Comparing Design Ground Snow Load Prediction in Utah and Idaho

Brennan Bean[1], Marc Maguire A.M.ASCE[2], and Yan Sun[3]

ABSTRACT

Snow loads in the western United States are largely undefined due to complex geography and climates, leaving the individual states to publish detailed studies for their region, usually through the local Structural Engineers Association (SEAs). These associations are typically made up of engineers not formally trained to develop or evaluate spatial statistical methods for their regions and there is little guidance from ASCE 7. Furthermore, little has been written to compare the independently developed design ground snow load prediction methods used by various western states. This paper addresses this topic by comparing the accuracy of a variety of spatial methods for predicting 50 year (i.e. design) ground snow loads in Utah and Idaho. These methods include, among others, the current Utah snow load equations, Idaho's normalized ground snow loads based on inverse distance weighting, two forms of Kriging, and the authors' adaptation of PRISM. The accuracy of each method is evaluated by measuring the mean absolute error using ten fold cross validation on datasets obtained from Idaho's 2015 snow load report, Utah's 1992 snow load report, and a new Utah ground snow load dataset. These results show that regression-based Kriging and PRISM methods have the lowest cross validated errors across all three datasets. These results also show that normalized ground snow loads, which are a common way of accounting for elevation in traditional interpolation methods, do not fully account for the effect of elevation on ground snow

[1]PhD Student, Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, Email: brennan.bean@aggiemail.usu.edu

[2]Assistant Professor, Department of Civil and Environmental Engineering, Utah State University, 4110 Old Main Hill, Logan, Utah 84321

[3]Assistant Professor, Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322

loads within the considered datasets. The methodologies and cautions outlined in this paper provide a framework for an objective comparison of snow load estimation methods for a given region as state SEAs look to improve their future design ground snow predictions. Such comparisons will aid states looking to amend or improve their current ground snow load requirements.

INTRODUCTION

Heavy snowstorms in the winter of 2017 filled local newspapers across the western United States with reports of snow related building collapses and fatalities (Lafferty 2017, Associated Press 2017, Mieure 2017, Kato and Florio 2017, Fisicaro 2017, Glover 2017). These snow-related failures can be catastrophic to local economies, like the recent \$100 million in losses incurred by Idaho/Oregon's onion industry (Ellis 2017). One study of 40 snow-induced building failures reported an average cost of \$166 per square meter and 122 days of business interruption for repairs (Strobel and Liel 2013). Snow-related damages can extend beyond building repairs, as a study of 1,100 domestic and international snow-induced building failures reported more than 300 fatalities (Geis et al. 2011). Few details are made public about the true causes of the above damages, as they could be agricultural buildings not designed to code or even suffer from construction error, but these reports and articles provide a sample of the serious consequences associated with snow load prediction.

Subtler costs are also associated with overly conservative load designs. As articulated by Nowak and Collins (2012): "Conceptually, we can design [a] structure to reduce the probability of failure, but increasing the safety... beyond a certain optimum level is not always economical." The following two examples demonstrate this point by exploring the relationship between design snow loads and roof construction costs. Roof costs are selected for these illustrations as they are likely the aspect of a structure most sensitive to snow load design.

The first example is found in the 2017 Craftsman National Building Cost Manual, which includes a table of estimated roof costs for manufactured homes rated for different snow loads. In this manual, a doubling of the roof snow load requirement from 1.44 to 2.88 kPa results in an approximate threefold increase in the estimated cost per unit meter of roof (\$11 to \$36) (Moselle

2016). While increases in cost outside the selected load range are not quite as drastic, the example demonstrates the influence of snow loads on roof costs. The second example comes from roof joist costs provided to the authors by Vulcraft Utah (Brigham City, Utah) in January 2018. These roof-only designs assume varying snow loads with the constant depths, typical joist spacings and a L/240 deflection limit, as indicated in Figure 1. These costs do not include the effects of the snow and larger roof components on the remainder of the gravity or seismic systems' cost. For this system, doubling the roof snow load requirement from 1.44 to 2.88 kPa leads to a 40-90% increase in the cost of the joists.

These two examples may represent highly sensitive situations with respect to cost and snow load. Other systems and components would likely not experience such dramatic cost increases. Regardless, the potential economic burdens created by overly conservative requirements likely explain recently amended ground snow load requirements in Rich County, Utah, where new requirements for major communities in the county (approximately 2.73 kPa) are less than half of those dictated previously (6.3-7.2 kPa) (Utah 2016).

American Society of Civil Engineers (ASCE) design ground snow load requirements have historically remained largely unspecified for the topographically complex western states up through ASCE 7-10 (ASCE 2013). This had led to the creation of a diverse set of state specific snow load estimation methods (Sack 2015). New snow load tables provided for many of the western states in ASCE 7-16 are derived from these state snow load reports (ASCE 2017). Design ground snow loads are defined in this paper as estimated 50 year ground snow loads. With the exception of the reliability-based snow loads in Colorado (Torrents et al. 2016), this definition is consistent with ASCE-7 and western state snow load reports. Many of these reports (or portions of them) are freely available to the public (NACSE 2012, SEAU 1992, Torrents et al. 2016, Al Hatailah et al. 2015, Theisen et al. 2004) and provide a wealth of information on dataset development, model predictions, and implications for building design. However, little is written regarding the accuracy of the methods used to predict design ground snow loads. While Sack (2015) and Sack et al. (2016) discuss differences between state methodologies and acknowledge discrepancies in predictions along state boundaries, no formal comparison of design ground snow load prediction methods is found in the literature. A lack of accuracy comparisons makes it difficult to determine whether differences in design ground snow load requirements along state boundaries are caused by differences in methodology, data, or both.

This paper begins such comparisons by determining the cross validated accuracies of several design ground snow load prediction methods on three independently developed datasets. These cross validation results are calculated using the R statistical software environment (R Core Team 2017) and visualized with the ggplot2 (Wickham 2009) and RColorBrewer (Neuwirth 2014) package extensions. These results will be preceded by a summary of the datasets and spatial prediction methods used in the comparisons and followed by a discussion of the challenges and limitations in predicting design ground snow loads. The authors conclude that regression-based spatial estimators that model the log-linear relationship between ground snow load and elevation consistently outperform all other methods in terms of minimizing the cross validated mean absolute error (MAE). Cross validation also highlights some of the limitations of normalized ground snow loads (NGSL), as explained in the "Prediction Comparisons" section of this paper. These results in Utah and Idaho provide a framework for a formal comparison of methodologies used by each of the western states, an important step for states looking to amend or improve their current ground snow load requirements.

DATA

The three datasets used in the cross validation comparisons are the authors' new Utah dataset (UT-2017), the 1992 Utah snow load report dataset (UT-1992) and the 2015 Idaho snow load report dataset (ID-2015). The variable of interest in each dataset is the design ground snow load. These design ground snow loads are calculated by fitting the annual maximum snow water equivalents (SWE) at each station location to a probability distribution and extracting the 98th percentile. Nearly all low elevation stations do not provide direct measurements of SWE. At locations where SWE is not measured, estimates of SWE are made from snow depth measurements using either the Rocky Mountain Conversion Density (RMCD) (Sack and Sheikh-Taheri 1986), or an equation developed by Sturm et al. (2010) referred to hereafter as "Sturm's equation". Table 1 provides an overview of the methods used to estimate design ground snow loads within each dataset. These readily available datasets were selected to compare the effectiveness of various spatial methods in predicting ground snow loads for different climates, terrain, and station coverage. In addition, the development of each of considered spatial method is associated with one of these datasets, including the current Utah snow load equations (UT-1992), Idaho's normalized ground snow loads based on inverse distance weighting (ID-2015), Kriging (UT-2017) and PRISM (UT-2017). The consideration of these three independently developed data sources ensures that the cross validation comparisons are not limited to one isolated dataset.

Each of these datasets use observations from Natural Resources Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) and Snow Course (SC) stations, as well as data from the National Weather Service (NWS) cooperative observer network (COOP). Many SNOTEL stations were installed to replace discontinued SC stations, thus creating situations where two separate stations have the same location. Identical decimal degree locations for two distinct stations creates singularity issues in many spatial interpolation methods. This problem was resolved by adding an arbitrarily small number r, ($|r| < .001$) to the decimal degree locations to create well defined but negligible spatial separation between such stations.

Figure 2 (a-c) reveals the distinct log-linear relationship between station design ground snow load estimates and elevation for each dataset. These scatterplots include lines representing ordinary and generalized least squares regression estimates of this log-linear relationship (using elevation as the predictor). The development of these regression lines will be discussed further in the "Methods" section of this paper. In addition, the histogram of station elevations in Figure 2 (d) show that the Idaho dataset contains a larger proportion of high elevation stations than either Utah dataset. Cross validated results must be interpreted in the context of station elevation, as higher elevations tend to have higher snow loads and consequently more variability in predictive accuracy.

The New Utah Dataset (UT-2017)

This dataset contains 279 (192 COOP, 87 SNOTEL) Utah stations with an additional 136 stations (103 COOP, 33 SNOTEL), all located within 100km of the Utah border. Log-normal distribution parameter estimates were calculated using annual yearly maximums for years 1970 to 2017 via maximum likelihood estimation. This range focuses on years where SNOTEL station measurements are available, as the earliest available measurements from active SNOTEL stations in Utah is 1978 (NRCS 2017). Sturm's equation estimated SWE from snow depth when SWE measurements were missing. This equation is defined using the coefficients for a "prairie" and "alpine" terrains (Sturm et al. 1995) as

$$SWE_i = \begin{cases} h_i\left[.3608 * (1 - \exp(-.0016h_i - .0031d_i)) + .2332\right] & \text{Elevation} < 2113.6\text{m} \\ h_i\left[.3738 * (1 - \exp(-.0012h_i - .0038d_i)) + .2237\right] & \text{Elevation} >= 2113.6\text{m} \end{cases} \quad (1)$$

where h_i represents snow depth (in centimeters) and d_i represents the day of the snow year, ranging from -92 (October 1) to 181 (June 30), for any given observation i (2010). See Bean et al. (2018) for a copy of this dataset along with further details regarding its creation.

The 1992 Utah Dataset (UT-1992)

These data consist of 413 stations (210 SC, 203 COOP), all located in Utah. The method used to calculate the Log-Pearson type III parameters is not specified. Estimates of SWE using the RMCD were occasionally adjusted when the resulting snow water equivalents exceeded the station's winter cumulative precipitation. Tables of these data can be found in the Utah snow load report (SEAU 1992).

The 1992 Utah report does not provide precise station locations. Since 1979, many of the snow course stations used in this report have been discontinued, and precise location information is unavailable. Station locations were determined for all but seven stations through a combination of station number matching in NRCS and NWS station databases, as well as personal contact with Randall Julander at the Utah Snow Survey Office in Salt Lake City. Locations for the seven remaining stations were approximated using Google Earth to determine coordinates given approximate station location information from the snow survey office and county information given in the 1992 Utah report.

The 2015 Idaho Dataset (ID-2015)

These data consist of 394 (246 SC/SNOTEL, 148 COOP) Idaho stations with an additional 257 (222 SC/SNOTEL, 35 COOP) located near the Idaho border with the most recent measurements being taken in 2014. Log-Pearson type III distribution parameter estimates were determined using the sample mean, standard deviation skew of annual of yearly maximums at each station location (i.e. method of moments). The data and further details regarding the estimation of these 50 year events are given in Al Hatailah's Masters Thesis (2015).

METHODS

Each of the following methods predict design ground snow loads at a state level using design ground snow loads at surrounding station locations as input. These methods were selected due to their ability to be easily applied to datasets of varying size and location, an important pre-requisite for calculating the cross validated errors discussed later in this paper. Details of the following methods can be found at citations provided in the respective summaries. For comparative convenience, the primary methods of consideration are defined using a common set of notation. Let $p_g(\boldsymbol{u})$ denote the ground snow load at a location $\boldsymbol{u}$ (with p_g^* representing the predicted design ground snow load) and let $A(\boldsymbol{u})$ denote location elevation. Further, let $\boldsymbol{u}_\alpha$ represent the location of station $\alpha\,(\alpha = 1, \cdots, N)$ and let $D(\boldsymbol{u}_i, \boldsymbol{u}_j)$ represent the geographic distance between locations $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$.

The defining feature of each method is in the way that elevation is accounted for in the design ground snow load predictions. With the exception of the ground snow load equations in the 1992 Utah Snow load report, each of the considered methods use normalized ground snow loads (NGSL) or some variant of linear regression. Normalized ground snow loads (NGSL) are calculated as design ground snow load divided by elevation $\left(\frac{p_g^*(\boldsymbol{u}_\alpha)}{A(\boldsymbol{u}_\alpha)}\right)$. They "appear to mask out the effects of the environment on the snow-making mechanism" and "reduce the entire area to a common base elevation" (Sack et al. 2016). NGSL have a long history of use in western state snow load studies, including the current snow load reports of Idaho, Montana and Washington (Sack et al. 2016).

On the other hand, regression based estimators seek to characterize the log-linear relationship between design ground snow loads and elevation observed in Figure 2. This relationship can be characterized using simple linear regression (LR) defined as

$$\log(p_g^*(\boldsymbol{u})) = \beta_0 + \beta_1 A(\boldsymbol{u}) \tag{2}$$

where β_0 and β_1 are calculated using ordinary least squares regression. The cross validated results in the following section show that differences in method accuracy can be largely attributed to differences in the characterization of the elevation/snow load relationship.

Current Utah Ground Snow Load Equations

The Structural Engineers Association of Utah (SEAU) predict design ground snow loads from elevation using the following equation (referred to hereafter as SNLW):

$$p_g^*(\boldsymbol{u}) = \begin{cases} \left(P_0^2 + S^2\left(A(\boldsymbol{u}) - A_0\right)^2\right)^{\frac{1}{2}} & A(\boldsymbol{u}) > A_0 \\ P_0 & A(\boldsymbol{u}) \leq A_0 \end{cases} \tag{3}$$

where P_0 (base ground snow load), S (change in ground snow load with elevation), and A_0 (base ground snow elevation) are parameters whose values are uniquely defined for each county. County specific parameters were selected to be "an approximate upper bound" to both the design ground snow loads and the maximum observed ground snow loads for the set of stations in and near the county of interest (SEAU 1992).

Recently amended snow load requirements for the state replace the equation estimates at select locations in Utah (Utah 2016). These updated requirements generally result in a reduction of ground snow loads when compared to the original equation estimates (Bean et al. 2017).

Inverse Distance Weighting

In Idaho's normalized ground snow loads based on inverse distance weighting (IDW), the predicted ground snow load at a particular location is a weighted average of the NGSL of surrounding stations, multiplied by the location's elevation. This prediction is expressed as

$$p_g^*(\boldsymbol{u}) = \frac{A(\boldsymbol{u})}{\sum_{\alpha=1}^{N} D\left(\boldsymbol{u}_\alpha, \boldsymbol{u}\right)} \sum_{\alpha=1}^{n} \left[\left(\frac{1}{D\left(\boldsymbol{u}_\alpha, \boldsymbol{u}\right)} \right)^c \frac{p_g^*(\boldsymbol{u}_\alpha)}{A(\boldsymbol{u}_\alpha)} \right]. \tag{4}$$

The variable c allows for adjustments to the weighting factor, with larger values of c further reducing the influence of stations far away from the area of interest. The Idaho snow load report uses $c_1 = 2$ for locations with elevations below 1,219 m (4,000 ft) and $c_2 = 6$ for locations with elevations above $l = 1,219m$ (Al Hatailah et al. 2015). The cross validation results for all three datasets in the following sections use these parameter values. One key difference in this implementation is the use of geographic distances rather than euclidean distances from the Idaho Transverse Mercator Projection (Al Hatailah 2015). The use of geographic distances eliminates the spatial distortion that may occur when applying a euclidean based map projection to a large geographical area.

Linear Triangulation Interpolation

In linear triangulation interpolation (TRI), the area of interest is partitioned into a set of non-intersecting triangles with vertices at each station location. Predictions use a weighted average of the NGSL at the three stations forming the triangle overlaying the point of interest (Akima 1978). The R implementation of this strategy creates an entire grid of predicted values within the convex hull of the given data points (Akima and Gebhardt 2015). There are instances during cross validation when the convex hull of the training set does not encompass points in the test set, resulting in missing value predictions. These missing values are currently ignored when computing cross validated errors. These missing value predictions would need to be addressed prior to any serious consideration of this method in future work.

PRISM

PRISM (Parameter-elevation Relationships on Independent Slopes Model) uses weighted least squares regression to account for additional climatological factors in response variable predictions (Daly et al. 2002, 2008). This leads to an extension of Equation 2 with the form

$$\log(p_g^*(\boldsymbol{u})) = \beta_0(\boldsymbol{u}, \boldsymbol{X}) + \beta_1(\boldsymbol{u}, \boldsymbol{X})A(\boldsymbol{u}) \tag{5}$$

where $\beta_0(\boldsymbol{u}, \boldsymbol{X})$ and $\beta_1(\boldsymbol{u}, \boldsymbol{X})$ are estimated via weighted least squares regression. Final predictions exponentiate the log-scale predictions. The regression weights are a function of several factors defined in this adaptation of the algorithm as

$$\boldsymbol{W}(\boldsymbol{u}, \boldsymbol{X}) = \boldsymbol{W}_c \left[F_d \boldsymbol{W}_d^2 + F_z \boldsymbol{W}_z^2\right]^{\frac{1}{2}} \boldsymbol{W}_b, \tag{6}$$

where $\boldsymbol{X}$ is the matrix containing all station meta-data and

- $\boldsymbol{W}_c$ - a cluster factor (stations distributed in a tight cluster and similar in elevation receive less weight)
- $\boldsymbol{W}_d$ - distance weighting (stations closer to the area of interest receive more weight)
- $\boldsymbol{W}_z$ - elevation weighting (stations with altitudes similar to the area of interest receive more weight)
- $\boldsymbol{W}_b$ - basin weighting (stations located in the same water basin as the area of interest receive more weight)
- F_d and F_z - importance factors for distance weighting and elevation respectively

These weights create a unique linear model fit for each area of interest. For details regarding the calculation of these weights, refer to (Bean et al. 2017) with one noted difference. Originally the basin weights compared similarities in station watersheds from the United States Geologic Survey (USGS) Hydrologic Unit Codes (HUC) 2-12 (USGS 2016). These finest watershed levels (HUC 10 and 12) proved too small to be relevant in the weighting scheme, as nearly every station had its own HUC 12 designation. This in mind, the basin weighting function now only detects more coarse water basin associations in the following manner:

$$\boldsymbol{W}_{b_\alpha} = \left(\frac{s_\alpha + 1}{5}\right)^c, \tag{7}$$

where s represents the number of common watersheds (*four* levels ranging from HUC 2 through 8) shared by station α and the target grid cell, and c is a user defined weighting factor that changes the shape of the weighting function.

Kriging

The family of Kriging estimators leverage the spatially dependent correlations between observations to make predictions. The gstat package extension of R (Pebesma 2004) provides a numerical implementation of many Kriging variations. Details regarding these family of estimators are given in Goovaerts (1997), and motivate the notation used in this paper. One Kriging extension of Equation 2 is called Simple Kriging with varying Local Means (SKLM) (Goovaerts 2000) defined symbolically as

$$\log(p_g^*(\boldsymbol{u})) = \beta_0 + \beta_1 A(\boldsymbol{u}) + \sum_{\alpha=1}^{N} \lambda_\alpha(\boldsymbol{u}) r(\boldsymbol{u}_\alpha). \tag{8}$$

This method proceeds in three steps. First, a linear model is calculated identical to Equation 2. Then, simple kriging uses the residuals of the linear model to predict a residual value at the location of interest. Finally, this residual value is used to update the original linear model prediction. The simple kriging coefficients ($\lambda_\alpha(\boldsymbol{u})$) are calculated by solving the kriging system

$$\sum_{\alpha=1}^{N} \lambda_\beta(\boldsymbol{u}) C_R\left(D(\boldsymbol{u}_\alpha, \boldsymbol{u}_\beta)\right) = C_R\left(D(\boldsymbol{u}_\alpha, \boldsymbol{u})\right) \quad \beta = 1, \cdots, n. \tag{9}$$

where C_R represents the covariance between any two observations and is assumed to be a function of distance.

An alternative method for accounting for elevation in kriging predictions is through universal kriging (UK), which calculates the trend implicitly within the kriging system, rather than separately as in SKLM (Goovaerts 1997). When elevation is the only trend coefficient, the universal kriging estimates are equivalent to

$$\log(p_g^*(\boldsymbol{u})) = \beta_0^* + \beta_1^* A(\boldsymbol{u}) + \sum_{\alpha=1}^{N} \lambda_\alpha(\boldsymbol{u}) r(\boldsymbol{u}_\alpha) \tag{10}$$

where β_0^* and β_1^* are calculated using generalized least squares regression. Figure 2 shows the difference in the trend lines resulting from SKLM and UK.

The semivariogram (i.e. variogram) is inversely related to the covariances between observations and provides the covariance matrix necessary for generalized least squares regression. A theoretical variogram function approximates the empirical variogram defined in this case as

$$\hat{\gamma}(\boldsymbol{h}) = \frac{1}{2N_{\boldsymbol{h}_1}} \sum_{\alpha_{\boldsymbol{h}}=1}^{N_{\boldsymbol{h}}} \left[r(\boldsymbol{u}_{\alpha_{\boldsymbol{h}_1}}) - r(\boldsymbol{u}_{\alpha_{\boldsymbol{h}_2}}) \right]^2 \tag{11}$$

where $\left[r(\boldsymbol{u}_{\alpha_{\boldsymbol{h}_1}}), r(\boldsymbol{u}_{\alpha_{\boldsymbol{h}_2}}) \right]$ represents each pair of regression model residuals located $||\boldsymbol{h}||$ distance away from each other. Figure 3 provides an example of the empirical and associated theoretical variograms for each dataset. It is the theoretical variogram that determines the values of the covariance function given in Equation 9.

Kriging predictions provide theoretical estimates of the prediction error uncertainty (often called kriging variance) (Moral 2010). A better understanding of prediction uncertainty could be used to make conservative adjustments to snow load predictions in volatile areas. Because error uncertainty cannot be compared across all methods, the authors leave the discussion of kriging variance as applied to snow load predictions to future work.

CROSS VALIDATION

This paper now proceeds with a comparison of the predictive accuracies of the previously described methods. Two common ways of measuring method accuracy are with new test data or cross validation. Test set error measures model accuracy on new observations not used in model fitting, which is often impractical as available observations beyond those used in model fitting are scarce. Cross validation seeks to approximate test set error without requiring additional data. This is done by randomly dividing the given observations into groups, using all but one of these groups to fit a model that makes predictions for the remaining group. This process is then repeated until a prediction is made for each observation in the dataset. In this paper, the data are separated into ten groups. Cross validation is a common tool used for model selection and refinement in many disciplines (Arlot and Celisse 2010), including structural engineering (Chang et al. 2017) and will be used to compare the spatial prediction methods defined in the preceding section.

The use of cross validation is limited to replicable methods that are separable from the input observations. For example, snow load predictions in the Colorado snow load report involve a contour map of input parameter values that includes allowed discontinuities along mountain ridges (Torrents et al. 2016). These contours and discontinuities are inextricably connected to the station observations and thus eliminate the option to use cross validation. In addition, the Montana and Oregon snow load reports do not include enough details to replicate their methods on new datasets (Theisen et al. 2004, NACSE 2012). Because of these limitations, the accuracy comparisons for snow load prediction methods in these states are not included in the following results.

Cross validated errors are defined as

$$E(\boldsymbol{u}_\alpha) = \hat{P}_g(\boldsymbol{u}_\alpha) - P_g(\boldsymbol{u}_\alpha) \tag{12}$$

where $\hat{P}_g(\boldsymbol{u}_\alpha)$ and $P_g(\boldsymbol{u}_\alpha)$ are the predicted and actual ground snow loads at station location $\boldsymbol{u}_\alpha$ respectively. Defined in this way, a positive error indicates over-predictions and a negative error indicates under-predictions. These errors are heteroskedastic and occasionally very large as observed in Figure 4.

Overall comparisons of method accuracies are measured with mean absolute error (MAE) and mean error (ME) defined similarly in Maguire et al. (2014) as

$$
\begin{aligned}
\text{MAE} &= \frac{1}{N}\sum_{\alpha=1}^{N} |E(\boldsymbol{u}_\alpha)| \\
\text{ME} &= \frac{1}{N}\sum_{\alpha=1}^{N} E(\boldsymbol{u}_\alpha)
\end{aligned}
\tag{13}
$$

where N represents the total number of weather stations with ground snow load measurements and $\hat{P}_g(\boldsymbol{u}_\alpha)$ represents model predictions for each station location $\boldsymbol{u}_\alpha$.

Parameter Selection

Many of the parameters associated with the previously described spatial prediction methods must be manually specified. In practice, values of these parameters are selected by cross validation. To illustrate such a procedure, Table 2 of Bean et al. (2017) selected PRISM parameters as follows:

- Create a vector of possible values for each of the eight PRISM parameters using recommendations from Daly et al. (2002).
- For every possible combination of the parameters, fit the PRISM model and record the prediction error (such as MAE) resulting from cross validation.
- Select a parameter combination that minimizes the prediction error.

Each dataset uses the log-PRISM parameters provided in Table 2 of Bean et al. (2017) during cross validation.

In addition, each dataset uses the Kriging variogram developed for UT-2017, rather than the dataset-specific variograms shown in Figure 3. Cross validation quantifies the effect of using a single variogram for predictions on ID-2015 and UT-1992. The MAE for ID-2015 and UT-1992 using the dataset-specific variograms in Figure 3 are within 0.01 kPa of the MAE using the UT-2017 variogram (as averaged over 100 iterations of cross validation). Such results show that the cross validated errors are fairly insensitive to modest changes in the theoretical variogram for SKLM and UK on the considered datasets.

Error and Elevation

The locally weighted regression (loess) (Cleveland and Devlin 1988) curves in Figure 5 reveal the elevation dependent structure of the error scatter-plots previously shown in Figure 4. These curves compute local weighted averages of raw and absolute station errors across elevation and map these local averages as smooth polynomial curves. The gray tick marks drawn between each set of plots represent the elevations of the individual stations locations. These tick marks help to visualize station density across elevation. This characterization of density gives context to plotted curves, as the loess estimates will be more reliable at elevations with a higher density of stations.

Figure 5 shows that PRISM, SKLM, and UK are fairly unbiased at low elevations (2000 meters or less) and tend to under-predict at higher elevations (2000 - 3000 meters). The errors of all methods are very unstable in ID-2015 at high elevations. The sinusoidal shape of the ME curves for IDW reveal the tendency of this method to over-predict design ground snow loads at low elevations and under-predict at high elevations. This behavior is a result of the relationship between NGSL and elevation as discussed in the "Practical Limitations" section of this paper. Finally, Figure 5 shows the strong tendency of SNLW to over-predict design ground snow loads. In terms of relative errors, the Utah equations on average predict design ground snow loads 34% higher than station design ground snow load estimates from UT-2017 and 57% higher than estimates from UT-1992 (with median relative errors of 25% and 41% respectively). Recall that Equation 3 was intentionally designed to over-predict design ground snow loads, and it is no surprise that this method would have higher cross validated errors when compared to models designed to minimize error. However, these accuracy comparisons are still useful as they quantify the magnitude of the over-prediction of design ground snow loads using SNLW. Such over-predictions are understandable when considering the consequences of under-predictions discussed earlier in this paper. However, reliability-based engineering widely holds that snow load estimates should be as accurate and reliable as possible, with conservative adjustments being made to load predictions through the selection of load factors from a proper reliability analysis (Nowak and Collins 2012).

Accuracy Comparisons

Cross validated error measurements are partially subject to the random separation of observations into groups. To account for this randomness, cross validation is performed 100 times, recording the MAE and median absolute error (Med-AE) for each method at every iteration. The large difference between MAE and Med-AE illustrates the skewness invoked by the exceptionally high prediction errors that occasionally occur at high elevation locations. Figure 6 visualizes the average MAE and Med-AE of the 100 cross validation iterations. Black whiskers on each bar indicate the minimum and maximum MAE and Med-AE for the 100 iterations.

Figure 6 shows that PRISM, SKLM and UK notably outperform all other methods on both Utah datasets, with an MAE approximately 40-45% lower than SNLW and IDW on UT-2017. These improvements are not as pronounced for ID-2015, likely due to the less pronounced log-linear relationship between ground snow loads and elevation. However, the accuracy UK on ID-2017 remains notably better than all other methods, highlighting the accuracy improvements associated with the universal Kriging paradigm. These results demonstrate the accuracy improvements offered by PRISM and Kriging when compared to current snow load estimation methods used in Idaho and Utah. More importantly, the methodology used to obtain these results provides a pattern for comparing all snow load estimation methods used in the western states. Using cross validation in future snow load studies will provide state and federal officials with a universal and defensible standard for final model selections. Such a standard will ultimately improve the ground snow load estimation methods used across the region.

PRACTICAL LIMITATIONS

Figure 7 compares the current ground snow requirements in Utah to the predictions of PRISM, UK, and IDW. This comparison is an extension of a similar comparison provided in Bean et al. (2017). In many cases, the current predictions lead to a reduction in ground snow load requirements, with some major reductions occurring in places like Kamas, UT. In other cases, each of the methods recommended increased to the ground snow load requirements like Monticello, UT.

It is critical that these predictions and the previously discussed accuracy comparisons be placed in the context of observational limitations. These predictions rely on accurate estimates of design ground snow loads within each dataset and there is no guarantee that predictive accuracy for terrain not represented in the input datasets will be comparable to cross validated accuracies reported previously. The following subsections discuss some of the inevitable limitations associated with predicting design ground snow loads.

Limitations of Regression-Based Estimators

There are extrapolation issues for the regression based estimators (PRISM, SKLM, UK, and LR) when attempting to predict snow loads at locations with elevations far exceeding all nearby station elevations. In Utah, these situations most often occur at mountain peaks lacking station measurements. In such cases, these estimators begin to predict unreasonably high snow load values, exceeding all observed snow load values in the dataset. This issue is resolved by restricting the PRISM, Kriging, and IDW predictions to extend no higher than the largest design ground snow load in the input dataset. In addition, the prediction of the global trend (as used in both kriging estimators and linear regression) is not allowed to extend beyond the predicted trend for the highest elevation station in the dataset. Such constraints are only imposed when predicting at the state level and are not imposed for the cross validation results presented in this paper.

Limitations of NGSL-Based Estimators

Figure 7 reveals an alarming IDW prediction that is more than double the other method predictions at Farmington, Utah (elevation 1316 meters). As observed in Table 2, three of the four stations nearest to Farmington are all located at elevations above 2000 meters with NGSL several times higher than the NGSL of the low elevation station. This results in a likely over-prediction of the design ground snow load at Farmington and highlights a key shortcoming of using NGSL to account for elevation. This shortcoming is due to the strong positive correlation between elevation and the log transform of NGSL at station locations in Utah as observed in Figure 8. This correlation leads to the sinusoidal error patterns for IDW observed previously in Figure 5. If NGSL fully accounted for the effect of elevation on ground snow loads, then NGSL should be independent of elevation with a non-significant correlation coefficient. However, the Pearson correlation coefficient associated with

Figure 8 is 0.63, which is highly statistically significant (p value < 0.0001). This correlation can be insignificant globally yet significant locally. For example, the overall Pearson correlation between elevation and log-NGSL on ID-2015 is only 0.14 (p value = .0004), while the Pearson correlation for stations located in 12 counties comprising the south-eastern corner of the state, is 0.71 (p value < .0001). The separation of locations into high and low elevation layers partially mitigates this effect. For example, the 1219 m separating elevation used in the Idaho report (Al Hatailah et al. 2015) results in non-significant log-NGSL/Elevation correlations (i.e. p value > .01) of 0.05 and 0.35 for the low and high elevation layers in the ten counties comprising the Idaho panhandle. However, this same separating elevation fails to eliminate the strong 0.71 correlation observed for the 12 south-eastern counties, as all stations in this region are located in the upper elevation layer.

The prediction patterns associated with NGSL observed in Figures 5 and 7 are likely to occur in topographically complex regions where the correlation between NGSL and elevation is strong. Recalling the cost implications shown in Figure 1, differences in ground snow load prediction similar in magnitude to those observed at Farmington, Utah could easily double or triple the cost of the roof of a structure at these locations if this issue is not recognized and addressed.

Limitations of 50 Year Estimates

When fitting probability distributions to annual SWE maximums to predict 50 year ground snow load events, the convergence rate of the estimated parameters via maximum likelihood is on the order of $O_p\left(n^{-\frac{1}{2}}\right)$ (Casella and Berger 2002). This means that a fourfold increase in the sample size will reduce the estimation error by roughly half. However, the sample size necessary to achieve an acceptable level of error will vary for every research project. The minimum number of yearly observations required for distribution fitting were twelve in UT-2017, ten in ID-2015 (Al Hatailah et al. 2015), and seven in UT-1992 (SEAU 1992). These relatively small thresholds for the distribution fitting process reflect practical efforts on the part of researchers to produce reasonable 50 year estimates at stations with short periods of record. However, distributions fit with only ten or so years of record are likely attempting to predict 50 ground snow loads with magnitudes larger than all observations in the period of record.

Even with an "adequate" sample size, the inherently messy nature of real data (outliers, missing values, inaccurate measurements, and poor estimates of snow density from snow depth) adds uncertainty to 50 year estimates. In addition, potential violations of two assumptions inherent to the distribution fitting process add additional uncertainty to 50 year estimates. The first assumption is that the yearly maximums at each station all come from the same distribution, implying that the measurement conditions at each station location remain constant over the life of the station. Documented changes in measurement tools, sampling site conditions, and human influence (Julander and Bricco 2006) bring this assumption into doubt. The second assumption is that the yearly maximums are statistically independent, implying that snow measurements at each station location are uncorrelated across time. However, there is a wealth of evidence that suggests that time cannot be ignored when measuring climatic events. Researchers claim that the proportion of precipitation falling as snow in Utah has declined by nine percent in the last half century, accompanied by long term decreases in overall snow cover (Gillies et al. 2012). This agrees with multiple sources indicating that yearly snow packs are declining across the Pacific Northwest (Mote 2006, Scott and Kaiser 2004). These sources indicate that the assumption of independence between yearly maximums is most likely violated. These unaccounted sources of uncertainty are important to acknowledge but difficult to quantify. The effect of such uncertainties will inevitably become more prevalent when trying to predict recurrence intervals beyond 50 years, such as those explored in Debock et al. (2017). Further work is required to determine precise influence these assumption violations have on station ground snow load estimates.

One way to illustrate the effect of these uncertainties is through a comparison of estimated 50 year ground snow loads for COOP station USC00109638 in Weiser, Idaho (NOAA 2017). This station was selected due to the series of snow related collapses occurring in Weiser during the winter of 2017, where ground snow loads were estimated to be as high as 1.89 kpa (Arcement 2017). The reader should be cautioned that the reported collapses could be due to any number of factors (design, construction, etc.), not just snow load prediction. The authors can not comment on the safety of those structures, only to illustrate the uncertainty in 50 year ground snow load based on the distribution and SWE prediction. Station records at this location extend as far back as 1912. Data from this station were processed using the same procedures and filters used in the creation of UT-2017, resulting in a sample size of 73 yearly maximum snow loads. The normal, log-normal, gumbel and generalized extreme value (GEV) distributions each predict the 50 year ground snow load estimate at this location, the latter two distributions being fit using the extRemes package (Gilleland and Katz 2016). Efforts to fit a log-Pearson type III distribution via maximum likelihood estimation were non-convergent and thus were excluded from the comparison. Each distribution was fit twice: once using Sturm's equation to convert snow depths to SWE and again using the RMCD. Table 3 compares each of the resulting 50 year estimates to the 0.81 kPa 50 year ground snow load estimate in the Idaho snow load report.

Table 3 shows that different distributions can provide notably different estimates of 50 year events. The differences in distribution estimates shown in Table 3 are relatively larger than distribution comparisons at the Denver-Stapleton, Colorado snow site provided in DeBock et al. (2017). Perhaps more important, however, is the difference in 50 year predictions resulting from changes to the snow depth to SWE conversion method. Table 3 shows that, using the same distribution, SWE estimates using Sturm's equation results are more than 50% higher than design ground snow loads using RMCD. Differences of this magnitude are not unique to this particular station, but are most pronounced at low elevation locations such as Weiser. Table 4 shows the median absolute relative difference of 50 year estimates for 261 stations on UT-2017 relative to the original log-normal distribution estimates. Of the 415 stations, 120 stations were excluded as they did not require any SWE conversions and 21 stations were excluded for not having stable GEV 50 year estimates. These results confirm that differences in SWE conversion method are more influential on design ground snow loads than differences in distribution selection. These large differences reinforce the need for increased scrutiny in the process used to estimate design ground snow loads.

CONCLUSION

Great care has been taken by each of the western states to develop ground snow load prediction methods. However, little work has been done to formally compare the accuracy of these methods. This paper began a formal comparison of methods using cross validation to compare a variety of snow load prediction methods on three independently developed datasets for Utah and Idaho. The cross validation results show that both Kriging methods and PRISM were the most accurate (in terms of cross validated error) across all three datasets. For UT-2017, these methods had a 40-45% lower mean absolute error the current method used in Utah and Idaho. Further, the cross validation results show that UK performed the same as PRISM on the Utah datasets, but noticeably better than PRISM on the ID-2015, suggesting that Universal Kriging may be the best method for predicting ground snow loads across varying datasets. The relative ease of implementation for SKLM, UK, and PRISM demonstrate the feasibility of using these methods on a consolidated dataset to make predictions for multi-state regions. In addition, these prediction methods readily lend themselves to other SWE-based topics, especially when making predictions across time. For example, the authors have used PRISM to visualize changes in the water content of Utah's April 1st snow-pack from 1930-2015.

This paper also discussed the limitations underlying the current distribution based methods for estimating 50 year ground snow loads (or similar variants) at station locations. Comparisons of various distributions and snow load conversion methods in Tables 3 and 4 show that estimated design ground snow loads are very sensitive to changes in the SWE conversion method.

This in mind, the following conclusions can be made:

- SWE and distribution fitting assumptions provide differing design ground snow load station predictions by up to a factor of nearly 290% based on the case study in Weiser, Idaho and more than 40% on average when comparing stations from UT-2017.
- The top three considered methods (in terms of low cross validated MAE) account for log-linear relationship between ground snow loads and elevation. The improvements in cross validated accuracy using these methods was as much as 45% on UT-2017 when compared to the current prediction methods used in Idaho and Utah.
- Normalized ground snow loads (NGSL) do not fully remove the effect of elevation in spatial

- interpolation methods, with a Pearson correlation of 0.63 on UT-2017. This correlation, when present, leads to a tendency for IDW to over-predict snow loads at low elevations, and under-predict at high elevations.
- UK was similar in accuracy to PRISM on UT-2017 (MAE $\approx$ 0.9kPa) and UT-1992 (MAE $\approx$ 1.2kPa) and more accurate on ID-2015 (MAE $\approx$ 1.4kPa vs MAE $\approx$ 1.7kPa). Given its relative simplicity, well defined prediction variance, and robustness to differences in input data, the authors recommend Universal Kriging as the optimal method for predicting ground snow loads in Utah and Idaho.

The framework for cross validation outlined in this paper can be readily adapted for larger scale comparisons of snow load estimation methods across the country. Such a framework owes its existence to the individual efforts of many of the western states, which have provided numerous state-level ground snow load datasets for comparison. Leveraging these datasets for formal cross comparisons of methods will accelerate the development of new and better models as well as the improvement of existing ones. Consolidating the advancements made by each of the western states will continue to improve the consistency and reliability in design ground snow load estimates across the region.

ACKNOWLEDGEMENTS

The authors would like to thank the Structural Engineers Association of Utah, who provided partial funding of the research conducted in this paper. The authors would also like to thank the support of Bruce Brothersen and Jason Fisher at Vulcraft Utah for their help with contemporary local costs.

REFERENCES

Akima, H. (1978). “A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points.” *ACM Transactions on Mathematical Software (TOMS)*, 4(2), 148–159.

Akima, H. and Gebhardt, A. (2015). *akima: Interpolation of Irregularly and Regularly Spaced Data*, <https://CRAN.R-project.org/package=akima>. R package version 0.5-12.

Al Hatailah, H., Godfrey, B. R., Nielsen, R. J., and Sack, R. L. (2015). "Ground snow loads for Idaho–2015 edition.

Al Hatailah, H. A. (2015). "Ground Snow Loads for the State of Idaho." M.S. thesis, University of Idaho, Moscow, Idaho.

Arcement, K. (2017). "'a lot of scared people': Relentless snow collapses hundreds of Idaho roofs, devastates rural county." *The Washington Post*, <https://www.washingtonpost.com/news/morning-mix> (January). Accessed: 05-15-2018.

Arlot, S. and Celisse, A. (2010). "A survey of cross-validation procedures for model selection." *Statistics surveys*, 4, 40–79 Accessed: 2017-09-08.

ASCE (2013). *Minimum Design Loads for Buildings and Other Structures*. American Society of Civil Engineers, asce/sei 7-10 edition, <http://ascelibrary.org/doi/abs/10.1061/9780784412916>.

ASCE (2017). *Minimum Design Loads and Associated Criteria for Buildings and Other Structures*. American Society of Civil Engineers, asce/sei 7-16 edition, <http://ascelibrary.org/doi/abs/10.1061/9780784414248>.

Associated Press (2017). "Roof on Provo building collapses under snow weight, <https://www.ksl.com/?nid=148&sid=42776523> (January). Accessed: 05-15-2018.

Bean, B., Maguire, M., and Sun, Y. (2017). "Predicting Utah ground snow loads with PRISM." *Journal of Structural Engineering*, 143(9), 04017126.

Bean, B., Maguire, M., and Sun, Y. (2018). "The Utah snow load study, <https://digitalcommons.usu.edu/cee_facpub/3589/>. Accessed: 2017-09-08.

Casella, G. and Berger, R. L. (2002). *Statistical inference*, Vol. 2. Duxbury Pacific Grove, CA, 337,472.

Chang, M., Maguire, M., and Sun, Y. (2017). "Framework for mitigating human bias in selection of explanatory variables for bridge deterioration modeling." *Journal of Infrastructure Systems*, 23(3), 04017002.

Cleveland, W. S. and Devlin, S. J. (1988). "Locally weighted regression: an approach to regression analysis by local fitting." *Journal of the American statistical association*, 83(403), 596–610.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). "A knowledge-based approach to the statistical mapping of climate." *Climate research*, 22(2), 99–113.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P. (2008). "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology*, 28(15), 2031–2064.

DeBock, D. J., Liel, A. B., Harris, J. R., Ellingwood, B. R., and Torrents, J. M. (2017). "Reliability-based design snow loads. i: Site-specific probability models for ground snow loads." *Journal of Structural Engineering*, 04017046.

Ellis, S. (2017). "Snow damage to Idaho-Oregon onion industry nears $100 million, <http://www.capitalpress.com/Idaho/20170127/snow-damage-to-idaho-oregon-onion-industry-nears-100-million> (January). Accessed: 05-15-2018.

Fisicaro, K. (2017). "Snow-load removal began day before gym collapse." *The Observer*, <http://www.lagrandeobserver.com/newsroomstafflist/5033603-151/snow-load-removal-began-day-before-gym-collapse> (February). Accessed: 05-15-2018.

Geis, J., Strobel, K., and Liel, A. (2011). "Snow-induced building failures." *Journal of Performance of Constructed Facilities*, 26(4), 377–388.

Gilleland, E. and Katz, R. W. (2016). "extRemes 2.0: An extreme value analysis package in R." *Journal of Statistical Software*, 72(8), 1–39.

Gillies, R. R., Wang, S.-Y., and Booth, M. R. (2012). "Observational and synoptic analyses of the winter precipitation regime change over Utah." *Journal of Climate*, 25(13), 4679–4698.

Glover, J. (2017). "Wet, heavy snow in Spokane sparks concern that weak roofs could collapse, <http://www.spokesman.com/stories/2017/feb/06/wet-heavy-snow-in-spokane-sparks-concern-that-weak/#/0> (February). Accessed: 05-15-2018.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press.

Goovaerts, P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." *Journal of hydrology*, 228(1), 113–129.

Julander, R. P. and Bricco, M. (2006). "An examination of external influences imbedded in the historical snow data of Utah Accessed: 2018-05-15.

Kato, D. and Florio, G. (2017). "Montana theater's roof collapses under weight of snow, <http://billingsgazette.com/news/state-and-regional/montana/montana-theater-s-roof-collapses-under-weight-of-snow/article_5e61c0b7-c58d-5560-9b67-6a348da2d63f.html> (February). Accessed: 05-15-2018.

Lafferty, K. (2017). "Snow buildup on porch roof causes collapse fatally injuring Deary woman, <http://klewtv.com/news/local/snow-buildup-on-porch-roof-causes-collapse-fatally-injuring-deary-woman> (January). Accessed: 05-15-2018.

Maguire, M., Moen, C. D., Roberts-Wollmann, C., and Cousins, T. (2014). "Field verification of simplified analysis procedures for segmental concrete bridges." *Journal of Structural Engineering*, 141(1), D4014007.

Mieure, E. (2017). "Snow causes partial roof collapse at Sears, Axis, Hole Bowl." *Jackson Hole Daily*, <http://www.jhnewsandguide.com/jackson_hole_daily/local/snow-causes-partial-roof-collapse-at-sears-axis-hole-bowl/article_09e3f73d-cfcb-5be3-9462-a05f0fad1cce.html> (February). Accessed: 05-15-2018.

Moral, F. J. (2010). "Comparison of different geostatistical approaches to map climate variables: application to precipitation." *International Journal of Climatology*, 30(4), 620–631.

Moselle, B. (2016). *National Building Cost Manual 41st Edition*. Craftsman Book Company, Carlsbad, CA (October).

Mote, P. W. (2006). "Climate-driven variability and trends in mountain snowpack in western North America." *Journal of Climate*, 19(23), 6209–6220.

NACSE (2012). “An updated snow load map and internet map server for Oregon (May).

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer Palettes*, <https://CRAN.R-project.org/package=RColorBrewer>. R package version 1.1-2.

NOAA (2017). “Cooperative observer network (COOP). Accessed: 2017-09-07.

Nowak, A. S. and Collins, K. R. (2012). *Reliability of structures*. CRC Press.

NRCS (2017). “Active SNOTEL stations, <www.wcc.nrcs.usda.gov/snow/sntllist.html> (June).

Pebesma, E. J. (2004). “Multivariable geostatistics in S: the gstat package.” *Computers and Geosciences*, 30, 683–691.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.

Sack, R. L. (2015). “Ground snow loads for the western United States: State of the art.” *Journal of Structural Engineering*, 142(1), 04015082.

Sack, R. L., Nielsen, R. J., and Godfrey, B. R. (2016). “Evolving studies of ground snow loads for several western US states.” *Journal of Structural Engineering*, 04016187.

Sack, R. L. and Sheikh-Taheri, A. (1986). *Ground and roof snow loads for Idaho*. University of Idaho, Department of Civil Engineering.

Scott, D. and Kaiser, D. (2004). “5.2 variability and trends in United States snowfall over the last half century.

SEAU (1992). “Utah snow load study.

Strobel, K. and Liel, A. (2013). “Snow load damage to buildings: physical and economic impacts.” *Proceedings of the Institution of Civil Engineers-Forensic Engineering*, 166(3), 116–133.

Sturm, M., Holmgren, J., and Liston, G. E. (1995). “A seasonal snow cover classification system for local to global applications.” *Journal of Climate*, 8(5), 1261–1283.

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., and Lea, J. (2010). “Estimating snow water equivalent using snow depth data and climate classes.” *Journal of Hydrometeorology*, 11(6), 1380–1394.

Theisen, G. P., Keller, M. J., Stephens, J. E., Videon, F. F., and Schilke, J. P. (2004). “Snow loads for structural design in Montana.

Torrents, J. M., DeBock, D. J., Harris, J. R., Liel, A. B., and Patillo, R. M. (2016). “Colorado design snow loads.

USGS (2016). “Watershed boundary dataset, <https://nhd.usgs.gov/wbd.html>.

Utah Legislature (2016). *15A-3-107 Amendments to Chapter 16 of IBC*, <https://le.utah.gov/xcode/Title15A/Chapter3/15A-3-S107.html>. Accessed: 2018-05-15.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, <http://ggplot2.org>.

List of Figures

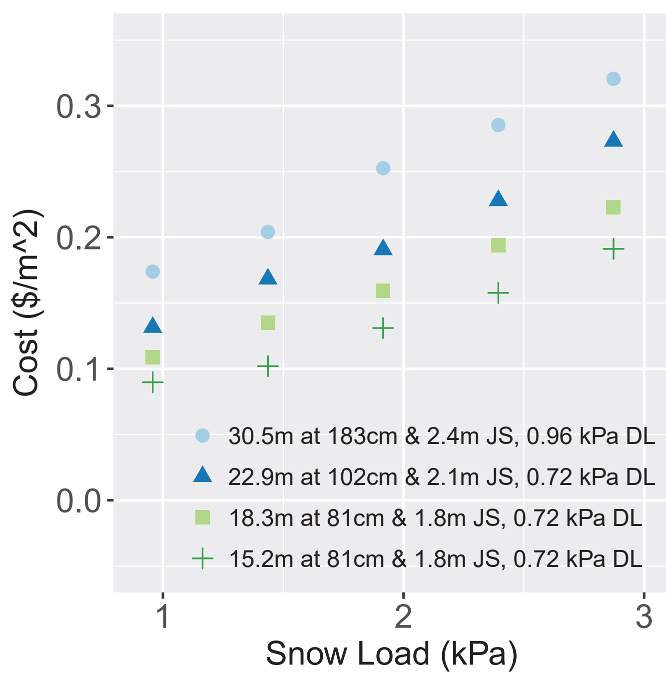

Fig. 1. Cost to snow load comparison for five different roof joist types.

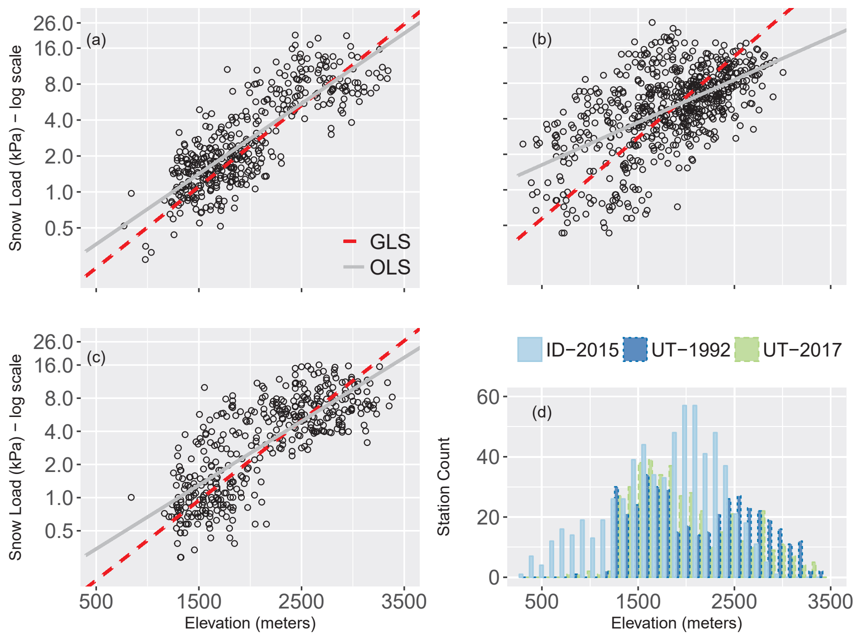

Fig. 2. Station elevation plotted against design ground snow loads (log scale) for (a) UT-2017, (b) ID-2015, and (c) UT-1992. Lines for ordinary (OLS) and generalized least squares (GLS) are given in each case. In addition, (d) shows histograms of station elevations for each dataset.

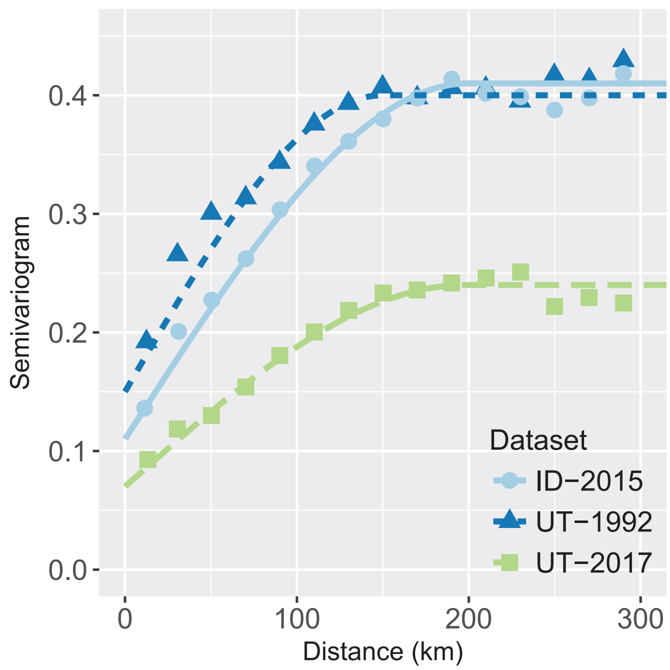

Fig. 3. Empirical (points) and theoretical (lines) semivariograms for each of the three datasets.

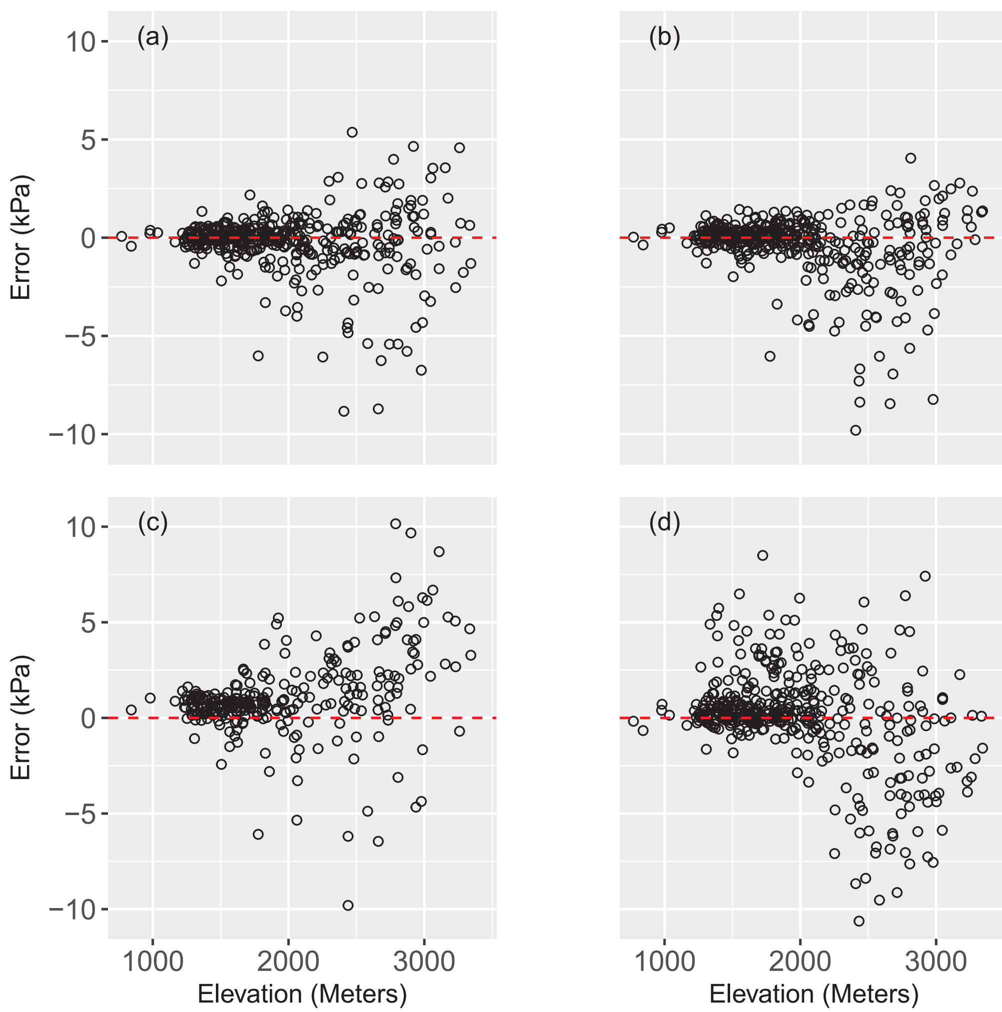

Fig. 4. Scatter plot of cross validated errors for (a) PRISM, (b) SKLM, (c) SNLW, and (d) IDW on UT-2017.

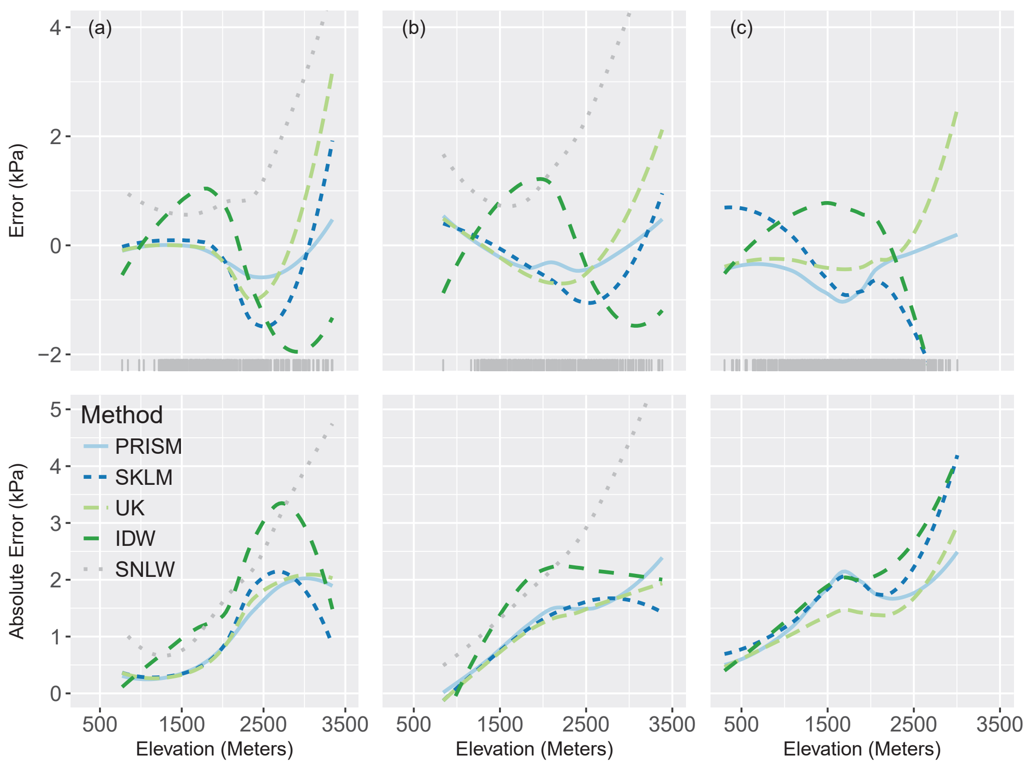

Fig. 5. Smoothed errors and absolute errors for (a) UT-2017, (b) UT-1992, and (c) ID-2015. The gray tick marks plotted along the x-axis of the three upper figures denote the individual station elevations.

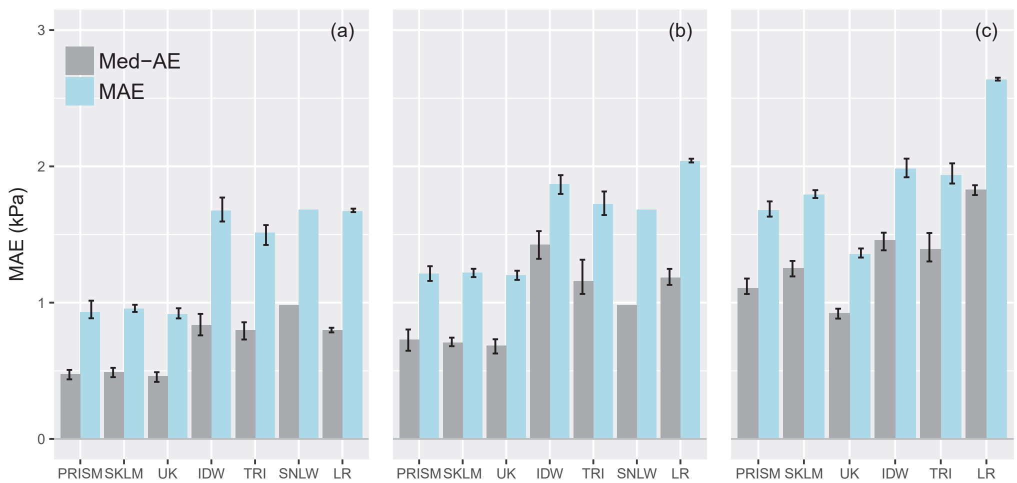

Fig. 6. Barchart of mean absolute errors (MAE) and median absolute errors (Med-AE) of spatial prediction methods for (a) UT-2017, (b) UT-1992, and (c) ID-2015.

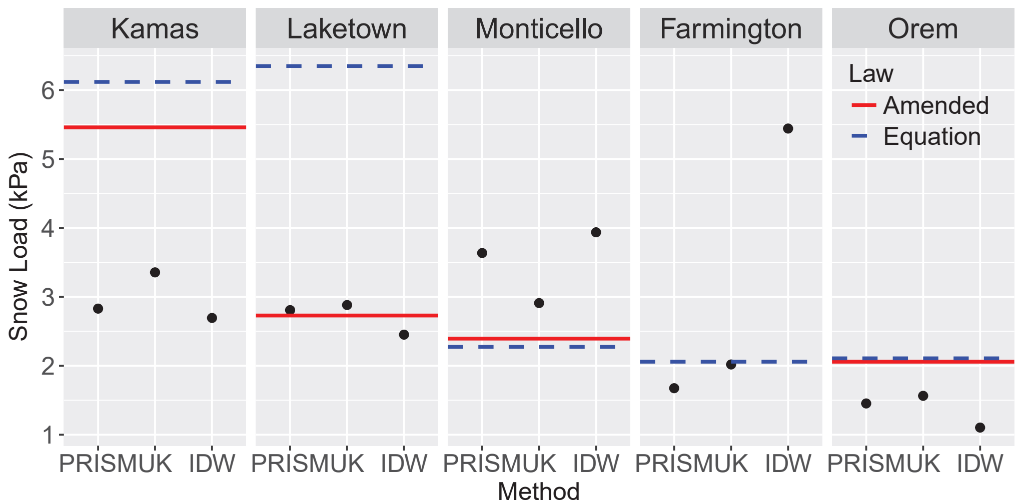

Fig. 7. Comparisons of spatial prediction methods to the 1992 Equations and recent 2016 amendments at select cities in Utah.

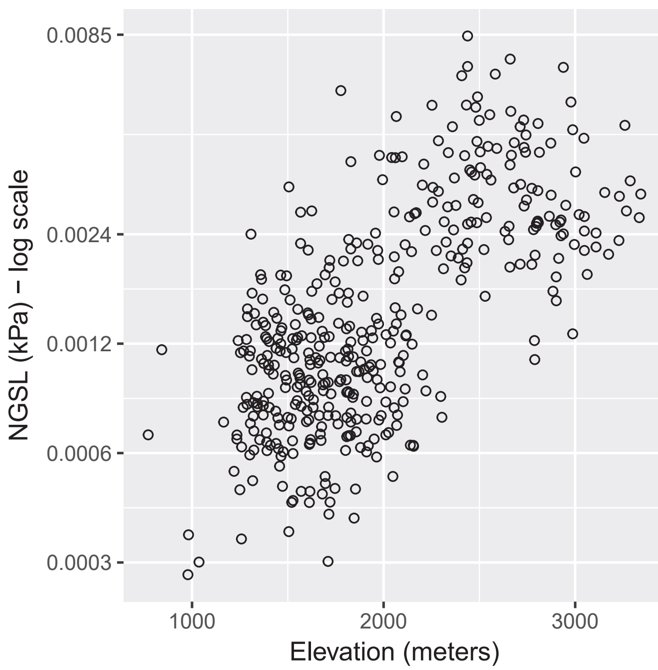

Fig. 8. Station elevation plotted against NGSL (log scale), showing that there is still a clear (and unaccounted for) log-linear relationship between NGSL and elevation.

List of Tables

TABLE 1. Summary of the three design ground snow load datasets used in method comparisons.

Dataset	Stations	SWE Conversions	Distribution
UT-2017	415	Sturm's Equation	Log-Normal
UT-1992	413	RMCD	Log-Pearson Type III
ID-2015	651	RMCD	Log-Pearson Type III

TABLE 2. NGSL at four nearest locations to Farmington, UT (111.884 W, 40.981 N).

Station	Elevation (m)	Distance to Location (km)	NGSL (kPa/m)
USC00422726	1335	5.4	0.0013
USS0011J11S	2438	5.5	0.0070
USS0011J12S	2066	6.4	0.0050
USS0011J68S	2359	8.4	0.0047

TABLE 3. 50 year ground snow load estimates for Weiser, Idaho using a variety of distributions.

	50 year estimate (kPa)	
Method	**Sturm**	**RMCD**
Log-Normal	1.64	1.04
Normal	1.54	1.07
Gumbel	1.55	0.99
GEV	2.34	1.25
Idaho Report		0.81

TABLE 4. Median absolute relative difference in 50 year estimates as compared to original log-normal distribution estimates.

	Absolute Relative difference (%)	
Method	**Sturm**	**RMCD**
log-Normal		35%
Normal	13%	42%
Gumbel	8%	40%
GEV	21%	29%