Assessing High-Resolution Satellite Imagery for Detailed Snow Cover Estimation: An Ecological Perspective

Tony Cannistra, Kavya Pradhan
Department of Biology
University of Washington
Seattle, WA 98105

Abstract

Understanding the responses of species to climate change requires a robust understanding of the impact of abiotic (environmental) factors on habitat suitability and habitat availability at spatial scales relevant to the study species. When studying plant species in montane ecosystems, snow cover is a particularly relevant abiotic factor in determining habitat suitability and predicting phenological changes. However, remotely-sensed snow cover measurements are either captured at a spatial scale far too large to be relevant to the study species (e.g. MODIS SCA), or are appropriate in spatial scale but cost-prohibitive (e.g. LIDAR snow measurements). Herein we evaluate the suitability of Planet Labs imagery, a commercial satellite imagery product with narrow spectral bandwidth (4 bands) but high spatial (0.7-3.0 m) and temporal (1-2 day) resolution, for the purpose of acquiring detailed snow-covered area maps at ecologically-relevant scales. We use a machine learning classification approach and spectral mixture analysis based on airborne and ground-based snow observations, and evaluate these approaches using both quantitative and qualitative methods. We find that both approaches perform well in specific circumstances, but performance is quite varied and depends on the study region. Therefore we suggest that developing regional models is important when using narrow bandwidth imagery, and offer methods to improve performance in future studies.

1 Introduction

Understanding the responses of species to climate change requires a robust understanding of the impact of abiotic (environmental) factors on habitat suitability and habitat availability at spatial scales relevant to the study species (Hannah et al. 2014). When studying plant species in montane ecosystems, snow cover is a particularly relevant abiotic factor in determining habitat suitability and predicting phenological changes. However, remotely-sensed snow cover measurements are either captured at a spatial scale far too large to be relevant to the study species (e.g. MODIS SCA), or are appropriate in spatial scale but cost-prohibitive (e.g. LIDAR snow measurements). Herein we evaluate the suitability of Planet Labs imagery, a commercial satellite imagery product with high spatial (0.7-3.0 m) and temporal (1-2 day) resolution, for the purpose of acquiring detailed snow-cover and snow-melt data at ecologically-relevant scales. The challenge herein is the radiometric bandwidth available from Planet imagery, which makes standard spectral snow cover indices like the Normalized Difference Snow Index (NDSI; Hall et al. 1995) unusable.

CEE 573: Snow Hydrology. Winter 2018 (Cristea). University of Washington, Seattle.

The explosive growth and availability of remotely-sensed environmental information is expanding the potential of geospatial research. With satellite-derived environmental variables, ecologists can evaluate research questions at broader spatial scales, no longer limited by the installation of ground-based instrumentation. However, a majority of the studies that use such data are forced to operate under the assumption that coarse (> 500-1000m resolution) climate information is sufficient to describe the environmental conditions that an individual organism is facing. Additionally, the necessity for temporal interpolation of rarely-captured environmental variables leads to the reduction of temporal accuracy in these studies. High spatial and temporal accuracy in these remotely-sensed environmental variables is important for the credibility of future studies reliant upon satellite data sources. Of particular interest is the observation of snow-covered area, which especially important in montane systems as snow drives much of the seasonal hydrological regimes and can have significant ecological impacts on plant communities. The recent perfusion of imagery with high spatiotemporal resolution may be able to bridge the gap between ground-based instrumentation and coarsely-captured satellite data.

Planet Labs is a promising source of high resolution imagery that can be used in ecological studies. However, its immediate utility with respect to inferring snow cover is limited due to the narrowness of the near infrared band which makes distinguishing snow from clouds difficult using a radiometric index (such as NDSI). As an alternative, we develop two methods of determining snow covered area—a machine learning algorithm and a spectral mixture analysis—and assess the effectiveness of each method. We couple ground and airborne snow observations with Planet Labs imagery in two montane systems in Washington and California, USA to perform this analysis.

2 Materials and Methods

2.1 Study Regions and *in situ* Snow Observations

Two snow-dominated regions were examined in this assessment: Mount Rainier, Washington, USA and the upper Tuolumne Basin, California, USA. Mount Rainier is a 4392m stratovolcano located in western Washington state and is an excellent candidate for studying the impacts of climate change on ecosystem dynamics. The mountain's 4026m prominence permits a wide span of ecotypes, ranging from boreal forest to glaciated alpine. This breadth facilitates a variety of research.

In particular, plant communities in alpine and subalpine meadows are heavily studied, and snow cover is a central abiotic factor in these forest-alpine transition regions. Current studies that have attempted to elucidate the effects of snow cover and snow seasonality on subalpine ecosystems have installed point in-situ snow presence instrumentation in the form of ground surface temperature sensors. This methodology leverages the insulating properties of snow—during periods of snow presence, surface-level temperature sensors measure a constant 0 °C, and during snow absence, air temperature is measured. During our chosen study year (2017) we have access to snow cover period data for 20 *in situ* snow measurement locations across an elevational band spanning the boreal and subalpine regions of Mt. Rainier.

The other study region examined in this report is the Upper Tuolumne basin of eastern California. This region spans a considerably larger spatial extent but a narrower band of ecotypes, ranging from high boreal forest to alpine. The basin has been the focus of considerable hydrologic research for some time. In particular, the NASA/JPL Airborne Snow Observatory (ASO) project has used airborne Light Detection and Ranging (LIDAR) technology to measure weekly to monthly snow water equivalent and surface albedo at 30m resolution for the entire basin since 2014 (Painter et al. 2016). To establish ground-truth snow presence and absence information we acquire ASO SWE rasters for 2017-04-01 (majority snow) and 2016-06-20 (majority dry/melt). To assess snow extent we extract all pixels with SWE greater than 0.0m in the 2017-04-01 raster and consider the resultant pixel "snow-containing." Conversely we extract those pixels with SWE of 0.0m ($\pm$0.001m) in the melt season raster (2016-06-20) and consider those "snow free."

2.2 Satellite Imagery

We utilize Planet Labs PlanetScope "analytic" 4-band imagery captured during snow-containing and snow-free periods for both study sites as input to our analytical models. The PlanetScope imagery product contains 4 16-bit radiometric bands (Blue: 455-515 nm, Green: 500-590 nm, Red: 590-670 nm, NIR: 780-860 nm) captured at 3m resolution with a near-daily revisit across the Earth land surface (https://planet.com). Planet Labs imagery is available to researchers through the Planet Ambassadors program, which offers access to 10,000 km^2 / month of imagery without cost for a specific research objective.

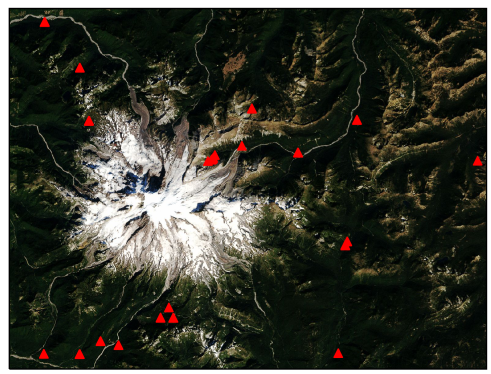

Figure 1: 20 *In situ* snow measurement stations on Mt. Rainier (2017).

On Mt. Rainier we choose at random 10 days from within the snow-cover period for each of the 20 in situ snow observation locations and acquire PlanetScope "analytic" imagery cropped to a rectangle 0.00° around the point measurement location. Similarly, we choose 10 days at random from within the snow-free period for each of the 20 observation sites and perform the same extraction. After georeferencing the in situ measurement to each image we further extract raw pixel values from the 10 pixels directly adjacent to the point measurement from each image as the focal pixels to be used in analysis.

The scale of the Upper Tuolumne basin precludes the above extraction from being practical due to the volume of data (the basin is covered by about 30 overlapping PlanetScope scenes, each containing about 340km^2). Instead we acquire entire PlanetScope scenes cropped to the snow-extent or snow-free raster regions such that the remaining pixels satisfy either a snow-on or snow-off condition as defined by ASO-derived SWE thresholds (see above). We then randomly select 10% of those resulting pixels as focal pixels for our models, described below.

2.3 Machine Learning Classification

A central motivator for this project is the good performance of machine learning techniques in other remote sensing applications, such as object identification (Cheng and Han, 2016). The success of algorithmic approaches in these domains led us to become interested in discovering whether a similar approach could be applied to the challenge of the spectrally-narrow Planet imagery for the purpose of snow cover classification. To this end we employ a classification algorithm to discriminate between snow-containing pixels and snow-free pixels.

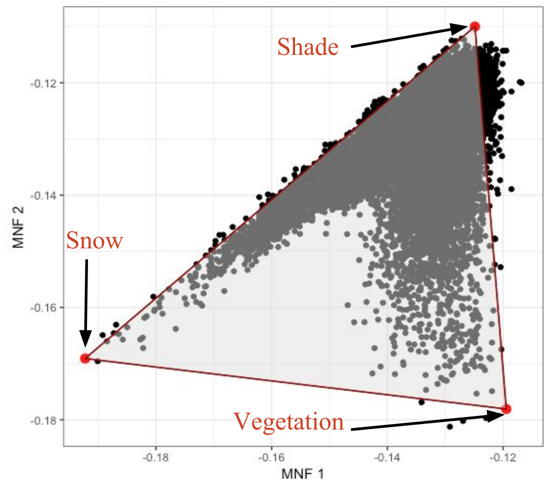

Figure 2: Example MNF feature space (of the first two axes) with selected endmember pixels (in red).

The core algorithmic framework for this project is a Gaussian Process classifier coupled with a radial basis function (RBF, or squared exponential) kernel, as described in Rasmussen and Williams, 2006. This approach is a general nonparametric machine learning algorithm capable of fitting to any function given by training data. Theoretical model performance was assessed by examining quantitative cross validation-based metrics of classification accuracy (percent accuracy) using pixels extracted from training data as described above. This process was performed differently for each study region. At Mt. Rainier we performed a 5-fold stratified cross validation analysis on the pixels adjacent to in-situ snow observation locations as described above. For the Upper Tuolumne basin, a sample of 10% of all ground-truth pixels was used in the same 5-fold stratified cross validation. Mean classification accuracy was computed across all folds for each site. We then trained the model on the entire training data and qualitatively inspected model predictions of pixels on actual images.

2.4 Spectral Mixture Analysis

Spectral mixture analysis (SMA; Adams and Adams 1984; Mustard and Peiters 1989) is based on the concept of a mixed pixel, i.e, pixels in remotely sensed images are rarely spectrally pure. As such, SMA can be used to identify the composition of a pixel based on the fractional contribution of each spectral component to the total pixel reflectance. SMA has been used in various remote sensing applications such as land cover classification, terrestrial ecosystem modelling, precision agriculture, and water resource management (Somers et al. 2011). In addition to these application, SMA has also been used to assess snow cover using AVIRIS (Painter et al. 1998), Landsat (Pimentel et al. 2015), and MODIS imagery (Rittger et al. 2013). Based on the successful use of SMA in various remote sensing applications, we wanted to assess whether it would be a good method for obtaining snow on/off from Planet imagery as well. In addition to this, SMA's application of multiple bands to identify components of a pixel the narrow spectral bandwidth of Planet data poses limitation for converting spectral information to snow cover information.

Although there are non-linear forms of SMA, for the purposes of this study SMA was implemented in its linear form wherein the spectral information in each pixel is considered to be the linear combination of the spectral signature of and the fraction of pixel occupied by each component of the pixel. Because SMA takes in information from many bands, image data is often subjected to dimensionality or noise reduction techniques. Here we used the Minimum Noise Fraction (MNF) transformation to our data and used the first two MNF components. MNF transformation is very similar to the commonly used Principal Component Axis (PCA) rotation but arranges information in order of increasing noise rather than decreasing variation.

After denoising the data, the pixel components were selected using the pixel purity index (PPI) endmember extraction algorithm (Boardman et al. 1995). These components of pixels are referred to as endmembers and are pixels that contain only one physical component, i.e., they are spectrally pure. In common application of SMA, endmembers can be obtained from a curated library, but in this study we chose to use endmembers from the images itself (Adams et al. 1993) since there is no such library available for Planet data with reference to snow assessment. Because we had four spectral bands, we chose to use three endmembers representing snow, shade (or water), and vegetation based on the PPI (higher is more likely to be selected; Figure 2). The selected endmembers for each image were then finally used to obtain the proportion of each endmember for each pixel. The snow fraction of each pixel was then converted to snow on/off using a manually selected threshold.

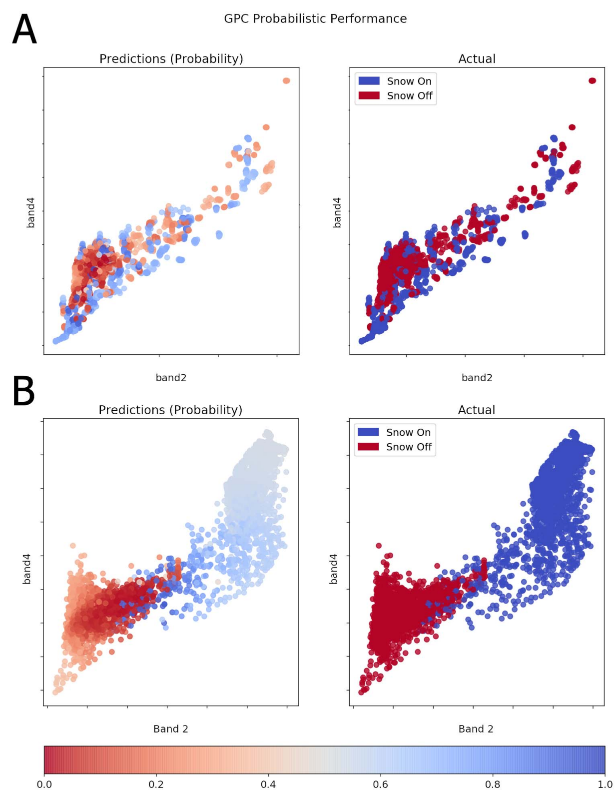

Figure 3: Machine learning models perform well on both datasets in cross validation. Classification performance on Mt. Rainier (row A) and Tuolumne Basin (row B) data, plotted across two imagery bands (Green and NIR). Each point represents an imagery pixel. Left column is model predictions, right column is actual ground-truth labels. Colorbar indicates model probability of snow presence.

The above method was applied to a total of 15 images, each containing a temperature sensor in MRNP. Note that this method was not applied to the Upper Tuolumne basin.

3 Results

Both spectral mixture analysis and the machine learning approach show some ability to discriminate snow-containing pixels from snow-free pixels. Cross-validation performance of the machine learning approach is excellent for both Mt. Rainier data (mean accuracy 96.2% in 5-fold stratified cross validation) and the Tuolumne Basin data (mean accuracy 99.9% in 5-fold stratified cross validation). **Figure 3** demonstrates this good classification performance on both datasets: the predictions are similar to the actual data labels. Significant departures from this exceptional performance can be observed in qualitative inspection of snow-covered area predictions on actual images, as evidenced in **Figure 4**. In particular, snow-covered pixels (as determined qualitatively in the RGB image) are assigned a 50% probability of containing snow by the model (white colors, center panel), which leads to an incorrect prediction of the snow-free state (black colors, right panel).

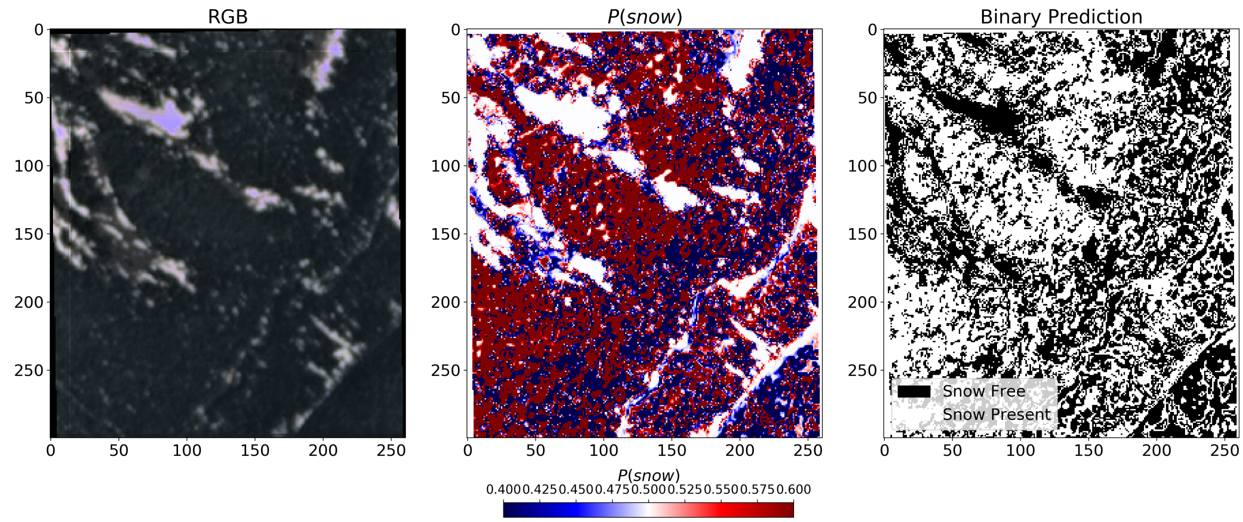

Figure 4: Poor qualitative model performance in snow cover estimation on a real image from Mt. Rainier. Left to right: RGB image, machine learning model derived probability of snow presence, binary model prediction. The model becomes quite uncertain how to predict in regions of snow (center panel, white colors), and chooses a snow-off label (right panel, black colors).

The model has observable qualitative improvement using data from the upper Tuolumne basin (Figure 5).

Although SMA showed initial promise in its ability to discriminate between snow and other image components (Figure 6), it did not perform well when images had a considerable amount of shade (Figure 6) or clouds (Figure 6). These results were obtained from the application of SMA to MRNP. Due to this poor performance of SMA within MRNP, we chose to forgo its application in the Upper Tuolumne Basin.

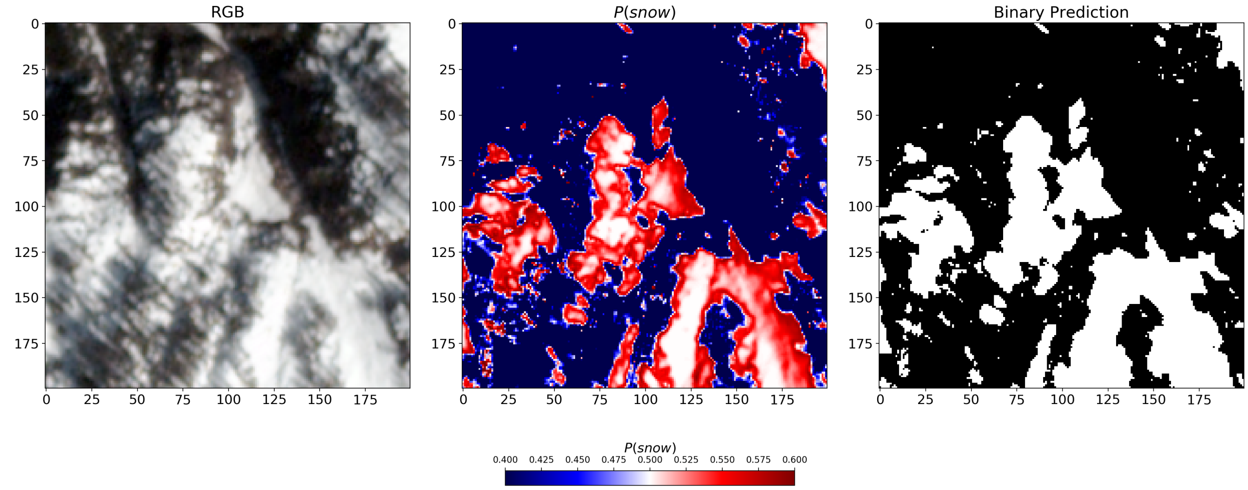

Figure 5: Improved snow cover identification performance in Tuolumne Basin. Left to right: RGB satellite image, model-derived probability of snow presence, and binary model prediction (white: snow).

4 Discussion

Herein we demonstrate, using both spectral mixture analysis and a machine learning classifier, the challenge of using imagery with low spectral bandwidth for the purpose of assessing snow cover in montane regions.

We observe that the machine learning method performs differently depending on the geographic region in which it was trained. This suggests that developing regional snow-covered area models is an important step when using narrow-bandwidth imagery. We hypothesize that the reason for significantly degraded qualitative performance (when compared to theoretical cross-validation performance) in the Mt. Rainier region is the pervasive cloudiness which plagues the Planet imagery during the snow season. (Conversely, the nature of the imagery used in the Tuolumne basin allows for a cloud-free dataset, which may explain better performance). Further research may add sensitivity to imagery cloudiness (perhaps via an additional cloud-detection step) to the image acquisition pipeline to improve the model training data, thereby enhancing qualitative snow-covered area identification.

Results from the spectral mixture analysis show mixed performance of this method depending on the composition of the image. Images without a considerable amount of shade or cloud are more tractable for use with SMA than images with a lot of shade, a lot of clouds or both. We suspect that the problem may lie in one of two possible steps. The first one is the application of a denoising using MNF transformation. In general MNF or other transformations are an important aspect of SMA since this technique is usually used with multiband or hyperspectral imagery. However, application of a noise reduction technique with a four band dataset may have led to information loss. An alternative to this issue would be to forgo denoising before the selection of endmembers. The other issue might be with the process used to select endmembers itself. Incorrect selection of endmembers would ultimately lead to the fractional division of an image

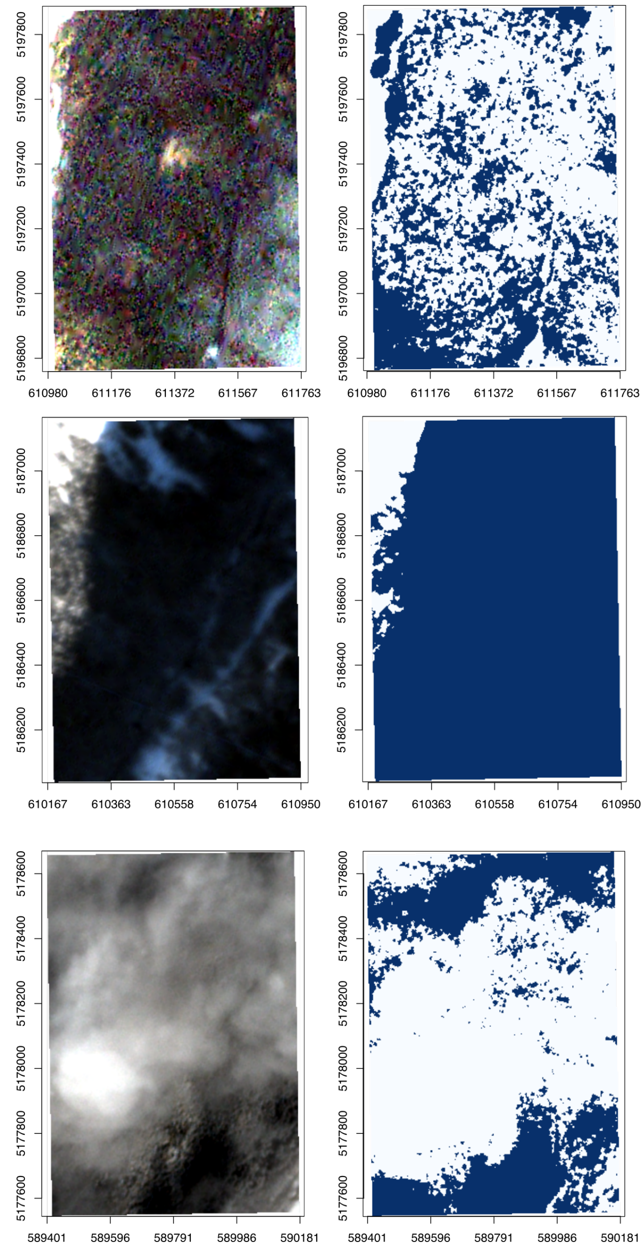

Figure 6: True color composites of Planet imagery (left column) compared to snow cover obtained using SMA (right column) in regions with less shade and cloud (first row), more shade (second row), and more clouds (third row). Images are located in MRNP.

into components that are either not spectrally pure, or not of interest. Further research can be targeted at improving endmember selection or developing an endmember library for Planet data.

5 References

Adams, J. B., Smith, M.O., Gillespie, A.R., (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. *Remote geochemical analysis: Elemental and mineralogical composition*, 145-166.

Adams, J.B., Adams, J.D., (1984). Geologic mapping using Landsat MSS and TM images: Removing vegetation by modeling spectral mixtures. *Third Thematic Conference on Remote Sensing for Exploration Geology*, 615-622.

Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). Mapping target signatures via partial unmixing of AVIRIS data.

Cheng, G., Han, J. A Survey on Object Detection in Optical Remote Sensing Images. (2016). *ISPRS Journal of Photogrammetry and Remote Sensing*. 117, 11-28. doi:10.1016/j.isprsjprs.2016.03.014

Hall, D.K., G.A. Riggs, V.V. Salomonson., (1995). Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectrophotometer data. *Remote Sensing Environment*, 54 , 127-140.

Hannah, L., Flint, L., Syphard, A.D., Moritz, M.A., Buckley, L.B., McCullough, I.M., (2014). Fine-grain modeling of species' response to climate change: holdouts, stepping-stones, and microrefugia. *Trends in Ecology & Evolution*. 29, 390-397.

Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning. (2009). *The Mathematical Intelligencer* 27, (Springer New York).

Mustard, J.F, Pieters, C. M., (1989). Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra. *Journal of geophysical research: Solid Earth*, 94, 13619-13634

Painter, T. H. et al., (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sens. Environ.* 184, 139–152 .

Painter, T. H., Roberts, D. A., Green, R. O., Dozier, J. (1998). The effect of grain size on spectral mixture analysis of snow-covered area from AVIRIS data. *Remote Sensing of Environment*, 65(3), 320-332.

Pimentel, R., Herrero, J., Polo, M. J. (2015). Snow evolution in a semi-arid mountainous area combining snow modelling and Landsat spectral mixture analysis. Proceedings of the International Association of Hydrological Sciences, 368, 33-39.

Rittger, K., Painter, T. H., Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, 51, 367-380.

Somers, B., Asner, G. P., Tits, L., Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment*, 115(7), 1603-1616.