Developing a Cloud-Reduced MODIS Surface Reflectance Product for Snow Cover Mapping in Mountainous Regions

Amin Khoramian and Alireza B. Dariane *

Department of Civil Engineering, K.N. Toosi University of Technology, Tehran 1996715433, Iran; akhoramian@mail.kntu.ac.ir
* Correspondence: borhani@kntu.ac.ir; Tel.: +98-218-877-9474

Academic Editors: Ulrich Kamp and Jesús Martínez-Frias
Received: 22 February 2017; Accepted: 12 April 2017; Published: 17 April 2017

Abstract: Cloud obscuration is a major problem for using Moderate Resolution Imaging Spectroradiometer (MODIS) images in different applications. This issue poses serious difficulties in monitoring the snow cover in mountainous regions due to high cloudiness in such areas. To overcome this, different cloud removal methods have been developed in the past where most of them use MODIS snow cover products and spatiotemporal dependencies of snow to estimate the undercloud coverage. In this study, a new approach is adopted that uses surface reflectance data in the cloud-free pixels and estimates the surface reflectance of a cloudy pixel as if there were no cloud. This estimation is obtained by subsequently applying the k-nearest neighbor and dynamic time compositing methods. The modified surface reflectance data are then utilized as inputs of a Normalized Difference Snow Index (NDSI)-based algorithm to map snow cover in the study area. The results indicate that the suggested approach is able to appropriately estimate undercloud surface reflectance in bands 2, 4 and 6, and can map the snow cover with 97% accuracy, which is a substantial improvement over the conventional method with an accuracy of 86%. Finally, although a clear underestimation of snow cover (about 15%) is observed by applying the proposed approach, still, it is much better than the 30% underestimation obtained by the conventional method.

Keywords: snow cover; MODIS; cloud removal; surface reflectance; mountainous regions

1. Introduction

Snow is an important component of hydrological cycle. Accumulation in cold season and melt in warm season provide a large portion of water resources in different regions around the world. Thus, many efforts have been made to measure and accurately estimate it in river basins for better management of water resources. In situ measurements are the most common way to quantify snow variations. Such measurements are costly and limited to a few points that are mostly located in lower elevations of basins. However, snow varies drastically in temporal and spatial extent and it is mainly concentrated at higher elevations. In recent decades, the development of remote sensing technology has offered an alternative for snow monitoring. This technology provides near global coverage, high temporal and spatial resolution, and is able to measure different parameters.

Among others, the Moderate Resolution Imaging Spectroradiometer (MODIS) images are widely used for snow cover monitoring in recent years. Fine 500 m resolution, daily revisit time, well scientific developer team, and ease of access have made MODIS images a popular dataset for hydrological modeling [1–4] and climatologic studies [5–9]. However, the obscuration of earth surface by clouds in visible and near-infrared wavelengths causes a major challenge in snow cover mapping through MODIS images [10–13]. Different methods have been proposed to overcome this problem [14–17].
Most of these studies have focused on MODIS snow cover products (SCPs) (NASA, Washington, DC, USA), namely MOD10A1 and MYD10A1. These products contain binary information indicating either the presence or absence of snow or cloud. Therefore, every pixel in an image would be classified as one of the categories including snow, land or cloud. The applied methods employ the spatiotemporal dependencies to reclassify the pixels that are assigned as cloud to snow or land. For example, Hall et al. [16] represented a methodology to replace cloud-covered pixels by most recent cloud-free observation. The results indicated a 96% increase in observable pixels. Spatial dependencies were applied by Parajka et al. [18], where they used relative position of cloudy pixels to regional snow-line elevation to reclassify them. The applied method was validated using climate stations with snow depth measurements. The results indicated that the proposed method could reduce cloud coverage by 50% without any considerable change in snow-mapping accuracy. Dong and Menzel [19] introduced a conditional probability interpolation method using in situ snow depth measurements to overcome cloud obscuration in southwestern Germany. The presented methodology showed an accuracy of 92% during snow season. Combination of different methods was investigated by several studies [14,15,20,21]. They applied a set of spatial, temporal and multi-sensor methods in a sequence-based algorithm to reduce cloud coverage. For example, Da Ronco and De Michele [14] combined five cloud removal methods subsequently including combination of Aqua and Terra SCPs, conservative temporal combination, regional snow-line method, eight-day backward temporal and seasonal filter. The results indicated that combination of methods could remove clouds efficiently with an accuracy of 95%. Considering seasonal variations in accuracy of different cloud removal methods [20], Dariane et al. [21] developed a cloud removal algorithm to combine different methods to reduce cloudiness along with maintaining the accuracy of snow mapping. Their algorithm was able to remove 94% of clouds and preserve the accuracy by 93%. All of these studies use MODIS SCPs as input dataset of cloud removal methods. In this study, we investigated an innovative approach to achieve a cloud-reduced snow cover map. This approach utilizes MODIS surface reflectance data and estimates cloud-covered pixels’ surface reflectance as it would have been measured at earth surface level and as if there were no clouds. It means that we changed the problem of estimating pixel coverage in cloudy condition (i.e., land or snow) to the estimation of earth surface reflectance when the pixel is cloud-covered. The estimated surface reflectance would be used as input of an algorithm to map snow cover in the study area. This new approach is implemented and compared to the conventional approach (i.e., using snow cover maps) that has been applied by the previous studies.

This paper is structured as follows: study area and the data sets are described in Section 2. Section 3 displays the methodology used in this paper. Results and discussions are presented in Section 4 and finally the study findings are concluded in Section 5.

2. Study Area and Data Sets

2.1. Study Area

The developed methodology is applied for Central Alborz region (35.2°–36.8° N. latitude, 50.3°–53.3° E. longitude) in north of Iran (Figure 1). Central Alborz is the middle part of the Alborz Mountain Ranges that supplies water for a population of about 21 million in the region. It covers an area of about 26,800 km² and its elevation ranges from 0 to 5604 m a.s.l.. The region is subdivided into two main slopes. The northern slope climate is humid due to movement of moist air from the Caspian Sea. The mean precipitation over this slope is about 700 mm per year. The southern slope is located on the leeward side of the Alborz Mountain Ranges. Its climate is semi-arid with mean precipitation of about 250 mm per year [22]. The vegetation cover in Central Alborz is not dense and is dominated by alpine meadows and consists mainly of herbaceous and xerophilous plants [23]. The study area consists of nine major river basins including Latian, Haraz, Taleghan, Lar, Mamloo, Karaj, Firoozkuh, Sehezar, and Chaloos, with several tributaries.
2.2. Remotely Sensed Data

MODIS sensor captures data in 36 bands from visible to thermal infrared. We utilized MODIS/Terra surface reflectance daily L2G (MOD09GA) product, which provides bands 1–7 surface reflectance data with daily revisit time and 500 m spatial resolution. This product is corrected for aerosols and atmospheric gases and is associated with data product state quality assurance scientific datasets that could be used to determine cloud state in every pixel [24]. This dataset is obtained from the MODIS cloud mask (MOD35_L2) product. To cover the study period from 1 October 2002 to 31 September 2015, 4748 of h22v05 tiles of MOD09GA version 6 are acquired through the Reverb website [25] and are preprocessed. In this study, bands 2, 4 and 6 would be used in snow cover mapping algorithm and cloud state dataset would be utilized to determine cloud condition as cloud flag.

3. Methodology

The adopted methodology estimates undercloud surface reflectance using a sequence-based combination of k-nearest neighbor method and dynamic time compositing. Then, estimated values would be used in snow cover detection algorithm to map snow cover in the study area. To identify the superiority of the proposed approach over the conventional method, a cloud removal algorithm using the same methods but with priority to snow mapping is implemented in parallel. The conventional approach uses cloud-contaminated surface reflectance data to map the snow cover and clouds. Then, it applies a cloud removal procedure to detect the surface coverage under the cloud. A scheme of both algorithms is given in Figure 2. To validate both algorithms, a validation methodology is adopted that fills cloud-free images with factitious clouds and assumes cloud-free images as the ground-truth for cloud-filled images. Detailed descriptions are presented in the sections that follow.
3.1. Cloud Removal Methods

3.1.1. K-Nearest Neighbors

This method uses eight neighbor pixels to estimate surface reflectance in a cloud-covered pixel. The estimation is obtained using the following equations:

\[
SF(x, y) = \frac{\sum_{i=-1}^{1} \sum_{j=-1}^{1} w(x + i, y + j) \times f(x + i, y + j)}{\sum_{i=-1}^{1} \sum_{j=-1}^{1} w(x + i, y + j)} \quad i, j \neq 0 \text{ and } \text{cloudflag}(x + i, y + j) \neq 1
\]

\[
w(x', y') = \frac{1}{\sqrt{(x - x')^2 + (y - y')^2}}
\]

where \(SF(x, y)\) is the target function and is equal to surface reflectance of a pixel in \(x\)-th column and \(y\)-th row, \(w(x', y')\) is the weighting function that returns the weight of every neighbor pixel at \(x'\)-th column and \(y'\)-th row based on its Euclidean distance. \(\text{Cloudflag}\) is the cloud state function that is equal to 1 and 0 for cloud and non-cloud conditions, respectively. In conventional approach, the target is estimating undercloud coverage (snow or land) and is obtained by assigning the most common coverage among eight neighbor pixels to the cloudy pixel. It means that if most of the neighbor pixels are classified as snow, the cloudy pixel would be classified as snow and vice versa. In the event that the number of cloudy and snow-covered pixels is equal, no decision would be made.

3.1.2. Dynamic Time Composite

This method employs images of prior and subsequent days to estimate undercloud surface reflectance in a cloudy pixel. The time window that would be employed is two, four and six days. At the first step, this method tries to estimate surface reflectance using images of the preceding and following days. If the corresponding pixels were cloudy on both days, the method would use images of two days before and two days after. Finally, if these images were also cloud-covered, then images of three days before and after would be utilized. The employed days in each step are shown in Figure 3. The estimated undercloud surface reflectance would be equal to the (average) observed surface reflectance of non-cloudy day(s).
3.1.3. Combination of Methods

The two above-described methods are combined in a sequence-based scheme. In this way, the k-nearest neighbor method is applied at first. This method is useful for the marginal pixels of clouds and/or cases with sparse small clouds. In the situation of large cloud areas, this method cannot work well. In such cases, the dynamic time composite method would be suitable and can remove the remaining parts of clouds. After applying each method, if the cloudiness is higher than a predetermined desired value (here 1%) the next method is applied. Figure 4 shows the flowchart of methods combination.

In conventional approach, the same steps would be used, but instead of surface reflectance values, the undercloud observation (snow or land) of the preceding or following days would be assigned to the cloudy pixel. In case of inconsistency, no action would be taken for the cloudy day.

3.2. Snow Mapping

The applied snow-mapping algorithm is based on the Normalized Difference Snow Index (NDSI). This index is the normalized difference of band 4 (visible) and band 6 (shortwave infrared) in MODIS onboard Terra satellite. A pixel is classified as snow if the three following conditions are met [26].

1. \[ \text{NDSI} = \frac{\text{Band 4} - \text{Band 6}}{\text{Band 4} + \text{Band 6}} > 0.4 \]
2. Band 2 reflectance > 11%
3. Band 4 reflectance > 10%

![Figure 3. Utilized days in each step for dynamic time composite.](image)

![Figure 4. Flowchart of methods combination.](image)
Figure 5 shows the flowchart of the applied snow-mapping algorithm.

3.3. Validation Methodology

Due to the absence of in situ measurements, we used a validation methodology presented by Gafurov and Bardossy [15] and improved by Dariane et al. [21]. This method is based on assuming selected images with lowest cloudiness as ground truth and filling them with a cloud mask that is borrowed from among highly clouded images. Then, by applying cloud removal methods on cloud-filled images and comparing the resulted images with the corresponding ground truths (images with least cloud), the accuracy of cloud removal procedure is assessed. Hence, the validation results would show the agreement between low cloud images and cloud-removed ones. Therefore, in this study, the accuracy term is equal to agreement between these images in this study.

Using this methodology, 156 images (one image in every month during 13 years of study period) are selected as clear-sky images and filled with low, medium and high cloud masks. Cloud masks are picked out based on 25%, 50% and 75% probability of occurrences [21].

The results are validated in two forms. In the first, estimated surface reflectance values are validated using three criteria, including Root of Mean Square Error (RMSE), Mean Bias (MB), Coefficient of Determination ($R^2$) and Pearson’s Correlation Coefficient ($r$). In the second form, snow-mapping accuracy is investigated. In this form, agreement between reclassified pixels and their corresponding pixels in ground truth images is assessed and is reported via a confusion matrix. This matrix shows the number of pixels that are correctly classified as snow (snow to snow), pixels that are misclassified as snow (land to snow), pixels that are correctly classified as land (land to land) and pixels that misclassified as land (snow to land). The total degree of agreement could be obtained as the rate of correctly classified pixels (snow to snow and land to land) to all the reclassified pixels.

4. Results

4.1. Cloud Removal Performance

The performance of a cloud removal method was assessed by the amount of clouds that could be removed applying the presented methodology. Figure 6 shows the time series of cloud coverage...
before and after cloud removal. In addition, monthly average cloud coverage during the study period is given in Table 1.

![Cloud coverage graph](image)

**Figure 6.** Time series of cloud coverage before and after cloud removal.

<table>
<thead>
<tr>
<th>Cloudiness</th>
<th>Winter</th>
<th>Autumn</th>
<th>Spring</th>
<th>Summer</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original amount of cloud (%)</td>
<td>56.3</td>
<td>38.0</td>
<td>18.7</td>
<td>38.3</td>
<td>37.8</td>
</tr>
<tr>
<td>Remained cloud (%)</td>
<td>4.3</td>
<td>1.7</td>
<td>1.0</td>
<td>2.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Percent reduction</td>
<td>92.3</td>
<td>95.6</td>
<td>94.9</td>
<td>94.6</td>
<td>94.1</td>
</tr>
</tbody>
</table>

As can be seen from Table 1, the cloud removal substantially reduces the cloud contamination and removes about 94% of clouds during the study period. It is worthwhile to mention that both approaches would perform similarly in reducing the cloud coverage. In fact, they are essentially similar in this step although having different target functions, i.e., a continuous function in the presented approach and a discrete function in conventional approach. The main difference between the two approaches is on detecting the nature of area under the cloudy pixels.

### 4.2. Undercloud Surface Reflectance Estimation

Undercloud surface reflectance was estimated in bands 2, 4 and 6 following the proposed approach. These bands were used in snow detection algorithm. In overall, surface reflectance of 112,737 pixels were reconstructed. Observed and estimated surface reflectance values are plotted against each other as shown in Figure 6. As can be seen from Table 2 as well as Figure 7, the proposed algorithm works well in bands 2 and 4; however, the accuracy of band 6 is not as good as the other two. The MB and RMSE are both under 0.1 in this band.

<table>
<thead>
<tr>
<th>Validation Criteria</th>
<th>Band 2</th>
<th>Band 4</th>
<th>Band 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Bias</td>
<td>0.066</td>
<td>0.054</td>
<td>0.051</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.09</td>
<td>0.082</td>
<td>0.063</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.75</td>
<td>0.85</td>
<td>0.61</td>
</tr>
<tr>
<td>$r$</td>
<td>0.87</td>
<td>0.92</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Table 2. Validation criteria for bands 2, 4 and 6. RMSE: Root of Mean Square Error.
estimating error, the NDSI error does not effectively alter the precision of snow mapping by the proposed method.

### 4.3. Snow Mapping

As illustrated in Section 3.3, snow-mapping accuracy was investigated as the number of correctly classified and misclassified pixels in a confusion matrix form. Tables 3 and 4 are the validation results for the two applied approaches. As is indicated in Table 3, 85.3% and 100% of snow and land reclassified pixels agree with original non-clouded snow cover map using the proposed approach, whereas, by applying the conventional approach, 70.2% and 91.4% of snow and land pixels are correctly reclassified (Table 4).

#### Table 3. Confusion matrix for the presented approach.

<table>
<thead>
<tr>
<th>Original Snow Map</th>
<th>Cloud-Removed Snow Map</th>
<th>In Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow</td>
<td>Snow 22,146 (85.3%)</td>
<td>25,976</td>
</tr>
<tr>
<td></td>
<td>Land 0 (0%)</td>
<td>86,761</td>
</tr>
<tr>
<td></td>
<td>In overall 22,146</td>
<td>112,737</td>
</tr>
</tbody>
</table>

Overall accuracy: 97%

#### Table 4. Confusion matrix for the conventional approach.

<table>
<thead>
<tr>
<th>Original Snow Map</th>
<th>Cloud-Removed Snow Map</th>
<th>In Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow</td>
<td>Snow 18,224 (70.2%)</td>
<td>25,976</td>
</tr>
<tr>
<td></td>
<td>Land 7422 (8.6%)</td>
<td>86,761</td>
</tr>
<tr>
<td></td>
<td>In overall 22,146</td>
<td>112,737</td>
</tr>
</tbody>
</table>

Overall accuracy: 86%
The overall accuracy of the proposed and conventional approaches is 97% and 86% in the study area, respectively. These results show that the proposed approach can reasonably improve the cloud removal accuracy in comparison to the conventional approach. Although a clear underestimation of snow cover (about 15%) can be found by applying the proposed approach, still, it is much better than the 30% underestimation obtained by the conventional method. A sample of obtained snow maps using the presented approach is given in Figure 8. This figure shows snow and cloud coverage before and after cloud removal at 1 February 2015. Also, variation of snow cover extent over the study region is estimated using cloud-removed snow maps (Figure 8c).

Figure 8. Snow and cloud maps of study region (a) before and (b) after implementation of proposed cloud removal approach on 1 February 2015, and (c) time series of snow cover extent extracted from obtained snow maps during the study period (2002–2015).

5. Discussion and Conclusions

Cloud obscuration is a major challenge in using MODIS images for snow cover mapping. Several studies have addressed this issue and have developed different methods to overcome this problem. Using spatiotemporal dependencies between MODIS snow cover binary maps is the cornerstone of all the past studies. In this work, we presented a new approach to reach a cloud-reduced snow cover map. The proposed approach estimates undercloud surface reflectance in the cloudy condition and uses estimated surface reflectance to map snow cover by applying a snow detection algorithm. This approach was compared to the conventional approach that, in the first step, maps the snow cover and labels the cloud-contaminated pixels as cloudy, and then approximates the undercloud area condition. The proposed approach utilized two sequential methods including k-nearest neighbor and dynamic time composite to estimate undercloud surface reflectance values.

Although the in situ measurements have to be used as ground-truth data, due to the absence of such measurements, an alternative methodology was followed. According to this approach, cloud-free images are assumed as ground truth data. Hence, the accuracy assessment results showed the
agreement between cloud-removed images and cloud-free ones. Thus, cloud-removed snow cover maps would be with some uncertainties that are reported in previous studies [19,27].

Validation was done in two ways. In the first one, the estimated surface reflectance values in bands 2, 4 and 6 were validated. The results showed that the proposed approach is able to estimate the surface reflectance values with high precision. In the second way, mapped snow cover was compared with ground-truth images. The comparison indicated that snow condition could be classified correctly with the accuracy of 97% and 86% using the proposed approach and the conventional method, respectively, which indicates the superiority of the proposed approach. Finally, although a clear underestimation of snow cover (about 15%) was found by applying the proposed approach, still, it is much better than the 30% underestimation obtained by the conventional method. The superiority of the presented approach in the validation results could be due to better performance of cloud removal methods in approximating surface reflectance values (presented approach) in comparison to estimating binary snow cover conditions (conventional approach). In other words, approximating a real value could cause less error than a binary value potentially. In a binary system, only two values (0 or 1) can be selected, but in real numbers system, a continuous set of values are feasible. Therefore, a slight induction bias may produce less error in an obtained snow map using the presented approach. This reason is subjective and needs to be investigated more precisely in future studies.

The presented methodology is applicable in other snow-covered mountainous regions all over the world. However, the performance of method in terms of cloud reduction and the accuracy depends on the persistency of clouds and heterogeneity of land coverage. More continuous cloud persistency in the region demands an increased number of needed pervious and following day’s images. Under this situation, the uncertainty with estimated snow condition would increase. Also, the heterogeneity of land surface would pose the k-nearest neighbor method with difficulties in accurate estimation of undercloud condition.

Author Contributions: This research article is part of the Ph.D. level dissertation of the co-author, Amin Khoramian, under the supervision of the corresponding author, Alireza B. Dariane.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Franz, K.J.; Karsten, L.R. Calibration of a distributed snow model using MODIS snow covered area data. J. Hydrol. 2013, 494, 160–175. [CrossRef]
7. Tahir, A.A.; Chevallier, P.; Arnaud, Y.; Ashraf, M.; Bhatti, M.T. Snow cover trend and hydrological characteristics of the Astore River basin (Western Himalayas) and its comparison to the Hunza basin (Karakoram region). Sci. Total Environ. 2014, 505, 748–761. [CrossRef] [PubMed]
8. She, J.; Zhang, Y.; Li, X.; Feng, X. Spatial and temporal characteristics of snow cover in the Tizinafu watershed of the Western Kunlun Mountains. Remote Sens. 2015, 7, 3426–3445. [CrossRef]
27. Dong, C.; Menzel, L. Improving the accuracy of MODIS 8-day snow products with in situ temperature and precipitation data. *J. Hydrol.* 2016, 534, 466–477. [CrossRef]