1 Kriging

1.1 Background

Physical surfaces vary continuously and must therefore be spatially correlated at short distance but statistically independent at large distances (Davis and Sampson, 1986). If sampling points are distributed throughout a surface, the degree of spatial correlation of the observed surface can be determined and the surface can then be interpolated between sampling points. Kriging is a geostatistical interpolation method that finds a set of optimal weights for the values at each sampling location based on the spatial correlation of measured values and then estimates values at unsampled locations (Davis and Sampson, 1986; Li and Heap, 2014). The kriging estimate ($\hat{z}$) at point x_0 is found using the equation

$$\hat{z}(x_0) - \mu = \sum_{i=1}^{n} \lambda_i [z(x_i) - \mu(x_0)], \tag{1}$$

where μ is a known stationary mean, λ_i is a kriging weight, $z(x_i)$ is the measured value of the surface at point x_i, n is the number of sampled points used to make the estimation (depends on size of sampling window), and $\mu(x_0)$ is the mean value of sampled points in the search window (Wackernagel, 2013; Li and Heap, 2008). Slight variations on Equation 1 result in various forms of kriging that are suited for different data sets. The kriging weights are estimated by minimizing the variance or the squared error, which is given by

$$\begin{aligned} \mathrm{var}[\hat{z}(x_0)] &= E[(\hat{z}(x_0) - z(x_0)^2)] & (2) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j C(x_i - x_j) + C(x_0 - x_0) - 2 \sum_{i=1}^{n} \lambda_i C(x_i - x_0), & (3) \end{aligned}$$

where $C(x_i - x_j) = \mathrm{Cov}[z(x_i), z(x_j)]$ is the covariance of the surface (Li and Heap, 2008). Kriging assumes that the variance does not depend on the sample location, but depends only on distance between samples. Anisotropic kriging allows for variance to change with direction but the variance still does not depend on sample location. Isotropic variance was assumed in this study.

A physical surface is usually assumed to be noisy and this noise is captured by a "nugget" parameter that allows for the surface to vary smoothly

and thus, not pass directly through each measured point. The nugget is a residual that encompasses sampling-error variance as well as the spatial variance at distances less than the minimum sample spacing (Li and Heap, 2008). Spatial correlation of data is estimated using imperical variance and is often plotted in the form of a variogram (see Section **??**).

There are many forms of kriging and each one is tailored to suit different data types (see Li and Heap (2014)). The two most appropriate types of kriging for this study are simple kriging and regression kriging.

1.1.1 Simple Kriging

Simple kriging is used to estimate residuals about a constant and stationary mean μ, which is typically calculated as the average of the data. Kriged estimates for simple kriging are found by slightly modifying Equation 1 to

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) + \left[1 - \sum_{i=1}^{n} \lambda_i\right] \mu. \tag{4}$$

Larger values of $[1 - \sum_{i=1}^{n} \lambda_i]$ result in estimates that are closer to the data mean (Li and Heap, 2008). The value of $[1 - \sum_{i=1}^{n} \lambda_i]$ generally increases in poorly sampled areas.

1.1.2 Regression Kriging

Regression kriging estimates values between measurement locations by combining a regression estimate (see Section **??**) with a kriged estimate of the regression residuals (Hengl and others, 2007). First, the regression estimate is determined using auxiliary variables (e.g. terrain parameters such as elevation and slope). Then, simple kriging is used to interpolate regression residuals, which have and assumed mean of zero. The two surface estimates are then added. The final estimate can be written as

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) \tag{5}$$

$$= \sum_{k=0}^{p} \hat{\beta}_k \cdot q_k(x_0) + \sum_{i=1}^{n} \lambda_i \cdot e(x_i), \tag{6}$$

where $\hat{m}(x_0)$ is the regression estimate and $\hat{e}(x_0)$ is the interpolated residual, $\hat{\beta}_k$ are the estimated regression coefficients, $q_k(x_0)$ are the regressors, p is the

number of regressors, λ_i are the kriging weights for the residuals, and $e(x_i)$ is the residual at x_i.

Regression kriging can be thought of as an intermediate between pure kriging (no regression) and pure regression (small residuals) and can be more strongly skewed to either end-member based on the strength of the regression correlation (Hengl and others, 2007). Regression kriging is mathematically equivalent to universal kriging, in which auxiliary variables are used directly to determine the kriging weights (Hengl and others, 2007). However, separating the trend analysis and kriging steps has the advantage of being able to test regression methods that go beyond a basic linear trend. Kriging combined with regression has been found to produce better estimates of spatial fields when compared to simple kriging and co-kriging (Knotters and others, 1995).

1.2 Methods

Simple kriging is implemented using the R package `DiceKriging`, which approximates functions by first generating a kriging model from input data and then estimating data values at new locations based on their distance from observed data (Roustant and others, 2012). For the SWE data, a nugget value needs to be estimated by the `DiceKriging` package because the covariance length varies considerably over short distances. For this study, the kriged surface cannot not be estimated without a nugget, indicating that the function is not purely deterministic. Maximum likelihood is used to estimate both nugget and covariance values in `DiceKriging`. Negative values of kriged SWE were set to zero. For more details on implementation of `DiceKriging` see Appendix 2.

1.3 Results

1.3.1 Simple Kriging

There are large different in spatial pattern of interpolated SWE distributions for the three study glaciers found using simple kriging (Figure 1). The lower half of Glacier 4 has a relatively uniform SWE, while the accumulation area has both low and high values of SWE. The low density of sampling points in the accumulation area of Glacier 4 result in large gradients in SWE. Glacier 2 has two distinct and relatively uniform areas — the lower ablation area

has low SWE ($\sim$ 0.1 m w.e.) and the upper ablation and accumulation areas have higher SWE values ($\sim$ 0.6 m w.e.). The boundary between two these zones closely follows the outline of the ice dune area observed during field data collection. Glacier 13 does not appear to have any strong patterns and accumulation is generally low ($\sim$ 0.1–0.5 m w.e.).

The glacier-wide winter mass balance on each glacier (B_w), calculated as the mean SWE, that is found using simple kriging is considerably different to that found using a topographic regression (Figure **??**). Accumulation on Glacier 4 is 0.09 m w.e. (15%) higher and has larger spatial gradients in SWE when simple kriging is used. Conversely, simple kriging estimates are 0.13 m w.e. (23%) and 0.08 m w.e. (22%) lower for Glaciers 2 and 13, respectively, and the range of SWE values is smaller. The accumulation distribution estimated by simple kriging and by regression is qualitatively similar for Glacier 13 but differs considerably for Glacier 4 and 2. Glacier 4 has large SWE values in the accumulation area when simple kriging is used, which contrasts strongly with the mostly uniform accumulation estimated with regression. The converse is observed on Glacier 2, where the simple kriging estimate has a small SWE range and two distinct regions of SWE values while the regression estimate has a large range and a significant elevation gradient in accumulation.

Estimates of the nugget found using maximum likelihood in the DiceKriging package also vary between glaciers (Table 1). Nugget values are small for both Glacier 2 and 13 ($\sim$ 0.004 m w.e.) and do not vary when different methods for interpolating snow density (Section **??**) are used. The nugget values for Glacier 4 are an order of magnitude larger and vary by $\sim$80% of the minimum nugget value for this glacier.

Comparing estimated and observed values of SWE (Figure 2) shows that kriging is best able to predict SWE values on Glacier 2 and that kriging has the lowest predictive power on Glacier 4. This is similar to topographic regression estimates (Figure **??**), where the least variance was explained on Glacier 4 and the most variance was explained on Glacier 2. When compared to regression estimates, the R^2 values between SWE estimated using simple kriging and observed SWE are higher for all glaciers. Correlation values increase from 0.11, 0.63 and 0.36 to 0.25, 0.84 and 0.49 for Glaciers 4, 2 and 13, respectively. The largest difference in R^2 between these two methods is on Glacier 2, which also has the highest R^2 values. Note that the scatter of estimated and observed SWE values in Figure 2 is due to the nugget, which encompasses sampling-error variance and variance at distances smaller than

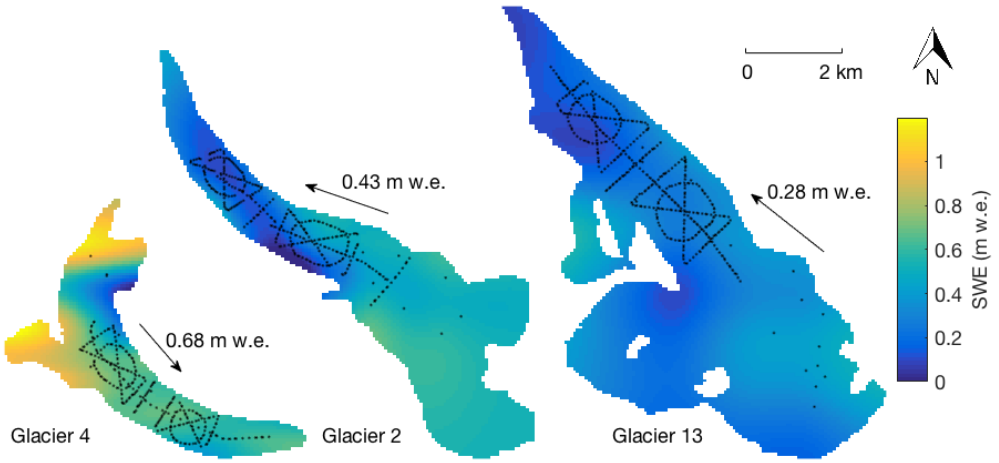

Figure 1: SWE distributions estimated by kriging. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

Table 1: Nugget (m w.e.) values for SWE data with various snow density interpolation schemes estimated using maximum likelihood in DiceKriging package. S = Snowpit density values, F = Federal Sampler density values. See Section **??** for details on density options.

Density Option	**Glacier 4**	**Glacier 2**	**Glacier 13**
S1	0.015	0.004	0.004
F1	0.013	0.003	0.003
S2	0.009	0.004	0.004
F2	0.009	0.003	0.003
S3	0.016	0.004	0.005
F3	0.017	0.002	0.003
S4	0.014	0.004	0.004
F4	0.009	0.002	0.003

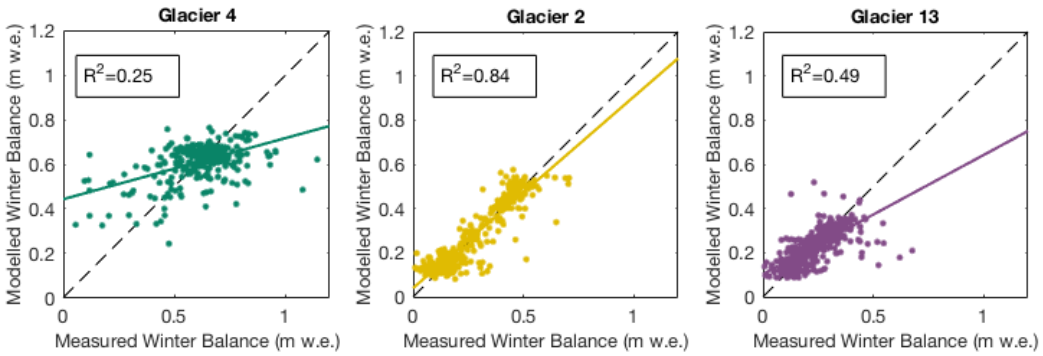

Figure 2: Comparison of estimated (simple kriging) and observed (original) snow water equivalent (SWE) for three study glaciers. The SWE values were calculated using inverse-distance weighted snowpit densities (S4).

the sample spacing.

1.3.2 Residual Kriging

Estimating SWE values with a regression oftopographic parameters is described in Section **??**. For this section, the BMA regression is used because it resulted in a better fit than MLR between estimated and observed SWE.

The range, magnitude and spatial pattern of estimated regression residuals found using simple kriging varies between the three study glaciers (Figure 3). Generally, the range of residual values is highest on Glacier 4 and lowest on Glacier 13. Extreme value are located in the accumulation area of Glacier 4 with both strongly negative and strongly positive residuals located within a kilometre of each other. The low density of sampling points in the accumulation area biases the interpolation of residuals to fit the over- and underestimation of SWE at the two uppermost sampling locations. Residuals show less variation on Glacier 2, although relatively large residuals of $\sim \pm 0.4$ m w.e. are present in the upper ablation area along the ice margins. Glacier 13 has the smallest range of residuals but residual values are approximately equal to those of estimated SWE. The mean value of distributed residuals is positive for Glacier 4, indicating that the topographic regression may have underestimated the winter balance. Conversely, the mean residual for Glacier 2 is negative, indicating that the regression may have over

estimated the winter balance for this glacier.

The SWE distribution found using simple kriging and regression kriging (Figure 4) shows a gradient in accumulation across the mountain range. Glacier 4 has the highest mean SWE and Glacier 13 has the lowest mean SWE. However, the accumulation gradient is steeper for simple kriged estimates than for regression kriged estimates. Glacier 4 has a similar mean SWE between the two methods but mean SWE on Glaciers 2 and 13 are much lower. All three glacier have the largest values of SWE in the upper part of the accumulation areas.

SWE distribution estimated using regression kriging is similar to the winter balance estimated using topographic regressions for Glaciers 2 and 13 (Figure 4). The similarity is due to the relatively high explanatory power of the regressions, so the regression kriging estimate is closer to a pure regression. The distribution of SWE on Glacier 4 is more strongly affected by the kriged residuals because the regression had low explanatory power, resulting in larger residual values that changed the snow distribution, particularly in the accumulation area. The regression kriging for Glacier 4 is therefore closer to pure kriging.

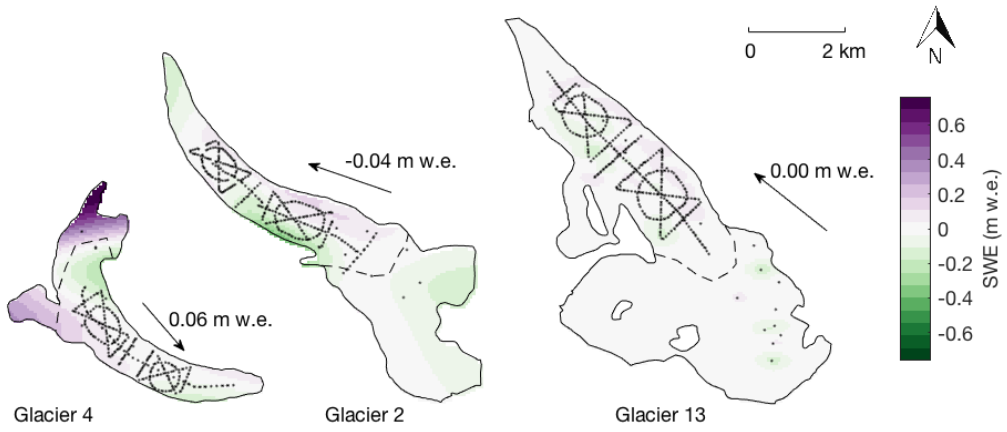

Figure 3: Distributed BMA residuals estimated by simple kriging. Arrows indicate glacier flow direction and black dots show snow depth sampling locations. Dashed line indicates approximate ELA.

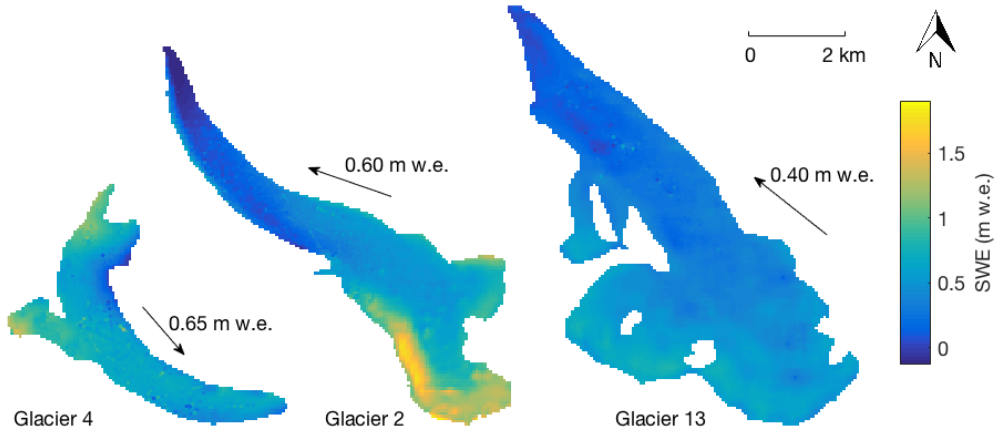

Figure 4: SWE distributions estimated by adding the kriged residuals to the SWE distributions estimated using topographic regression with BMA. Observed SWE values are overlain on the maps. Arrows indicate glacier flow direction.

Regression kriging produces the highest overall set of R^2 values (Figure 5). There is a large increase in correlation for both Glaciers 4 and 13 when compared to simple kriging and pure regression (Figure ???). The R^2 value for Glacier 2 is only slightly lower than that for simple kriging.

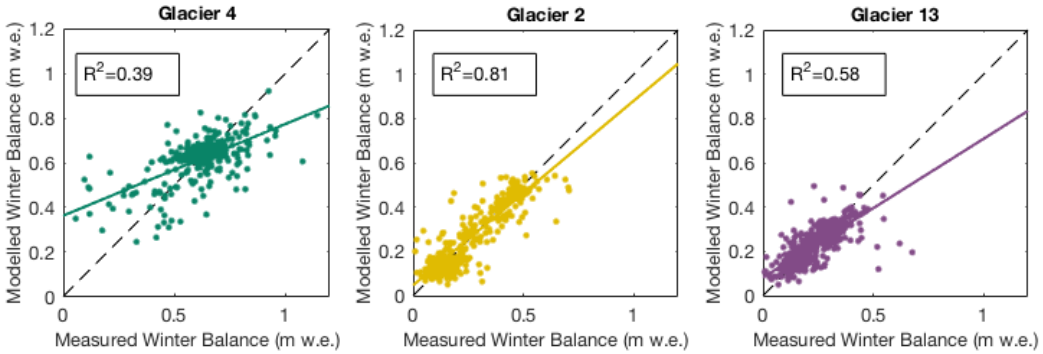

Figure 5: Comparison of estimated (regression kriging) and observed (original) snow water equivalent (SWE) for three study glaciers. The SWE values were calculated using inverse-distance weighted snowpit densities (S4).

References

Davis JC and Sampson RJ (1986) *Statistics and data analysis in geology*, volume 646. Wiley New York et al.

Hengl T, Heuvelink GB and Rossiter DG (2007) About regression-kriging: From equations to case studies. *Computers & Geosciences*, **33**(10), 1301–1315 (doi: 10.1016/j.cageo.2007.05.001)

Knotters M, Brus D and Voshaar JO (1995) A comparison of kriging, cokriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma*, **67**(3), 227–246 (doi: 10.1016/0016-7061(95)00011-C)

Li J and Heap AD (2008) A review of spatial interpolation methods for environmental scientists No. Record 2008/23. *Geoscience Australia*

Li J and Heap AD (2014) Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, **53**, 173–189 (doi: 10.1016/j.envsoft.2013.12.008)

Roustant O, Ginsbourger D and Deville Y (2012) DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, **21**, 1–55

Wackernagel H (2013) *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media

2 APPENDIX - Kriging software

Data kriging is executed using the DiceKriging package in R (Roustant and others, 2012). The Matlab function `KrigingR()` moves SWE data to R and then uses DiceKriging to find the estimated kriged surface of input data, as well as the upper and lower confidence intervals, the cross-validated (leave-one-out) estimates of SWE, as well as parameters that describe the kriging model fit (nugget, maximum log likelihood and mean constant). Input parameters of `KrigingR()` are the snow data (e.g. SWE values), locations of measurements, and the glacier that the data is located on. The function follows these steps:

1. The working directory is changed to the 'Kriging' folder
2. The DiceKriging.R script is run
   (a) The function `km()` estimates the best fit kriging model with a constant mean using the Matérn $\nu = 3/2$ covariance kernel. The maximum log likelihood, model intercept (mean) and nugget estimates from the model are then extracted.
   (b) Leave-one-out cross validation is then completed using the function `leaveOneOut.km()`.
   (c) Surface prediction is then completed for the entire glacier. First, a grid that is the same size as the glacier is created (point spacing every 40 m to match the size of the DEM). Then, the function `predict()` is used to predict the surface values at all grid locations. The predicted values, as well as the upper and lower 95% confidence intervals are returned.
   (d) All returned values are then saved as a .mat file.
3. The .mat file is imported into Matlab and the data is organized into a structure. Cells outside of the glacier outlines are set to `NaN` and negative SWE values are set to zero. The final data structure is then outputted from the Matlab function.