Machine Learning Classification and Derived Snow Metrics from Very-high-resolution Multispectral Satellite Imagery in Complex Terrain

Michelle Hu[1] and David Shean[1]

[1]University of Washington

November 22, 2022

Abstract

Satellite remote sensing often requires a compromise between spatial resolution and spatial coverage for timely and accurate measurements of earth-system processes. But in recent years, increased availability of submeter-scale imagery dramatically altered this balance. Commercial satellite imagery from DigitalGlobe and Planet offer on-demand, very high-resolution panchromatic stereo and multispectral (MS) image collection over snow-covered landscapes, with individual image coverage of up to ~1900 km2. Repeat stereo-derived digital elevation models can be used to accurately estimate snow depth. Integration of contemporaneous ~1–2 m land cover classification maps can provide precise snow-covered area (SCA) products and improved processing, analysis, and interpretation of these snow depth estimates. We are developing machine learning classification algorithms to identify snow, vegetation, water, and exposed rock using varying combinations of available bands (panchromatic, 4/8-band multispectral, SWIR) and band ratios (e.g. NDVI, NDSI) from these products. We present findings for NASA SnowEx campaign sites (Grand Mesa and Senator Beck Basin, CO) and other snow monitoring sites in the Western U.S. using WorldView-3, PlanetScope, and Landsat 8 imagery. Preliminary results show that a tuned random forest algorithm using WorldView-3 MS and SWIR bands yielded the most accurate estimates of SCA of all band combinations and imagery products. With the power to resolve individual trees, these products offer direct measurements of SCA, without the need to account for mixed pixels and fractional SCA as with lower-resolution products. This open-source workflow will be used to process longer time-series and larger areas in a semi-automated fashion, allowing for rapid analysis, increased portability, and broader utility for the community.

Machine Learning Classification and Derived Snow Metrics from Very-high-resolution Multispectral Satellite Imagery in Complex Terrain

#C33E-1619

Michelle Hu and David Shean

Terrain Analysis and Cryosphere Observation Lab
University of Washington | Civil and Environmental Engineering
https://jmichellehu.github.io/   jmhu@uw.edu

This work presents a preliminary semi-automated, open-source workflow to train machine learning models for classifying very-high-resolution panchromatic and multispectral imagery from commercial vendors (DigitalGlobe and Planet).

These models identify a subset of land cover classes and can directly quantify snow-covered area at sub-meter resolution. Resulting classifications will be used to improve snow depth measurements from contemporaneous stereo imagery through point cloud filtering and improving DEM co-registration.

Random forest model results for WorldView-3 imagery are shown here, but deep learning models that leverage spatial information are under development.

Background

Satellite remote sensing has often been used to measure snow-covered area (SCA) across vast regions. Most operational SCA products are created from medium to coarse-resolution imagery (i.e. tens to hundreds of meters per pixel) with variable revisit times, offering daily to biweekly imagery of a single location.

These pixels often contain signals from multiple land cover types, resulting in **the mixed pixel problem**, which requires spectral unmixing models to obtain fractional snow-covered area (fSCA).

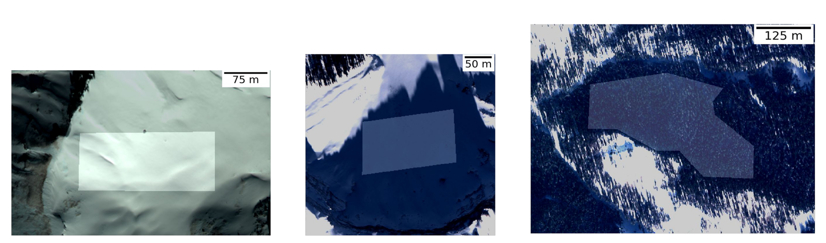
Widely-used fSCA products from MODIS and Landsat 8 (more recently) incorporate these types of routines.

Why snow-covered area?

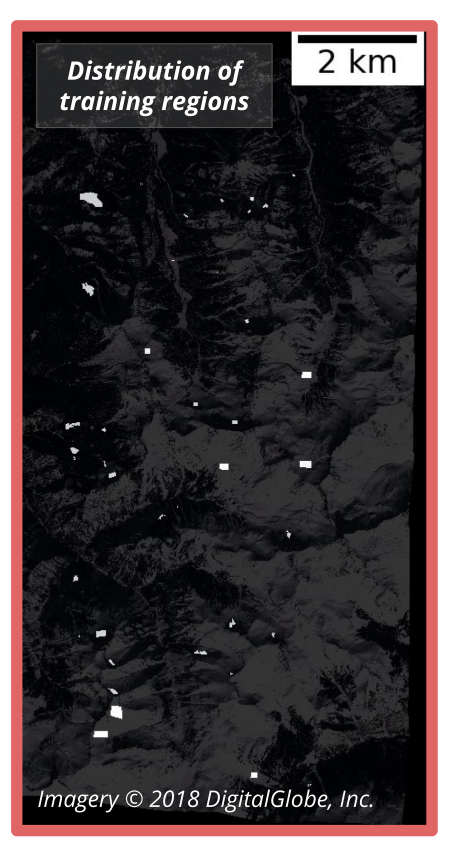
- Impacts **radiation** balance
- **Insulates** ground surfaces & regulates temperature
- Determines when and where **snowmelt** can occur [1]
- One **indicator** for fire season severity [2]

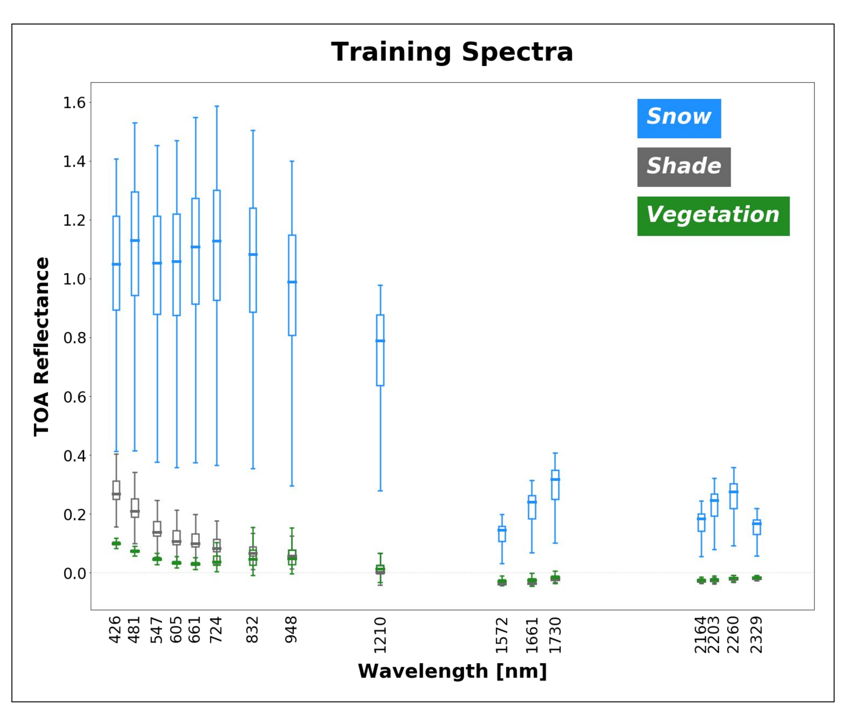
Training dataset

Training regions for land cover classes were created using panchromatic and multispectral WorldView-3 imagery. Regions were distributed throughout the test image with classes between 97,000–140,000 total pixels each. For this generation of models, spectral differences between land cover classes was maximized, resulting in more homogeneous, spectrally distinct classes.



Above: example training regions for each land cover class (snow, snow in shade, and vegetation). Imagery © 2018 DigitalGlobe, Inc.



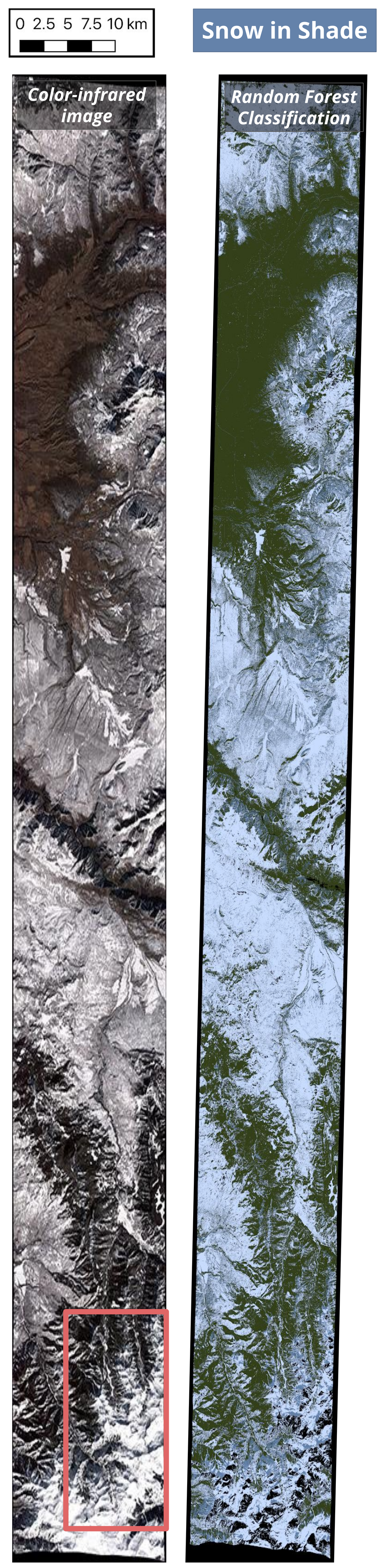


Left: Training spectra for subset of land cover classes. "Shade" denotes snow in shade. Note that all values are top-of-atmosphere reflectance.

Sample WorldView-3 imagery for San Juan Mountains, Colorado

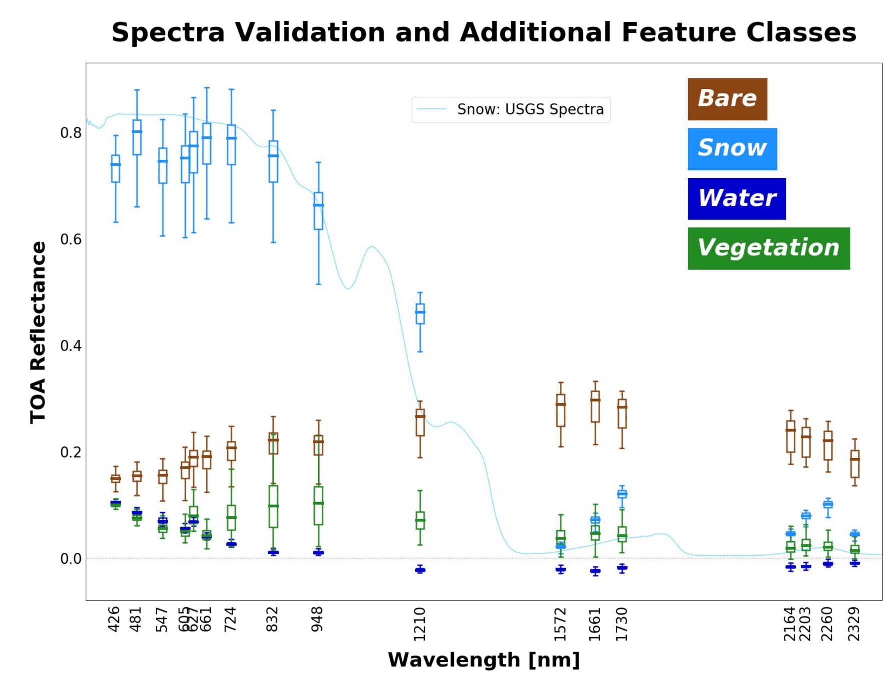
Random forest models quickly classify unseen areas demonstrating both the portability and increased utility of pre-trained models in this production-style workflow. Leveraging high performance computing resources, classification products can be generated **at scale** for large archives of very-high-resolution image data.

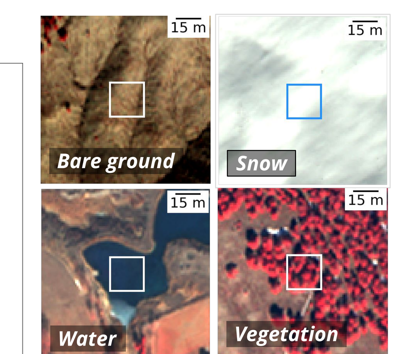
Legend: Snow | Vegetation | Snow in Shade



Above: Color-infrared composite with location of training regions delineated in red. Corresponding random forest classification on the right. Imagery © 2018 DigitalGlobe, Inc.

Strong water and mineral absorption in the **short-wave infrared (SWIR)** helps to **discriminate between clouds and snowy/icy surfaces** and **identify bare ground**. Subsequent model generations will leverage these regions to increase model robustness and classification accuracy.



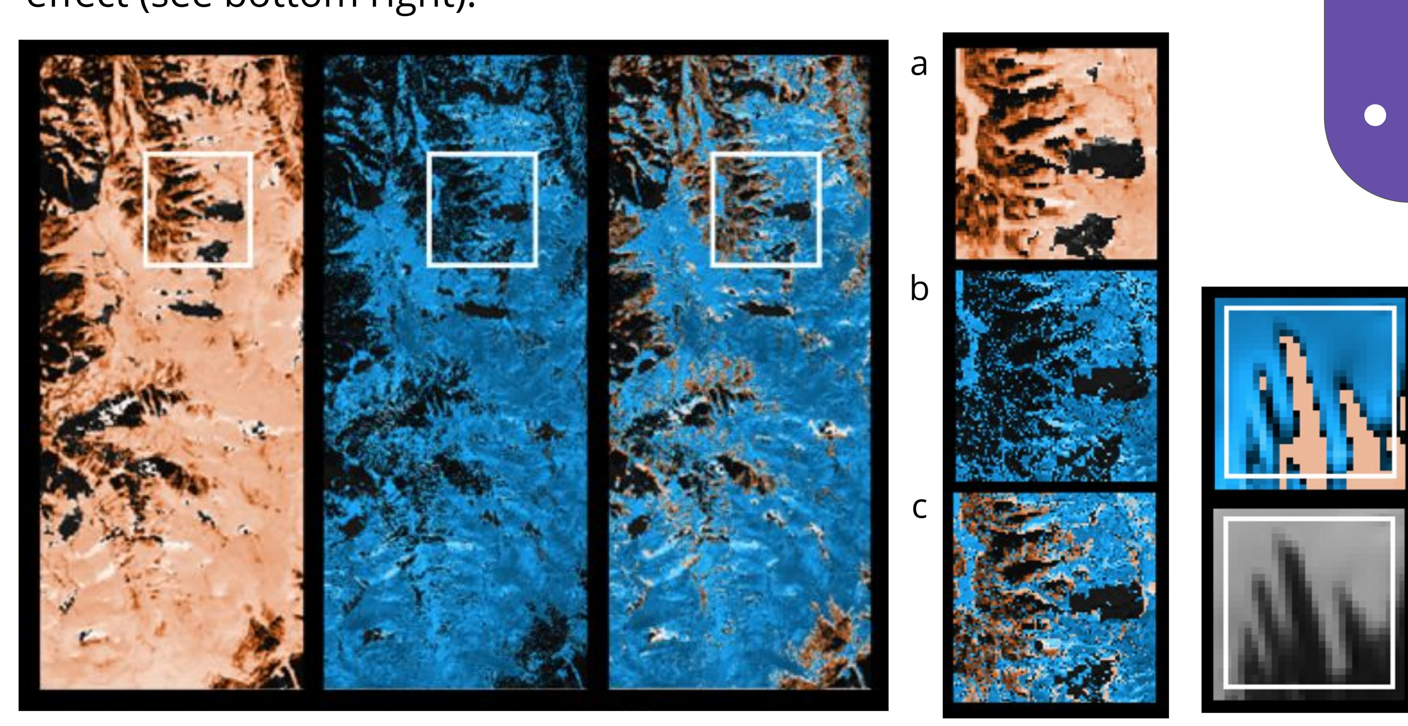


Left and above: WV-3 spectra for additional land cover classes. SWIR bands (1210–2329 nm) helpful in feature identification. Imagery © 2018 DigitalGlobe, Inc.

Meter-Scale Validation

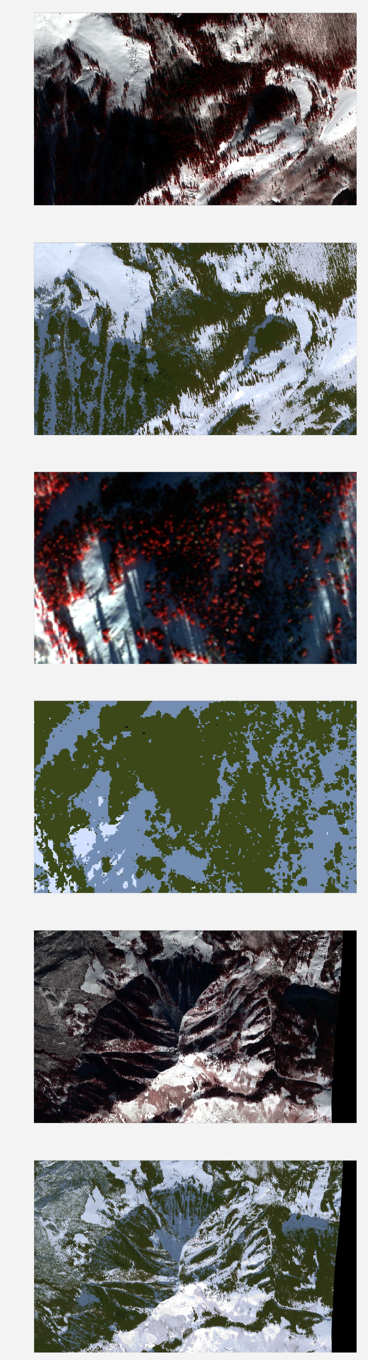
Preliminary comparison work between Landsat 8 and WorldView-3 suggests very-high-resolution imagery can serve as a **validation** for coarse-resolution classification and fSCA products.

Initial results in the test area shown in the bottom panel showed that **Landsat 8 imagery classified 16.3% more snow than WorldView-3** (83.6% vs 67.3% of total area). Most of this disparity was attributable to overestimates from the mixed pixel effect (see bottom right).



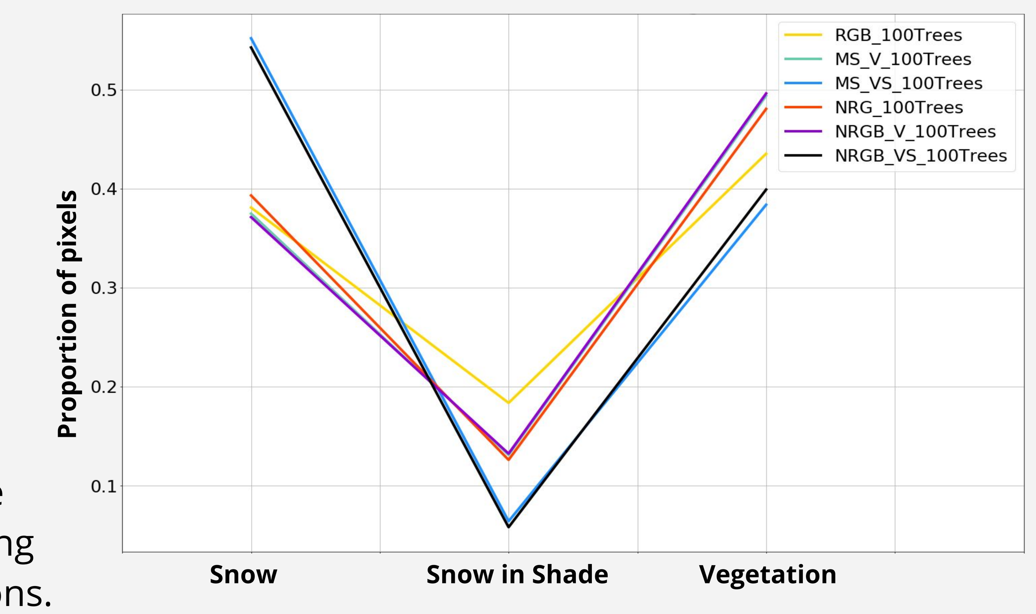
Upper left and middle panels: Pixels classified as snow within the training area for Landsat 8 (left: bronze) and WorldView-3 (middle: blue). White-bordered insets show locations of panels a–c. **Right panel:** Mixed pixel panel with WorldView-3 snow pixels in blue overlaying a single Landsat 8 snow pixel in bronze (top). Corresponding WV-3 image (bottom) reveals snow and forest cover.

Adaptive model selection based on available imagery

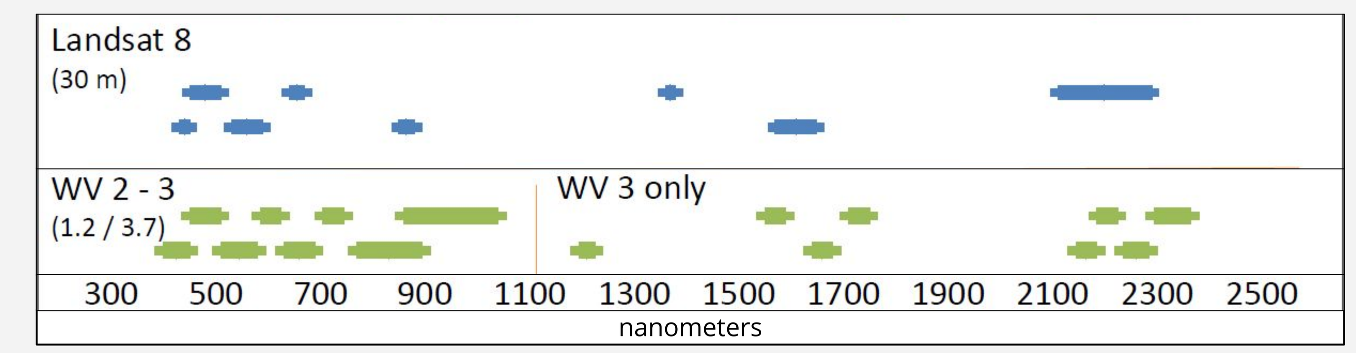


Left: Classifiers resolve shaded snow, individual trees, and the shaded snow between trees. *Imagery © 2018 DigitalGlobe, Inc.*

Right: Models trained on image stacks with varying band combinations.



Models trained with SWIR imagery identify higher proportions of snow.



Above: Spatial and spectral resolution and spectral comparison between Landsat 8 and WorldView-3. *Image © 2019 DigitalGlobe, Inc.*

Takeaways

- **Semi-automated** approach adaptively trains and selects models to reduce manual intervention
- **Very-high-resolution classifications** produced by single models can serve as validation for coarser products
- Workflow **generates classification products at scale**

Next Steps

- Incorporate **additional data layers** (elevation, texture, microwave datasets)
- Build more **robust models** by refining training classes
- Continue development of **neural networks**
- Create **time series** for SCA and land cover change

References

1. D.R. DeWalle and A. Rango. 2008. Principles of snow hydrology. Cambridge., Cambridge University Press. 410pp. ISBN 978-0-521-82362-3, hardback.
2. Medler, M.J., P. Montesano, and D. Robinson. 2002. Examining the relationship between snowfall and wildfire patterns in the western United States. Physical Geography 23: 335–342. doi: 10.2747/0272-3646.23.4.335

Acknowledgments

This work was financially supported by NASA grant #80NSSC18K1405. Imagery provided by DigitalGlobe, a MAXAR company. Non-financial support provided by friends, loved ones, and the rest of the members of the Terrain Analysis and Cryosphere Observation Lab: Shashank Bhushan, Friedrich Knuth, and Joe Meyer. Thanks also to Karl Rittger for feedback during preliminary discussions.