Estimating spatial distribution of daily snow depth with kriging methods: combination of MODIS snow cover area data and ground-based observations

C. L. Huang\textsuperscript{1,2}, H. W. Wang\textsuperscript{1,2}, and J. L. Hou\textsuperscript{1,2}

\textsuperscript{1}Key Laboratory of Remote Sensing of Gansu Province, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, China
\textsuperscript{2}Heihe Remote Sensing Experimental Research Station, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu, 730000, China

Received: 10 August 2015 – Accepted: 1 September 2015 – Published: 22 September 2015

Correspondence to: C. L. Huang (huangcl@lzb.ac.cn)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Accurately measuring the spatial distribution of the snow depth is difficult because stations are sparse, particularly in western China. In this study, we develop a novel scheme that produces a reasonable spatial distribution of the daily snow depth using kriging interpolation methods. These methods combine the effects of elevation with information from Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover area (SCA) products. The scheme uses snow-free pixels in MODIS SCA images with clouds removed to identify virtual stations, or areas with zero snow depth, to compensate for the scarcity and uneven distribution of stations. Four types of kriging methods are tested: ordinary kriging (OK), universal kriging (UK), ordinary co-kriging (OCK), and universal co-kriging (UCK). These methods are applied to daily snow depth observations at 50 meteorological stations in northern Xinjiang Province, China. The results show that the spatial distribution of snow depth can be accurately reconstructed using these kriging methods. The added virtual stations improve the distribution of the snow depth and reduce the smoothing effects of the kriging process. The best performance is achieved by the OK method in cases with shallow snow cover and by the UCK method when snow cover is widespread.

1 Introduction

Snow is an important type of land cover that affects the energy and water balances of drainage basins in alpine regions, which have high albedos and sensitivity to weather and climate changes (Jin et al., 2006). The contribution of snowmelt to runoff is an important aspect of water resources in mountainous regions, in addition to rainfall and glacial melt. Snow stores at least one-third of the water used for irrigation and crop growth worldwide. For example, snowmelt contributes up to 90% of the annual runoff in the high-elevation basins of the Rocky Mountains (Schmugge et al., 2002).
Because of the extreme spatial variability in the snow depth in mountainous terrain, estimating the spatial distribution of snow depth is difficult. Snow depth is also influenced by numerous processes, such as snow drift (Savelyev et al., 2006), avalanches (Elder et al., 1991; Birkeland et al., 1995), and snow accumulation and ablation (Pomeroy et al., 1998; Hardy et al., 1998; DeBeer and Pomeroy, 2009). Therefore, point measurements of snow depth from snow courses or meteorological stations may not adequately represent the average conditions of a study area, particularly in alpine areas. Because collecting ground measurements of snow in harsh, remote environments of snow-dominated areas is logistically difficult and expensive, ground observations of snow are usually very sparse or unavailable. Therefore, the only feasible alternative for collecting comprehensive snow information is through more robust approaches (Marofi et al., 2011). Recently, interpolation methods, including geostatistics, have been used for snow distribution modeling, including regression models (Molotch et al., 2005), statistical relationships (Elder et al., 1991), binary regression tree methods (Elder et al., 1998), and kriging (Balk and Elder, 2000; Erxleben et al., 2002). Rice et al. (2011) also used visible satellite snow cover estimates to constrain the interpolation of snow depth/SWE. However, the performance of these interpolation methods is dependent on the sampling scheme, sampling distribution, and quantity of the sample. Therefore, a suitable interpolation method for distributing point measurements of the snow depth over an area must be selected.

In western China, which has mountainous terrain and only limited daily snow depth observations at meteorological stations, predicting the spatial distribution of snow depth is a considerable challenge.

In this study, we propose a new scheme for estimating the spatial distribution of the snow depth. This scheme combines elevation data with Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover area (SCA) products. The new scheme is evaluated with daily snow depth observations at 50 meteorological stations in northern Xinjiang Province, China. Four types of kriging methods are evaluated: ordinary kriging (OK), universal kriging (UK), ordinary co-kriging (OCK), and universal co-kriging (UCK).
This paper is organized as follows. Section 2 describes the methodology and data used in this research. Section 3 presents the results of the snow depth interpolation using the different methods. Finally, a brief conclusion is provided in Sect. 4.

2 Study area and data

Northern Xinjiang Province in western China is selected as the study area (Fig. 1). This area comprises temperate continental arid and semi-arid climates. Mountains, plains and deserts are the three major geomorphologic units. The area is influenced by the Siberian circulation, so snowfall is very frequent during winter and spring. In fact, the area is one of the three snowiest locations in China. The average snow depth is 60 cm, with a maximum of 1–2 m in the mountains (Liang et al., 2008). Therefore, this area is very suitable for validating our snow depth interpolation scheme. Snow depths were recorded at fifty meteorological stations within the study area (Fig. 1). Of those stations, the snow depth measurements from forty-six stations are used to produce the distribution of the snow depth, while the remaining measurements from the Fukang, Beitashan, Kelamayi, and Tekesi stations are used to evaluate the proposed interpolation methods quantitatively.

2.1 AMSR-E derived snow depth

The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) brightness temperature data were obtained from the United States National Snow and Ice Data Center (NSIDC) (http://nsidc.org/). The modified snow depth retrieval algorithm was developed by Che et al. (2008) were used in this study to produce the daily snow depth product with 25 km spatial resolution, which can be downloaded from Cold and Arid Regions Science Data Center (http://card.westgis.ac.cn/). In this study, the coarse-resolution snow depth product were used as auxiliary data to produce cloud-free MODIS snow cover area data.
2.2 MODIS SCA data

MODIS snow cover data from the NSIDC (http://nsidc.org/) were used in this study. The MODIS data are collected by the morning satellite Terra and the afternoon satellite Aqua at a 500 m resolution; these data are used to produce the daily snow products MOD10A1 and MYD10A1. First, we use the MRT tool to splice and reproject these data. Second, we combine Terra MODIS and Aqua MODIS snow cover products (for a daily synthesis) and apply adjacent-pixel analysis to remove clouds based on the local similarity of the spatial distributions according to the snow line method (SNOWWL). The AMSR-E derived snow depth data and MODIS image fusion are used to remove clouds in northern Xinjiang snow products and to identify areas without snow (Huang et al., 2012).

2.3 DEM data

We use raw data from the 90 m resolution Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM). The dataset is provided by the International Scientific and Technical Data Mirror Site, Computer Network Information Center, Chinese Academy of Sciences (http://www.csdb.cn/). The raw data are also mosaicked and reprojected to the Universal Transverse Mercator (UTM) projection with a WGS84 datum at a 500 m resolution.

3 Method

3.1 Kriging methods

Kriging is an interpolation technique that uses observations \( z(x_i) \) at location \( x_i \) to estimate the values \( z(x_0) \) at a point \( x_0 \) where observations are not available. The random variable \( z \) at any location can be written as the sum of a deterministic component called the trend, \( m(x) \), and a stochastic error component, \( r(x) \). In this study,
\( z(x_i) \) is the measured snow depth at each meteorological station, and four types of kriging methods are adopted: ordinary kriging, universal kriging, ordinary co-kriging, and universal co-kriging. Detailed descriptions of these kriging methods are given in textbooks (Goovaerts, 1997), and all these methods can be implemented through the gstat R package (Pebesma and Wesseling, 1998; Pebesma, 2004). For completeness, we provide brief descriptions of these kriging methods in the following sections.

### 3.1.1 Ordinary kriging (OK)

Ordinary kriging is a spatial interpolation estimator used to find the best linear unbiased estimate at non-sampled locations with an unknown constant mean as follows:

\[
\hat{z}_{\text{ok}}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i)
\]

(1)

\[
\sum_{i=1}^{n} \lambda_i = 1
\]

(2)

Here, \( \hat{z}_{\text{ok}}(x_0) \) is the estimated snow depth at the non-sampled location \( x_0 \), \( z(x_i) \) is the measured snow depth at the meteorological station \( x_i \), and \( \lambda_i \) is the weighting factor for \( z(x_i) \).

### 3.1.2 Universal kriging (UK)

Universal kriging (UK) is also called kriging with an external trend, which is a procedure that uses a regression model as part of the kriging process, typically modeling the unknown local mean values with a local linear or quadratic trend. UK usually assumes the model

\[
z(x) = m(x) + r(x) = \sum_{p=1}^{l} a_p f_p(x) + r(x)
\]

(3)
where $z(x)$ is the variable of interest at location $x$, $m(x)$ is a deterministic function that is sometimes called drift, and $r(x)$ is a random variable. The coefficient $a_p$ is the $p$th coefficient, $f_p$ is the $p$th basis function that describes the trend, and $l$ is the number of functions used to model the trend. In this study, $f_p$ is the elevation basis function, and $l$ is equal to 1. Thus, the estimation model of snow depth can be expressed as

$$z(x) = ah(x) + b + r(x)$$

(4)

where $h(x)$ is the elevation at location $x$ and $a$ and $b$ are coefficients that are fitted to observations each day. The best estimate of the snow depth at location $x_0$ is created by combining the estimated drift $\hat{m}(x_0)$ and the estimated residual $\hat{r}(x_0)$:

$$\hat{z}_{uk}(x_0) = \hat{m}(x_0) + \hat{r}(x_0)$$

(5)

3.1.3 Ordinary co-kriging (OCK)

OCK is often used to address multiple correlated variables: it incorporates both the main variable of interest and all other variables types into its predictions. In this study, the elevation of each meteorological station is considered as a co-variable, so the estimated snow depth $\hat{z}_{ock}(x_0)$ at location $x_0$ is obtained as a linear combination of the measured snow depth $z(x_i)$ and elevation $h(x_i)$ given by the OCK estimator

$$\hat{z}_{ock}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) + \sum_{i=1}^{n} \beta_i h(x_i)$$

(6)

where $\lambda_i$ and $\beta_i$ are the coefficients of the measured snow depth and elevation at each meteorological station, respectively; these values are obtained as Lagrangian multipliers for a constrained optimization problem. The unbiasedness condition results in the following conditions for the weighting coefficients:

$$\sum_{i=1}^{n} \lambda_i = 1, \text{ and } \sum_{i=1}^{n} \beta_i = 0$$

(7)
3.1.4 Universal co-kriging (UCK)

UCK is also called co-kriging with external drift. The method is a combination of OCK and UK; it improves predictions by considering the drift of the main variable and by introducing one or more co-variables. The UCK estimator of snow depth is also expressed as Eq. (6), with the addition that the main variable, snow depth \( z(x_i) \), has external drift that is related to the elevation. Additionally, the elevation \( h(x_i) \) is considered a co-variable.

3.2 Performance evaluation

To assess which spatial prediction method provides the most accurate estimates of snow depth for each study site, cross-validation is used to compare the estimated values with their true values. Cross-validation is accomplished by removing each data point and then using the remaining measurements to estimate the data value. This procedure is repeated for all observations in the dataset. The true values are subtracted from the estimated values. The residuals resulting from this procedure are then evaluated to assess the performance of the methods. The root mean square error (RMSE) is adopted as an evaluation indicator:

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2}
\]  

(8)

where \( n \) is the number of meteorological stations and \( x_i \) and \( y_i \) are the measured and estimated snow depth, respectively, at the \( i \)th station.
To evaluate the accuracy of the interpolated snow depth, the BIAS quantity and the correlation coefficient ($R$) are also used in this study.

$$BIAS = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)$$

$$R = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

3.3 Spatial interpolation scheme of snow depth

The flowchart of the scheme is shown in Fig. 2 and is briefly described as follows.

(1) Generate the spatial distribution of cloud-free snow cover. The cloud-removal method proposed by Huang et al. (2012) is used in this study; it removes most cloud-contaminated pixels by fusing TERRA and AQUA MODIS snow-covered area products (MOD10A1 and MYD10A1) with snow depth products derived from AMSR-E.

(2) Set up virtual stations. We randomly choose snow-free pixels from the cloud-removed MODIS SCA image as virtual stations with snow depths equal to zero. These virtual stations not only increase the number of stations but also reduce the “smoothing effects” caused by kriging. This feature improves the estimated snow depth, particularly in areas with shallow snow.

(3) Interpolate snow depth observations using the kriging methods. We first calculate the correlation coefficients between the snow depth and elevation based on the 10-day history of the snow depth observations at available meteorological stations. Then, four types of kriging methods are used to calculate the spatial distribution of the snow depth. The kriging method with the minimum RMSE, as determined by the cross-validation, is chosen to produce a high-quality spatial distribution of snow depth for each day. The
results of the interpolated snow depth with and without virtual stations are hereafter referred to as scheme-1 and scheme-2, respectively.

4 Results

4.1 Cloud-removed MODIS SCA

Figure 3 shows the evolution of the cloud-covered and snow-covered areas in northern Xinjiang from the MOD10A1 and MYD10A1 products from November 2004 to March 2005. Both the MOD10A1 and MYD10A1 products are intermittently contaminated by clouds (Fig. 3). We cannot determine whether cloud-covered pixels contain snow, which leads us to underestimate snowfall in areas of heavy snow. Fig. 3 also shows that snow-covered areas can be recovered by integrating MOD10A1-, MYD10A1- and AMSR-E-derived snow depth data. The cloud-removal scheme is very efficient and sufficiently accurate in mountainous regions (Huang et al., 2012).

Figure 4 shows snapshots of the snow cover area and its spatial distribution from MOD10A1, MYD10A1, and the cloud-removed images on 1 December 2004, and 1 February 2005. Clearly, MOD10A1 and MYD10A1 contain large quantities of cloud cover that obscures the ground conditions. Fortunately, microwave radiation can penetrate clouds and determine the actual land surface conditions, so fusing MOD10A1-, MYD10A1-, and AMSR-E-derived snow depths can provide a reasonable estimate of the snow cover.

4.2 Spatial distribution of the interpolated snow depth

Figure 5 shows the evolution of the SCA and the correlation coefficient (R) between the snow depth and elevation in each 10 day period from 1 November 2004 to 31 March 2005. The varying trends between R and SCA during periods of the snow season are apparent. R and SCA have similar increases at the beginning of the snow season. R remains relatively stable (approximately 0.4) when the SCA exceeds 40%.
during the middle of the snow season. During the ablation period, \( R \) decreases sharply as SCA decreases.

Based on the observed snow depth at 46 meteorological stations, the MODIS-derived SCA, and the elevation, kriging methods are used to produce snow depth maps for each day. The cross-validation of the interpolated snow depth with the kriging methods on the first day of each month is shown in Table 1. Comparing the adopted kriging methods (OK, UK, OCK, and UCK), the RMSE of OK is the smallest for 1 November and 1 December 2004. Less snow exists in the study area, and the \( R \) between the snow depth and SCA is very small on these days. With increasing SCA and increasing \( R \) between the snow depth and SCA, the RMSE increases significantly; thus, UCK achieves the best performance among the adopted kriging methods on 1 January, 1 February, and 1 March 2005. Thus, elevation is an important co-variable when most regions are covered by snow. Furthermore, the results using scheme-2 are more accurate than the results using scheme-1 in all cases. Adding virtual stations increases the number of samples and can thus somewhat reduce the interpolation errors.

Figure 6 shows the spatial distribution of the estimated snow depth based on the best kriging method with the two schemes on 1 December 2004, 1 January 2005, 1 February 2005, and 1 March 2005. All maps predict deep snow in the northeast and southwest portions of the study area because of the mountainous terrain. Outside of these areas, the maps estimated by the two schemes differ significantly. The maps of snow depth estimated by scheme-2 are more detailed than those created by scheme-1. Additionally, the snow depths estimated with scheme-1 are much higher than those estimated with scheme-2, particularly in areas of deep snow. The scheme-2 spatial distribution of snow depth better matched the DEM and provided more detail. More complicated and heterogeneous features of the snow depth distribution can also be seen in the results of scheme-2 because this scheme considers a greater number of virtual stations, thus increasing the number of samples.
4.3 Validation at independent meteorological observation sites

To assess the performance of the proposed scheme for snow depth interpolation, the observed and estimated snow depths at four independent meteorological stations are used. These data are shown in Fig. 7. The BIAS, RMSE, and $R$ of the estimated snow depth at these four meteorological stations are listed in Table 2. The estimated snow depth is computed by the optimal kriging method at each time step. Comparing these values to the measured snow depth at each station, the estimated snow depths with scheme-1 and scheme-2 capture the trend of the snow depth variation over the entire snow season (the $R$ is greater than 0.89 for all stations), but the results produced with scheme-2 are more similar to the observations than the results produced with scheme-1. The statistical indicators in Table 2 also suggest that the results of scheme-2 are more accurate than the results of scheme-1. The BIAS ranges from 2.84 to 5.80 cm for scheme-1 but only from 0.90 to 3.8 cm for scheme-2. The RMSE is reduced from 1.98 to 1.13 cm at the Fukang station, from 3.46 to 2.31 cm at the Beitashe station, from 2.65 to 1.37 cm at the Kelamayi station, and from 1.67 to 0.89 cm at the Tekesi station. These results indicate that adding virtual stations without snow is useful for increasing the accuracy of the snow depth interpolation.

5 Conclusion and discussion

Blending visible satellite snow cover estimates to constrain the interpolation of snow depth/SWE is an useful way to improve accuracy of snow depth interpolation (Rice et al., 2011). In this study, an optional interpolation scheme that produces a reasonable spatial distribution of daily snow depths using four types of kriging interpolation methods (OK, UK, OCK, and UCK) is proposed. The scheme combines the effects of elevation with MODIS SCA products. The daily snow depth observations at 50 meteorological stations in northern Xinjiang Province, China, are used to validate the
propose scheme. The results suggest that the spatial distribution of the snow depth can be accurately reconstructed using the proposed method.

It is found that the elevation as co-variable is not always taken effect in the whole snow season. Early in the snow season, when relatively few regions are covered by snow, the correlation coefficient between the elevation and SCA is very low, and the OK method produces the best snow depth estimation without using the elevation. Later in the snow season, when a large number of regions are covered by snow, the correlation coefficient between the elevation and SCA is relatively high; thus, the UCK method produces the best snow depth estimation. Therefore, the candidate kriging method should be dynamically determined each day according to the SCA conditions and the correlation coefficient between the elevation and SCA.

Additionally, the accuracy of the interpolated snow depth is mainly determined by the number of the available stations. In this study area, only 50 snow depth measurements in each day can be used to generate the spatial distribution of the snow depth by the kriging method. Therefore, smoothing effects are always apparent in the results of the estimated snow depths, particularly in mountainous regions where the snow depths are consistently overestimated. Adding a few virtual stations with zero snow depths, as indicated by the MODIS SCA data, can increase the number of observations and reduce the smoothing effect to yield more complicated and heterogeneous features in the snow depth distribution. The snow depth derived from passive microwave sensors such as AMSR-E and SSM/I are also an important data source thought their spatial resolutions are very low. How to fuse all kinds of snow information (snow depth measurements from meteorological stations and derived from passive microwave sensors, and MODIS SCA products) to produce snow depth distribution with high-spatial resolution will be studied in next step.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (grant numbers 91325106 and 41271358) and the Project of the One Hundred Person Project of the Chinese Academy of Sciences (grant number 29Y127D01).
References


Table 1. Cross-validation of the snow depth spatial interpolation using various kriging methods. Here, $R$ is the correlation coefficient between the snow depth and elevation, scheme-1 represents the results at 46 meteorological stations, and scheme-2 represents the results at 46 meteorological stations and virtual stations.

<table>
<thead>
<tr>
<th>Date</th>
<th>SCA (%)</th>
<th>$R$</th>
<th>RMSE (cm) (scheme-1)</th>
<th>RMSE (cm) (scheme-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>OK</td>
<td>UK</td>
</tr>
<tr>
<td>1 Nov 2004</td>
<td>6.701</td>
<td>0.001</td>
<td>0.292</td>
<td>0.297</td>
</tr>
<tr>
<td>1 Dec 2004</td>
<td>13.714</td>
<td>0.060</td>
<td>2.778</td>
<td>2.785</td>
</tr>
<tr>
<td>1 Feb 2005</td>
<td>69.710</td>
<td>0.375</td>
<td>12.68</td>
<td>11.69</td>
</tr>
</tbody>
</table>
Table 2. Bias and RMSE of the snow depth at four meteorological stations from November 2004 to March 2005. Here, scheme-1 represents the results at 46 meteorological stations, and scheme-2 represents the results at 46 meteorological stations and virtual stations.

<table>
<thead>
<tr>
<th>Scheme-1</th>
<th>Scheme-2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
</tr>
<tr>
<td></td>
<td>3.329</td>
</tr>
<tr>
<td>Fukang</td>
<td>5.801</td>
</tr>
<tr>
<td>Beitashan</td>
<td>4.396</td>
</tr>
<tr>
<td>Kelamayi</td>
<td>2.836</td>
</tr>
</tbody>
</table>
Figure 1. Distribution of meteorological stations in northern Xinjiang Province, China.
Figure 2. Spatial interpolation scheme for snow depth with kriging methods using ground-based snow depth observations, MODIS snow cover areas, and elevation.
Figure 3. Time series of snow-covered pixels among MOD10A1, MYD10A1, and cloud-removed MODIS snow cover area products from November 2004 to March 2005.
Figure 4. Maps of snow cover areas on 1 December 2004 (left column) and 1 February 2005 (right column), derived from (a1 and b1) MOD10A1, (a2 and b2) MD10A1, and (a3 and b3) the cloud-removed images.
Figure 5. Evolution of the snow cover area and the correlation coefficient between the snow depth and elevation in northern Xinjiang from 1 November 2004 to 1 March 2005.
Figure 6. Spatial distribution of the estimated snow depth on 1 December 2004 (first row), 1 January 2005 (second row), 1 February 2005 (third row), and 1 March 2005 (fourth row). Scheme-1 (left column) shows the interpolation results at 46 meteorological stations, and scheme-2 (right column) shows the results at 46 meteorological stations and virtual stations.
Figure 7. Comparison of the observed and simulated snow depth at four meteorological stations from November 2004 to March 2005. (a) Fukang, (b) Beitashan, (c) Kelamayi, and (d) Tekesi.