MER model. This was due to the MER model building an individual variogram for each site, and then combining. The GLS also does not estimate the variance components associated with the grouping structure, which, while not central to the results of this study, may be of interest. It was therefore concluded that the MER model with random effects provided significant advantage.

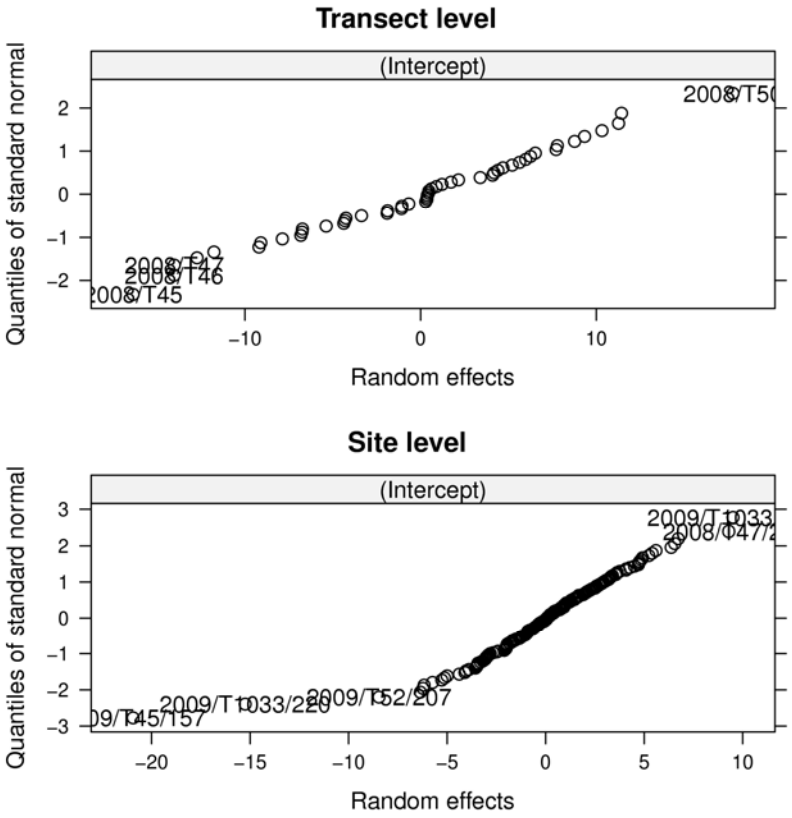
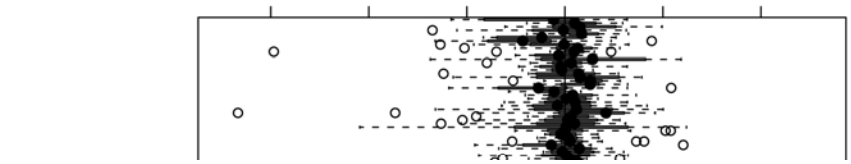



Figure 4: Q-Q plot to assess normality of random effects at transect and site level. Potential outliers to the normal distribution are labelled. Plotted values are the differences in intercept for each site (marked as Year / Transect number / site number), from the overall average intercept (in cm). A mean zero and a normal distribution are expected. The y-axis shows the expected value if the distribution is normal. Therefore a straight line of points show that the distribution of random effects is approximately normal.

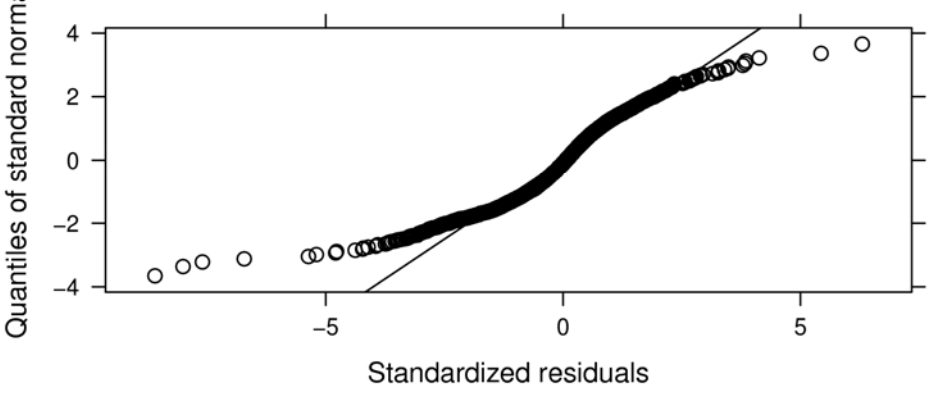
Examining the Quantile-Quantile (QQ) plot (Figure 4) for the random effects structure allows the assessment the distribution of the random effects. Ideally, the random effects follow the normal distribution. Site level random effect of the intercept was normal, with very few outliers.

Since there was only one random effect estimated, the intercept, there are no correlations within the random effects to check for, and no further adjustment of the covariance matrix was required.

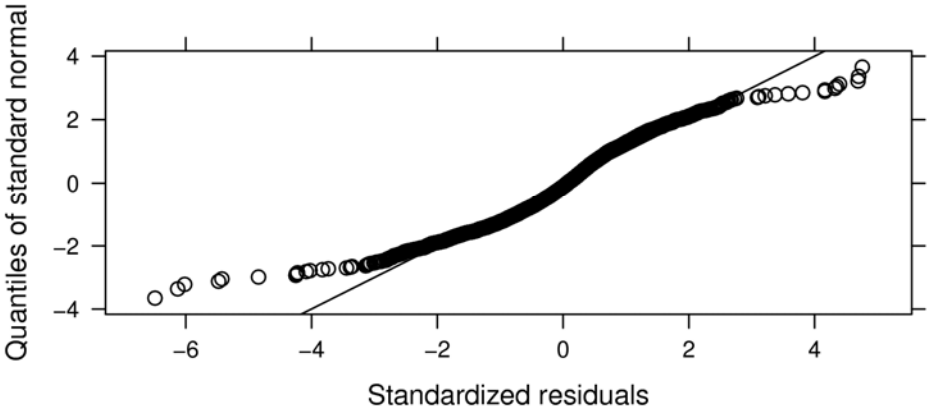
Second, the fixed effects component of the adjusted model was examined to ensure that the assumptions of the MER were met. The assumption of mean zero, constant variance in within-group errors are checked graphically in Figure 5.



**Full model**




**Adjusted Fit**



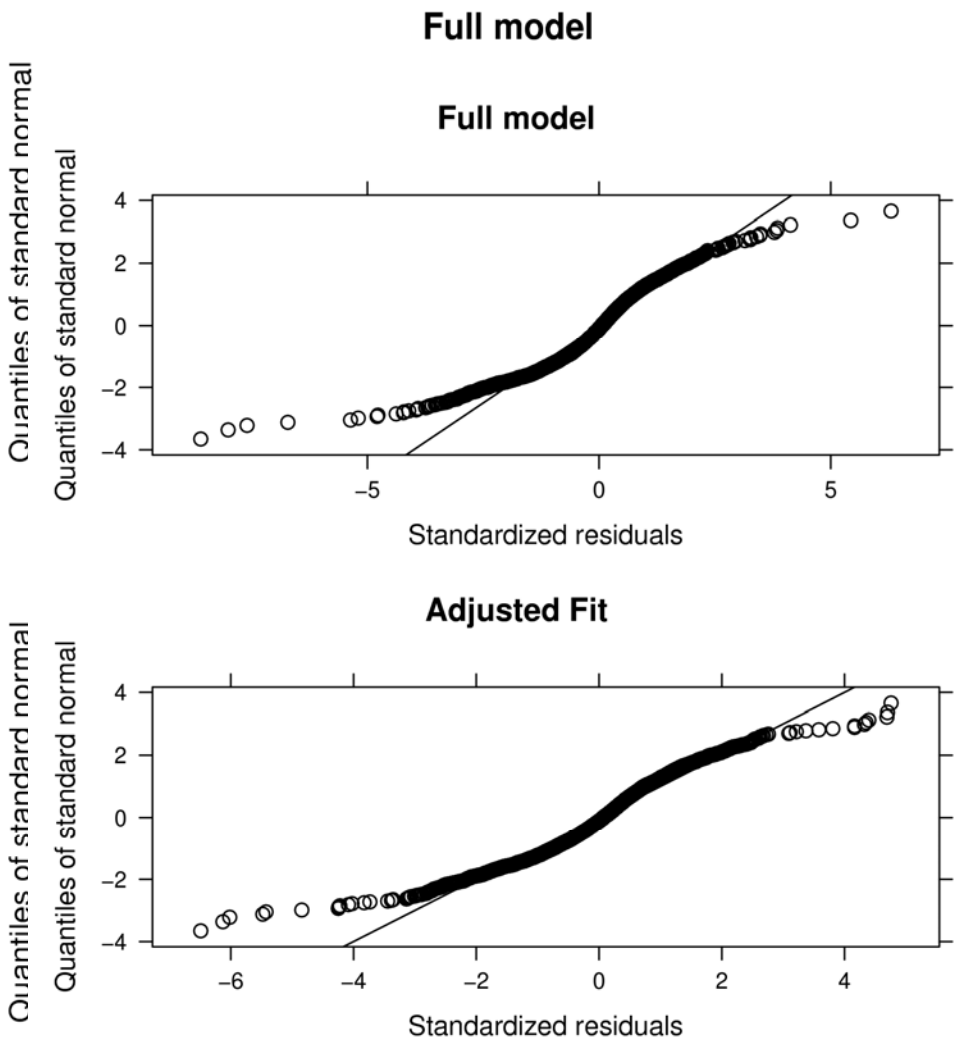



Figure 6: Q-Q plot to assess normality, full model vs adjusted and reduced model. Plot is of standardized residuals for fitted snow depth (open circles) plotted against their expected value if they follow the normal distribution. Points on or very near the solid line meet this assumption.

The adjustments to the model have improved the distribution towards normal, however the tails are still heavy. This may be due to snow depth being bounded by a minimum depth of zero, or by the minimum snow capture of even the smoothest sites (lakes). This departure from normality will inflate the standard error, and bias the F-tests for fixed effects. This is less of a concern as the model was chosen on its variance reducing capability rather than F-test for significance of fixed effects. Third, the modelling interaction terms between predictor variables has been avoided in this study for two reasons: (1): The number of available data points was not sufficient to work with the vast number of possible co-linearities produced by even a second order interaction. (2): The high achieved without interactions limits the potential explanatory power of interaction terms (Jost et al., 2007). Although, it was expected to be inflated by the correlations between predictors present within the model.

**Table 5: Accuracies of similar regression based studies.**

| Study | Mean SD (cm) | RMSE of SD prediction |
|---|---|---|
| This study: Central Yukon | 38.71 | 12.58 |
| Erxleben et al. (2002): St. Louis creek | 58 | 10.4 |
| Erxleben et al. (2002): Fool creek | 109 | 17.5 |
| Erxleben et al. (2002): Walton creek | 177 | 31.3 |
| Jost et al. (2007) | 10.75 | 8.77 |
| Molotch et al. (2005) | 255 | 77.64 |
| Lopez-Moreno and Nogues-Bravo (2006) | approx. 90 | 27.5 |
| Winstral et al. (2002) | 227 | approx. 40 |

**Model predictions (Snow map)**

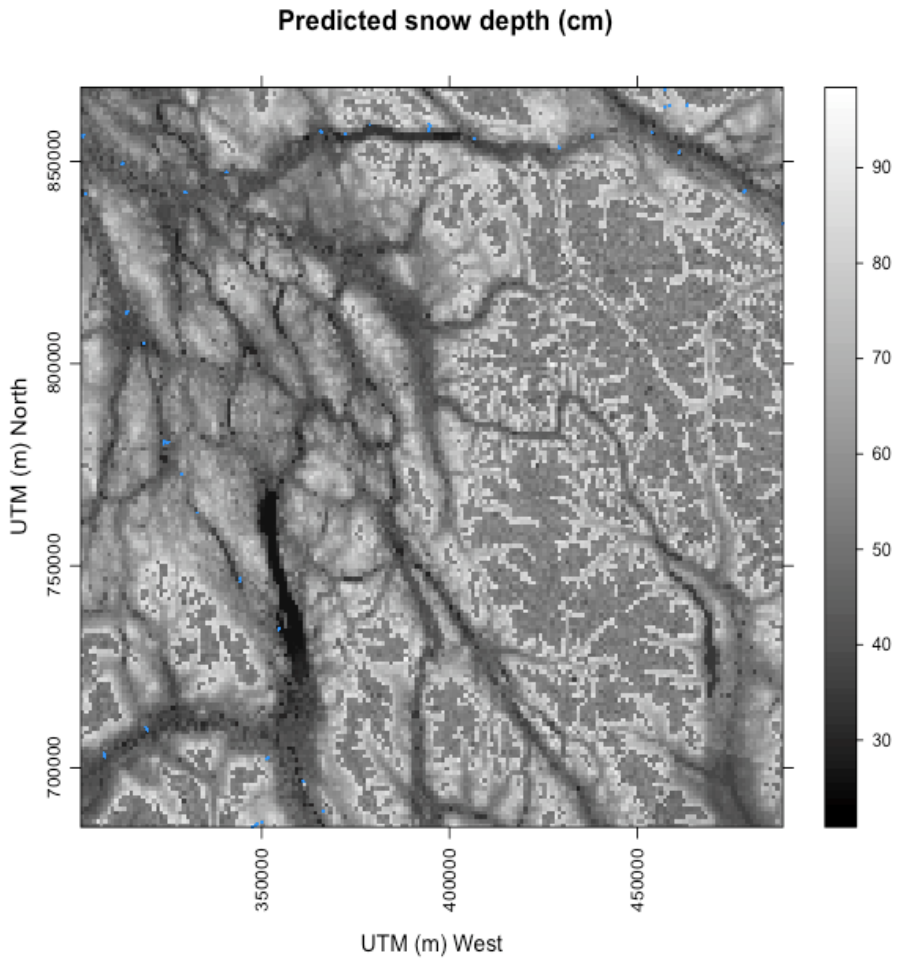



Figure 7: Predicted snow depth for central Yukon in cm. Coordinates are UTM map coordinates in meters. Actual survey sites are shown as blue dots.

The predicted snow depth from the model is presented in figure 7. Areas in high mountains (right hand region of the map) were well outside the sampling range of elevation in the model and are thus highly uncertain, and should not be relied upon.

**Snow water equivalent**

The common approach to modelling snow water equivalent (SWE) when snow depth is the primary measured variable, is to form a linear regression between SD and SWE for each area of the model, such as sub-basins or transects. This was justified as SWE has been found to be much less spatially variable than SD (Leydecker et al., 2001; Derksen et al., 2005; Erxleben et al., 2002; Winstral et al., 2002; Plattner et al., 2006). This approach was performed in this study, however when the fitted SWE was used as the predicted variable in the regression the quality of the fit decreased. In addition, the confidence intervals for the model degrade further.

## DISCUSSION AND CONCLUSIONS

**SD model interpretation**

Through mixed effects regression a snow depth and SWE map for the central Yukon has been produced. In constructing this map, the assumptions required of the mixed effects model have not been violated, and so the final product is statistically robust. The model indicates that snow depth was primarily predicted by elevation, but that land cover, and shelter contribute.

Previous regressions with sampling less than the range of autocorrelation violate the independence assumptions (Erickson et al., 2005). All previous regressions with sampling with spacing greater than the autocorrelation range miss the small scale variation of snow depth or SWE, a major contributor of its sampling variation. Mixed effects modelling is a requirement to use regression, as evidenced in the different estimations produced by the GLS model. Some assumptions are still broken, however, it is believed to be in a non-critical fashion which only adds to the error of the estimate, rather than invalidating the findings.  is a poor measure for goodness of fit of an autocorrelated model as it will be inflated. However, its simplicity has encouraged its use in a wide variety of snow models, and so to allow inter-comparability, it has been included here.

Many studies conclude that SD is related to SWE, and its ease of measurement makes it the preferred observation variable in the field. Measuring SWE directly could increase the accuracy of the model, at a cost of vastly increased time (at least six times longer). SD is strongly related to elevation, as shown in a number of previous studies. This indicates the transferability of this result is high.

**Sources of Error**

There are several potential sources of error for this model. The most probable are an error in the model function, as insufficient data points were acquired to test higher order interaction terms or polynomial transformations, thus ideal model terms may not have been tested. The 8-10 year gap between image acquisition and use may have led to some sites being classified under different conditions than they were sampled. One Landsat scene (lower left, WRS2 coordinates 61-17) was hazy which produced different DNs for the transformations and classification.

A major probable misspecification in the model was the lack of a wind term. While the effects of wind on snow are mitigated by sheltering terrain (included via elevation range, horizontal and profile curvature, elevation standard deviation), and vegetation (both classification and transformation express increasing capture capability), the directional effects were not explored. For example, relying on confounders can only increase the complexity of the model.

Shrinkage, the over-fitting of regression coefficients, was not tested for directly. Examining the cross-validation estimates for the coefficients, the degree of variance indicates that over-fitting was present, but the small RMSE indicates that it is not large. Royston et al. (2008) summarize a number of studies on the topic of sample size. A sample of 10 events per variable is concluded as the minimum to avoid having significant shrinkage, however, multi-level studies were not

discussed. Pinheiro and Bates (2000) state that the pooling effect of MER should add some robustness against shrinkage. This study includes 14 variables, along with a transformation for 11 of them, so should have a minimum of 140-260 observations. If taken at the site level, there are 212 observations, which unlikely to be sufficient; however, approximately 20 measurements make up each observation. Due to this a selection bias might be expected, causing weakly correlated predictors to not be selected, and shrinkage, the mis-estimation of the model parameters to be present, although mild. These effects will be concentrated on weakly correlated variables.

**Further work**

Further work to improve the estimates of this model should include sampling SWE(density) more intensively, and locating the samples at the corners of the study site triangles rather than only at the center, so as to improve the within-site scale variance estimate. (effectively only the 1-m and 100-m variance were measured, not the 30m) see Watson et al. (2006)

Furthermore, the model should include some measure of wind directly, as it has frequently been kept in regression models by other authors. Its inclusion makes increasing physical sense at higher altitudes where wind scouring and drifting play an increasingly important role (Elder et al., 1991).

Finally, additional surveying of snow depth and density to increase the sample size. This would allow us to increase the number of possible regressors or interactions considered, and increase the accuracy of the prediction. Doing so will vastly improve the usefulness of these results to other researchers.

## REFERENCES


Anderton S. White S. Alvera B. 2004. Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes*, **18(3)**: 435–453.

Brabets T, Wang B, Meade R, Geological Survey AWRDD Anchorage. 2000. Environmental and hydrologic overview of the Yukon River Basin, Alaska and Canada. Tech. Rep. 99–4204, United States Geological Survey.

Carroll S, Cressie N. 1997. Spatial modeling of snow water equivalent using covariances estimated from spatial and geomorphic attributes. *Journal of Hydrology.* **190(1-2)**: 42–59.

Cline D, Armstrong R, Davis R, Elder K, Liston G. 2001. NASA Cold Land Processes Field Experiment Plan 2002-2004. *Cold Land Processes Working Group, NASA Earth Science Enterprise. Land Surface Hydrology Program.* 202.

Demidenko E. 2005. *Mixed models: theory and applications*. Wiley-Interscience.

Derksen C. Walker A. Goodison B. Strapp J. 2005. Integrating in situ and multiscale passive microwave data for estimation of subgrid scale snow water equivalent distribution and variability. *IEEE Transactions on Geoscience and Remote Sensing.* **43(5)**: 960–972.

Elder K, Dozier J, Michaelsen J. 1991. Snow accumulation and distribution in an alpine watershed. *Water Resour. Res* **27(7)**: 1541–1552.

Elder K, Rosenthal W, Davis RE. 1998. Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes* **12(10-11):** 1793–1808.

Erickson T, Williams M, Winstral A. 2005. Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. *Water Resources Research* **41(4)**: W04014.

Erxleben J, Elder K, Davis R. 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. *Hydrological Processes* **16(18)**.

Janowicz J, Joe-Strack R. 2008. Yukon snow survey bulletin and water supply forecast. Tech. rep. Environment Yukon - Water Resources Branch.

Jensen JR. 2007. *Remote sensing of the environment: an earth resource perspective*. Prentice Hall, 2 ed.

Jost G, Weiler M, Gluns DR., Alila Y. 2007. The influence of forest and topography on snow accumulation and melt at the watershed-scale. *Journal of Hydrology* **347(1-2)**: 101 – 115.

Lapena DR, Martz LW. 1996. An investigation of the spatial association between snow depth and

topography in a prairie agricultural landscape using digital terrain analysis. *Journal of Hydrology* ***184*****(3-4):** 277 – 298.
Leydecker A, Sickman JO, Melack JM. 2001. Spatial scaling of hydrological and biogeochemical aspects of high-altitude catchments in the Sierra Nevada, California, U.S.A. *Arctic, Antarctic, and Alpine Research* ***33*****(4):** 391–396.
Lopez-Moreno J, Nogues-Bravo D. 2006. Interpolating local snow depth data: an evaluation of methods. *Hydrological Processes* ***20*****(10)**.
López-Moreno J, Stähli M. 2008. Statistical analysis of the snow cover variability in a subalpine watershed: Assessing the role of topography and forest interactions. *Journal of Hydrology* ***348*****(3-4):** 379–394.
Luce C. Tarboton D. Cooley K. 1999. Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model. *Hydrological Processes* ***13*****(12)**: 1921–1933.
Molotch N. Colee M. Bales R. Dozier J. 2005. Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrological Processes* ***19*****(7)**.
Pinheiro JC, Bates DM. 2000. *Mixed-effects models in S and S-PLUS*. Springer-Verlag.
Plattner C. Braun LN. Brenning A. 2006. The spatial variability of snow accumulation on Vernagtferner, Austrian Alps, in winter 2003/2004. *Zeitschrift für Gletscherkunde und Glazialgeologie* ***39:*** 43–57.
Pomeroy WJ, Gray MD. 1995. *Snowcover: Accumulation, Relocation and Management.*, vol. 88. Supply and Services Canada, Saskatoon.
Royston P. Sauerbrei W. 2008. *Multivariable model-building: A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. John Wiley.
Stahli M. Schaper J. Papritz A. 2002. Towards a snow-depth distribution model in a heterogeneous subalpine forest using a Landsat TM image and an aerial photograph. *Annals of Glaciology* ***34:*** 65–70.
Trujillo E. Ramirez JA. Elder KJ. 2009. Scaling properties and spatial organization of snow depth fields in sub-alpine forest and alpine tundra. *Hydrological Processes* ***23*****(11)**: 1575–1590.
Watson F. Newman W. Coughlan J. Garrott R. 2006. Testing a distributed snowpack simulation model against spatial observations. *Journal of Hydrology* ***328*****(3-4)**: 453–466.
Watson FGR. Anderson TN. Newman WB. Alexander SE. Garrott RA. 2006. Optimal sampling schemes for estimating mean snow water equivalents in stratified heterogeneous landscapes. *Journal of Hydrology* ***328*****(3-4, Sp. Iss. SI)**: 432–452.
Wilson J. Gallant J. 2000. *Terrain analysis: principles and applications*. Wiley.
Winstral A. Elder K. Davis R. 2002. Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology* ***3*****(5)**: 524–538.