USING REMOTE SENSING DATA FUSION MODELING TO TRACK SEASONAL SNOW COVER IN A MOUNTAIN WATERSHED

by

Allison N. Vincent

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Hydrologic Sciences

Boise State University

May 2021

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Allison N. Vincent

Thesis Title: Using Remote Sensing Data Fusion Modeling to Track Seasonal Snow-Cover in a Mountain Watershed

Date of Final Oral Examination: 05 March 2021

The following individuals read and discussed the thesis submitted by student Allison N. Vincent, and they evaluated their presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Alejandro N. Flores, Ph.D.	Chair, Supervisory Committee
Hans-Peter Marshall, Ph.D.	Member, Supervisory Committee
Nancy F. Glenn, Ph.D.	Member, Supervisory Committee
Caroline S. Nash, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Alejandro N. Flores, Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

ACKNOWLEDGMENTS

There are many people whom I would like to sincerely thank for their efforts and involvement in this project. First, I would like to thank my advisor, Dr. Alejandro "Lejo" Flores for giving me the opportunity to study under his direction and fulfill my dream of attaining my master's degree. I am so grateful for everything I have learned under your patience and guidance. I would also like to acknowledge my thesis committee members: Dr. Hans-Peter Marshall for contributing his expertise in snow science to this project and providing opportunities for me to attain hands-on experience in the field with the NASA SnowEx campaigns, Dr. Nancy Glenn for contributing her vast expertise in remote sensing to this project, and Dr. Caroline Nash. A special thank you goes out to Dr. Nash for the many hours she spent guiding me through every stage of my research and meticulously reviewing my work, all of which helped immensely in my professional development as a scientist. To Lejo and my entire committee, it has truly been a pleasure working with all of you.

Thank you to the members of the Lab for Ecohydrological Applications and Forecasting (LEAF) for providing such a supportive environment in which to grow, both personally and professionally, during my time as a graduate student. I would additionally like to thank those individuals who provided technical and computational assistance: Jake Graham and Josh Enterkine for helping me with coding, James Nelson for his wealth of patience with me during my many technological struggles, and the Boise State Research Computing Department, especially Mike Ramshaw, Jason Watt, and Forrest Burt, for

providing me with timely assistance and somehow always magically fixing all of my issues with R2.

Finally, I wish to thank several individuals from my undergraduate days at Iowa State University: Dr. Richard Schultz, Dr. William Simpkins, Dr. Thomas Isenhart, and Leigh Ann Long. Thank you all for your unwavering support as I stumbled through the process of finding my career path. The opportunities you provided for my initial research experiences helped me realize my passions and allowed me to end up where I am today.

This research utilized the high-performance computing support of the R2 compute cluster (DOI: 10.18122/B2S41H) provided by Boise State University's Research Computing Department. Funding for this project came from the Department of Energy Office of Science award DE-SC0019222 through the Subsurface Biogeochemical Research Program.

ABSTRACT

Seasonal snowfall is the largest component of the water budget in many mountain headwater regions around the world. In addition to sustaining biological water needs in drier, lower elevation areas throughout the year, mountain snowpack also provides essential water inputs to the Critical Zone (CZ) - the outer layer of the Earth's surface, which hosts a variety of biogeochemical processes responsible for transforming inorganic matter into forms usable for life. Water is a known driver of CZ activity, but uncertainty exists in its spatial and temporal interactions with CZ processes, particularly in the complex terrain of heterogeneous mountain areas. Increasing pressure on the CZ due to climate change and human land use needs creates an urgency to better understand the CZ system and how it may change in the future. An important variable for water driven CZ behaviors in mountain areas is the spatial extent of snow, also known as snow-covered area (SCA). SCA in mountain areas can change quickly over small scales of time and space with large impacts on the rest of the system. It has been difficult historically, however, to measure snowpack extent for large areas on very fine spatial and temporal scales due to a lack of remote sensing datasets with both of these fine scale characteristics. In this study we use the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) to fill this historic knowledge gap for the East River watershed in Colorado, USA. By fusing low spatial and high temporal resolution data from MODIS (500-m, daily) with high spatial and low temporal resolution data from Landsat (30-m, 16 days), a fine resolution, 30-m daily dataset can be created. This study is one of the first to

use this model with the primary intent of monitoring SCA in a mountain watershed.

The first component of the study in this thesis presents a comprehensive validation of STARFM for use in monitoring snow cover in mountain areas. Normalized Difference Snow Index (NDSI) values from MODIS and Landsat are used as input to the STARFM model, and synthetic NDSI values at 30-m resolutions are obtained for days without Landsat data acquisitions. After converting NDSI to binary snow cover, we then examine the temporal performance of STARFM for an entire calendar year. The model's performance is also analyzed for different landscape features known to influence snow cover. Accuracy, precision, recall, and F-score values indicate that the model is able to successfully predict the location of SCA in the landscape when validated with Landsat data.

The second component of the study describes the process of creating the daily, 30-m NDSI dataset with STARFM for 20 water years of analysis and provides examples of how these data can be used to monitor SCA in a mountain watershed. We examine patterns of percent annual snow cover for three of the water years from the dataset, a dry, average, and wet water year. Here we find that predictable patterns of SCA occur over those years, with the highest percent annual snow cover occurring during the wet year and the lowest occurring during the dry year. Despite these differences, however, elevation is clearly the dominating factor in determining the spatial variability of snow cover in the landscape for all three water years. We also connect our snow cover analysis back to CZ processes by examining the timing of snow cover disappearance with the peak of annual stream discharge at the watershed outlet.

The results of this work provide a multi-decadal dataset of snow cover information for the East River that can be used for future research into snowpack and streamflow forecasting, modeling of the movement of water through the CZ, and the effects that climate change may have on these processes. This study also provides examples of methods that can be used for further snow monitoring work in the East River watershed and other snow-dominated mountain catchments similar to it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AOI	Area of Interest
ASO	Airborne Snow Observatory
CZ	Critical Zone
DEM	Digital Elevation Model
DOE	Department of Energy
ETM+	Enhanced Thematic Mapper Plus (Landsat)
*f*SCA	Fractional Snow-Covered Area
GEE	Google Earth Engine
Landsat	Land Remote-Sensing Satellite System
Lidar	Light Detection and Ranging
OLI	Operational Land Imager (Landsat)
MODIS	Moderate Resolution Imaging Spectroradiometer
MODSCAG	MODIS Snow-Covered Area and Grain Size
NASA	National Aeronautics and Space Administration
NDSI	Normalized Difference Snow Index
SBR	Subsurface Biogeochemical Research
SCA	Snow-Covered Area
SFA	Scientific Focus Area
SNOTEL	Snowpack Telemetry
STARFM	Spatial and Temporal Adaptive Reflectance Fusion Model

SRTM	Shuttle Radar Topography Mission
SWE	Snow Water Equivalent
TM	Thematic Mapper (Landsat)
TMSCAG	Thematic Mapper Snow-Covered Area and Grain Size
USGS	United States Geological Survey
WY	Water Year

CHAPTER ONE: INTRODUCTION AND OVERVIEW

1.1 The Critical Zone

The Critical Zone (CZ) is the dynamic region of the Earth's near-surface environment where a wide range of hydrobiogeochemical processes transform inorganic ingredients into terrestrial life (NRC, 2001). The boundaries of the CZ (Figure 1.1) are broadly defined as the vertical region extending from the vegetation canopy to the base of the saturated groundwater zone (NRC, 2001). The exact locations of the CZ boundaries, however, vary across the literature depending on the process being studied and the scientific field of interest (Anderson et al., 2007; Lin, 2010; Akob & Küsel, 2011; Kim et al., 2017; Riebe et al., 2017). As is often the case in rapidly emerging fields of inquiry, new insights into CZ processes necessitate the evolution of the fundamental definitions of the field and the recognized boundaries of the system. For example, the lower boundary of the CZ is often defined as where life no longer influences rock, but microbiological life is regularly found at depths much lower than it was previously thought to exist (Akob & Küsel, 2011).

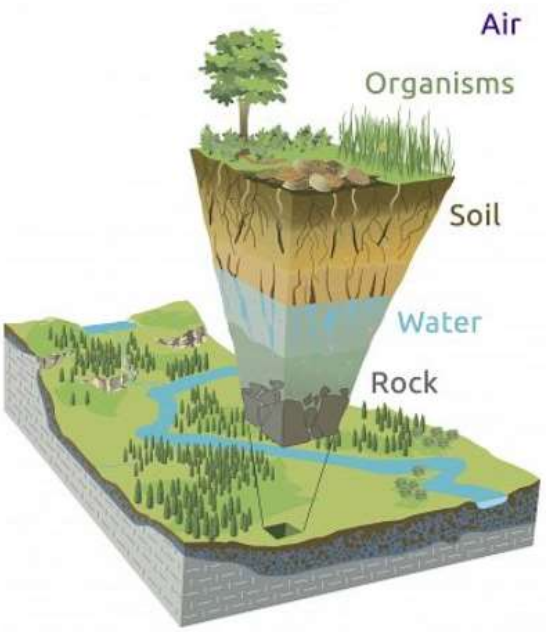


Figure 1.1: A conceptual model of the CZ with boundaries defined as the vertical region extending from the vegetation canopy to the base of the saturated groundwater zone. Source: czo-archive.criticalzone.org

The CZ is "critical" because it is responsible for many first-order controls on biological life. The bedrock base of the CZ contains a host of elements that are gradually released into the environment as the structure of the subsurface changes and evolves through chemical and physical weathering (Anderson et al., 2007). The release of these elements and the resulting structural changes, such as bedrock fracturing, directly control the distribution of vegetative cover at the surface (Hahm et al., 2014). Vegetative communities in turn have an important role in soil development, adding carbon through organic matter input and building a carbon storage reservoir that is significant to the global carbon cycle (Eswaran et al., 1993). Evidence suggests that, after the ocean, the CZ is the largest reservoir of carbon on the planet (Lal, 2004). Soils in particular contain four times more carbon than is found in living biomass and three times more carbon than is found in the atmosphere (Lal, 2004). Just 10% of this organic carbon is equivalent to

all anthropogenic CO_2 emissions over a 30-year period (Kirschbaum, 2000), underscoring just how invaluable the CZ is not just from a local, but also global, perspective.

Water is a key driver for the many biogeochemical reactions within the CZ, determining their rate, magnitude and oftentimes direction. Because of this, the hydrologic cycle is a useful organizing principle by which to gain insight into the myriad, complex processes at work within the CZ. For example, the presence of water, along with the size of pore spaces in the soil and bedrock substrate, determines the species of microorganisms found in the CZ, which in turn make up the foundations of the food and energy webs of their ecosystems (Akob & Küsel, 2011). Rock moisture, which is water from precipitation stored in shallow weathered bedrock, plays an important role in sustaining vegetation through times of drought (Rempe & Dietrich, 2018). Fluxes of water from the surface also bring solutions such as dissolved organic matter and gases like CO_2 and $O_2$to the subsurface (Anderson et al., 2007). These solutes and the amount and frequency of flow largely determine the degree and types of weathering that take place in the subsurface of the CZ (Lin, 2010). Weathering processes affecting the CZ locally make carbon and other key nutrients available not only to the surrounding environment, but also at a larger spatial scale as can be detected by the chemical signature in the groundwater exiting the watershed (Kim et al., 2017). As such, the timing and magnitude of water inputs to the CZ drive key processes fundamental to the CZ and the services it provides to living organisms.

In recent decades, the urgency to better understand the structure of and constrain the processes taking place in the CZ has been widely recognized due to the expanding resource needs of human society (NRC, 2001), made more uncertain by the projected

impact of climate and land use change on CZ functions (Banwart et al., 2013). Climate and land use change have the potential to significantly alter the timing and magnitude of water inputs to the CZ, which will, in turn, alter the biogeochemical CZ processes necessary for human and ecosystem health. For example, as global temperatures rise, the rates of uptake and release of carbon between the CZ and the atmosphere are expected to change non-uniformly around the world. Warmer temperatures are expected to speed up weathering rates in some areas but increase net primary production in others (Kirschbaum, 2000). This non-uniformity is difficult to constrain in part due to uncertainty around how climate change will alter soil moisture regimes, particularly through changing precipitation patterns (Kirshbaum, 2000; Falloon et al., 2011), the primary mechanism of water delivery to the CZ. Additionally, potential changes to CZ functions controlled by primary physical processes, such as climate, are expected to be exacerbated by the additional pressures of human land use. Land use practices such as deforestation and intensive agriculture without complementary conservation practices can rapidly deplete CZ resources, including removal of nutrients and vegetation substrate through erosion, reduced water holding capacity of the landscape, and degraded surface water quality (Lal, 2004). These pressures threatening water delivery mechanisms, and in turn the life-sustaining water driven CZ functions, highlight the importance of quantifying water driven behaviors of the current state of the CZ system.

This potential for CZ areas around the globe to respond non-uniformly to climate and land use pressures is largely due to heterogeneity within and between controls on CZ depth and reactivity throughout the globe (Brantley et al., 2007) driven by dynamic and uneven energy inputs into this open system (Lin, 2010). Despite growing

acknowledgment that water is the primary driver of biogeochemical reactions in the CZ, there remain gaps in our comprehension of hydrological processes within the CZ and how they vary across time and space. A better understanding of the heterogeneity of water inputs to the CZ and their influence on key CZ processes is a crucial piece to improving our understanding of the processes critical to life on Earth, as well as how the CZ will react to anthropogenically driven changes, and with what global consequences.

1.2 Mountain Critical Zones and the Importance of Snow

With the many unknowns regarding the influence of the hydrologic cycle on the CZ, these knowledge gaps are particularly relevant to CZs in the mountain headwaters of the world. These headwater areas serve as primary sources of water and solute exports to downstream ecosystems (Kim et al., 2017; Winnick et al., 2017), but the mechanisms by which they do so, particularly those attributed to the CZ, remain largely uncharacterized. Constraining CZ processes is especially problematic in these rugged regions whose nature lends itself to considerable variations in the landscape. Steep gradients in topography (St. Clair et al., 2015) and lithology (Hahm et al., 2014) drive similarly steep gradients in precipitation type and quantity (Marks et al., 2013), which together create heterogeneity in infiltration rates and subsurface water distribution (Sprenger et al., 2019), and lateral transfer of materials (Anderson et al., 2007). These variations in timing, quantity, and rate of water input drive differing patterns of weathering and subsequent variation in vegetation patterns (Roering et al., 2010; Carroll et al., 2018). These combined heterogeneities affect the properties and development of the CZ and, in turn, determine the timing, quantity and rate of transport of water and other key nutrients to downgradient systems.

Because water is so important to CZ mechanisms and their exports, the means of how and when a region receives precipitation is key to understanding CZ development and function. In many mountainous areas, especially those in arid and semi-arid climates such as the western United States, most of the yearly precipitation arrives in the winter as snow (Tennant et al., 2015). As temperatures warm and sun angles increase throughout the spring, variations in topography create gradients in temperature and radiation that melt existing snow and change the dominant precipitation phase from snow to rain (Klos et al., 2014) at increasingly higher elevations. The form of precipitation is important because precipitation that falls as snow can persist in the landscape for longer timescales, gradually releasing water into the rest of the system as it melts (Knowles et al., 2015; Carroll et al., 2018). This gradual melting provides water to the surrounding ecosystem and downstream areas throughout the dry summer months, including to reservoirs for human consumptive uses (Barnett et al., 2005).

Climate change is predicted to alter the patterns of snow precipitation, snow accumulation, and timing of snowmelt in these mountain systems. The past century has already seen a decrease in peak SWE on 1 April, the date which is traditionally considered to have the maximum amount of snow, in mountain areas throughout the western U.S. This is largely attributed to both earlier warming of the snowpack and less overall precipitation, especially as snow, in winter months (Mote et al., 2005). Global temperatures are expected to increase anywhere from 0.3-4.8°C during the coming century (IPCC, 2014). These changes are predicted to significantly affect the timing of seasonal, topographically driven shifts in temperature gradients, melting snow and

moving the rain/snow transition line higher in elevation and latitude at earlier times in the season (Klos et al., 2014; Tennant et al., 2015).

Compounding the effects due to likely changes in mountain snowpack from climate change are the effects of changes in land use, both locally and regionally, that result in disturbance of natural vegetative cover. When soil is left exposed, as sometimes happens with agricultural activities or biomass burning, increased aeolian transport of sediment can occur (Vicars & Sickman, 2011). If aeolian transported sediment is carried to a snow-covered area, it can cause dust on snow events that lower the albedo of the snowpack and contribute to increased melt rates (Skiles & Painter, 2017). It is reasonable to expect that an increasing human population will continue to drive demand for agricultural and commercial development, thereby expanding the amount of area disturbed by removed vegetative cover (Smith et al., 2014), and further influencing changes to the duration of seasonal mountain snowpack by increased aeolian deposition on snow.

With the changes anticipated to occur regarding mountain snow regimes, it is plausible to assume that rates and timing of CZ processes may also be altered in these areas. How exactly these changes will likely manifest, however, remains uncharacterized. Constraining the controls on snow as a unique and changing moisture input to the CZ in mountainous headwaters will allow for identification of the integrated water/CZ behavior responses that are essential to mountain CZ formation and function. Identifying these controls and their influences on mountain CZs begins with a fundamental understanding of how snow is distributed in these landscapes. As previously discussed, however, the compounding heterogeneity of CZs and mountain areas renders many traditional forms of

data acquisition unsuitable, requiring innovative thinking as to how we collect spatial snow data.

1.3 Remote Sensing of Snow

Remote sensing is a viable method for collecting data on landscape changes over wide spatial and temporal scales, especially in environments that are difficult to access. Gathering any data, but especially snow data, in situ in mountain areas is resource intensive and sometimes potentially hazardous. Some infrastructure does exist for the purposes of local snow observation and data collection. A primary example is the Snowpack Telemetry (SNOTEL) station, which collects observations of snow, weather, and other climatological data, and of which there are multiple stations located throughout mountain watersheds of the western U.S. (wcc.nrcs.usda.gov/snow/). While SNOTEL data are extremely valuable, they only offer information for one spatial point, which is often located in flat, easily accessible areas that may not be characteristic of the entire region they are meant to represent. Snow monitoring at basin or regional scales requires not only observations at the required intervals (temporal resolution), but also at a small enough scale where processes of interest can be observed (spatial resolution). Fortunately, a wide variety of remote sensing methods currently exist for measuring the various properties of snow, including snow depth, snow albedo and grain size, snow water equivalent (SWE), and snow covered area (SCA) over vast regions at a variety of spatial and temporal resolutions (Nolin, 2010).

Ground-based and airborne remote sensing methods typically provide the highest spatial resolution snow data. Light Detection and Ranging (lidar), both through airborne methods and ground-based Terrestrial Laser Scanning (TLS), has been used to monitor

snow depth and determine SWE at spatial resolutions of tens of meters or less (Deems et al., 2013). Imaging spectrometers are often used in conjunction with lidar airborne operations and provide snow albedo measurements (Painter et al., 2016). Temporal resolution of these methods is limited, however, by the frequency of both the ability to fly over the designated area and ground access of the site in the case of TLS. Another airborne method of snow monitoring is Structure from Motion (SfM) photogrammetry with airplane flights or Unmanned Aerial Vehicles (UAVs). This technology can generate highly accurate digital surface models (DSMs) with centimeter resolution that can be used to determine snow depth, SCA, and SWE (Buhler et al., 2016; Fernandes et al., 2018). Similar to lidar, however, SfM acquisitions are currently limited by flight availability, weather conditions, or in the case of UAVs, by aerial coverage due to the limitations in signal distance.

Satellite remote sensing solves problems of temporal frequency and areal coverage in that it can monitor large areas at routine intervals. To this effect, several satellites with sensors capturing data at varying wavelengths throughout the electromagnetic spectrum are currently in operation collecting snow data. Passive microwave remote sensing is possible with several different instruments, including the Special Sensor Microwave/Imager (SSM/I) and the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E), both operated by the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Association (NOAA). Synthetic aperture radar (SAR), or active microwave, data is currently being collected on instruments such as the European Space Agency (ESA)’s Sentinel-1 satellite and the Canadian Space Agency (CSA)’s Radarsat-2 satellite.

These methods are useful because they operate at wavelengths that can penetrate through clouds, therefore eliminating many data collection restrictions due to weather conditions. Both methods have primarily been used to determine snow depth and SWE with success (Bernier et al., 1999; Che et al., 2008; Daly et al., 2012; Vuyovich et al., 2014; Lievens et al., 2019). As a by-product of measuring snow depth and SWE, SCA can also be determined. Some passive microwave sensors, however, have been found to have difficulties in determining the spatial distribution of snow (Vuyovich et al., 2014). The spatial resolution of passive microwave data is also more coarse than other instruments (several kilometers to tens of kilometers), and therefore is not always suitable for studying snow processes at sub-basin scales. SAR imagery, on the other hand, can produce products with very fine spatial resolutions (as low as 1-m) at 6-day revisit cycles. SAR methods can encounter difficulties in measuring SCA, however, due to the differences in signal backscatter from dry and wet snow and from exposed rocks and vegetation during partial snow melt, leading to underestimations in snow depth and snow cover (Storvold et al., 2006; Lievens et al., 2019).

Satellite sensors operating in the visible and infrared spectrums, also known as optical sensors, can also be used for monitoring snow properties. Some examples include NASA's Advanced Very High-Resolution Radiometer (AVHRR) and Visible Infrared Imaging Radiometer Suite (VIIRS). These and other optical data products have been used to successfully monitor the spatial distribution of snow (Zhou et al., 2013; Zhang et al., 2020). Optical data are useful for this task due to snow's high reflectivity and the strong spectral reflectance contrast in the visible and infrared spectrums between snow-covered and snow-free areas, and can also be used to measure properties such as snow grain size

and albedo (Painter et al., 2009). Some optical remote sensing instruments provide data captured at high spatial resolutions, which can be used to identify snow cover patterns in highly heterogeneous landscapes. Others provide data at high temporal resolutions, which can be used to identify quickly changing snow patterns. Two optical remote sensing instruments that fit these descriptions and are of interest to this study are NASA's Moderate resolution Imaging Spectroradiometer (MODIS) and NASA's Land Remote-Sensing Satellite System (Landsat).

The MODIS instrument onboard the Terra and Aqua satellites, whose missions began in 1999 and 2000, respectively, collects data in 36 spectral bands covering the visible and infrared wavelengths at various resolutions of 250-m, 500-m, and 1000-m, and has a daily revisit time for most areas around the globe. Due to its high temporal resolution, MODIS data are often used to observe rapid changes in landscapes. As such, MODIS data have been used in conjunction with, or as validation for, other remote sensing datasets used in snow studies. For example, in a snow depth study conducted by Che et al. (2008), optical remote sensing data from MODIS was used as a validation for passive microwave estimates of SCA due to it having finer spatial resolution than the passive microwave data. MODIS data have also been used to create several SCA remote sensing products. One example is the widely used MOD10A1 (Hall et al., 2016) product, whose current version also contains snow albedo and fractional snow-covered area (*f*SCA) data. Another is the MODIS Snow-Covered Area and Grain Size (MODSCAG) product (Painter et al., 2009), which contains *f*SCA and grain size data obtained by using MODIS surface reflectance data with a multiple endmember spectral mixture analysis (Roberts et al., 1998).

NASA's Landsat satellite mission began with Landsat 1, launched in 1972, and continues with the current satellite, Landsat 8, with Landsat 9 expected to be launched in September 2021 (usgs.gov/core-science-systems/nli/landsat/landsat-9). The Landsat missions have captured data in the visible and infrared spectral bands at progressively higher spatial resolutions, with current data available in 30-60-m resolution. These high spatial resolutions allow for identification of heterogeneities in a landscape, including distribution of snow cover at the sub-basin scale (Vikhamar & Solberg, 2002; Salomonson & Appel, 2004) that is not always possible with data of coarser spatial resolution. Similar to MODIS data, SCA data products derived from Landsat are available, which include a Landsat Fractional Snow-Covered Area product (Selkowitz et al., 2017). Landsat data have also been used as a method of validating snow models and as "ground-truth" data for in the development of other remote sensing snow products. Landsat Thematic Mapper (TM) data were used as model validation in the development of the MODSCAG model and subsequent snow product. Salomonson & Appel (2004) also used Landsat data for validation in the development of the SNOMAP algorithm, a method of using the normalized difference snow index (NDSI) to find sub-pixel *f*SCA from MODIS surface reflectance data.

As useful as Landsat data is because of its high spatial resolution, its temporal resolution of 16 days limits its ability to observe rapid changes occurring in a landscape. This limitation is particularly disadvantageous for observing snowpack in mountain regions, where large changes in SCA can occur from day-to-day during the snow accumulation and snowmelt seasons. As such, developing an understanding of where water from snowpack enters the subsurface in a highly heterogeneous region requires

detailed characterization of the locations of early snow accumulation as well as areas of persistent snow cover as melting is occurring, a demand which cannot be met by Landsat data alone.

This trade-off between remote sensing datasets with high spatial or temporal resolution, but not both, has been a common problem with remote sensing datasets until very recently. The Sentinel-2 mission, launched in June 2015 by the ESA, provides a bottom-of-atmosphere (BOA) corrected reflectance data product at 10, 20, or 60-m spatial resolution and a temporal resolution of 5 days. Although this is a vast improvement in data resolution, any analysis of remote sensing data prior to this high-resolution data availability must deal with spatial and temporal resolution limitations. These limitations have often been dealt with by downscaling high temporal, low spatial resolution data, such as MODIS data (Emelyanova et al., 2012; Gevaert & García-Haro, 2014; Walters et al., 2014), to spatial resolutions at which fine landscape features can be identified. In this study we explore one such method of data downscaling using the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) and evaluate its potential to be used to monitor daily, fine-spatial resolution changes in SCA.

1.4 East River Watershed Scientific Focus Area

The location for this study is the East River watershed, located in the Upper Colorado River Region, USA and encompassing the town of Crested Butte, CO (Figure 1.2). Contained within it is a study watershed that supports scientific investigation through the Watershed Function Scientific Focus Area (SFA) supported by the Department of Energy (DOE)'s Subsurface Biogeochemical Research Program (SBR). The Watershed Function SFA is led by the Lawrence Berkeley National Laboratory

(LBNL) (watershed.lbl.gov), and uses the East River watershed as one of many testbeds in its network for investigating its grand challenge compound question of "How do mountainous watersheds retain and release water, nutrients, C and metals over episodic to decadal perturbations, and what are the ramifications for downgradient water availability and quality?" (Hubbard et al., 2018). Our study addresses two of the six supporting scientific questions being investigated specifically at the East River watershed (Hubbard et al., 2018):

1. How do perturbations to individual watershed subsystems, including early snowmelt and drought, lead to downgradient exports of water, C, N, and P from that subsystem?
2. Which insights and methods are critical for improving operational forecasting predictions of water quantity and quality in response to a range of pulse and press perturbations?

The East River watershed drains approximately 750 km², has a mean elevation of 3137-m, and a topographic relief of 1869-m. The area listed here is slightly larger than the boundaries of the East River SFA (Hubbard et al., 2018). This is because we include the entire drainage area of the East River as calculated from its confluence with the Taylor River at the south end of the watershed. The large topographic relief enables the catchment to sustain alpine, subalpine, montane, and riparian ecosystems, with most of the area covered by conifer and aspen forests (Carroll et al., 2018). The region has a continental climate with long, cold winters and short, cool summers. Two SNOTEL (wcc.nrcs.usda.gov/snow/) sites are included within the watershed area. The Schofield station (site number 737) and the Butte station (site number 380) are located at elevations

of 3,261-m and 3,097-m, respectively. For the almost 40-year period of record, mean annual temperature has been recorded as ~0.5°C at the Schofield station and ~2.4°C at the Butte station. Mean annual precipitation is 1222 mm at Schofield and 663 mm at Butte, the majority of which falls during the winter months as snow in both locations (Carroll et al., 2018). The geology of the SFA has been mapped and consists mainly of Paleozoic and Mesozoic sedimentary layers, including the large Mancos Shale layer, with multiple Cenozoic igneous intrusions (Gaskill et al., 1991).

Because this area is mid-continental and high elevation, it is expected that the timing and amount of winter snowpack will be less affected under a warming climate than other mountain regions. Climate modeling completed by Klos et al. (2014) showed that although the Colorado Headwaters, and specifically the Gunnison River watershed

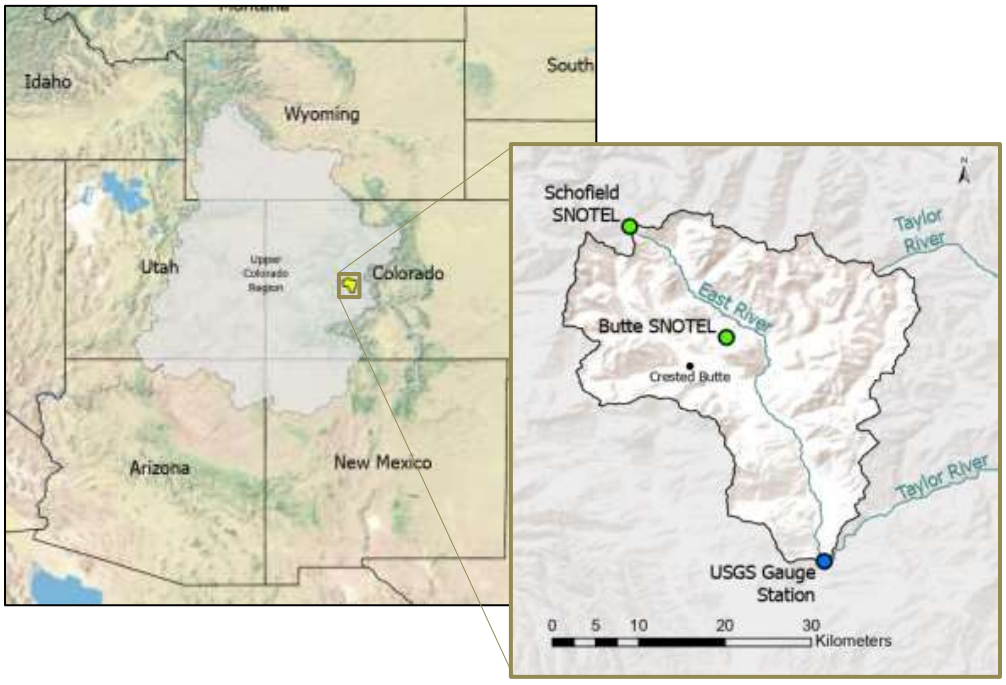


Figure 1.2: The East River watershed, located in western Colorado, USA in the headwaters of the Upper Colorado River Region. The Butte and Schofield SNOTEL sites located within the watershed are identified in green. The East River at Almont USGS gauging station is identified in blue.

(which encompasses the East River watershed), are predicted to experience a decrease in snow-dominated area by 30% and 26%, respectively, by the mid-21st century, these areas are expected to retain some of the largest snow-dominated area compared to other major watersheds in the western U.S. Although these predictions appear promising for the future of the Upper Colorado River Region, climate change is predicted to make regions that depend on the Colorado River as a major source of water, particularly the southwestern U.S. and northwestern Mexico, drier and hotter, thereby increasing the demand for water from the river (Garfin et al., 2013).

1.5 Thesis Organization

This study addresses the issue of the lack of historical fine spatial and temporal resolution remote sensing data needed to fully understand the distribution of snow cover in mountain watersheds. Snow cover patterns and the resulting rate and timing of water delivery to the CZ in mountain areas is a significant knowledge gap in the understanding of CZ processes. Using STARFM, our goal is to explore the use of a method that can realize the goal of observing small scale, day-to-day changes in snow cover in a mountain landscape. This is accomplished within the broader context of understanding how these changes in snow cover affect water distribution and delivery to the Critical Zone. Specific outcomes of this thesis will be discussed in the following chapters.

Chapter 2 provides a detailed assessment of the ability of the STARFM model to correctly predict SCA over time and space by creating synthetic snow cover data for dates when no high spatial resolution data were collected via satellite remote sensing. A sensitivity analysis is first completed to tune STARFM's parameters to the East River study area, a mid-latitude mountain watershed area in the Upper Colorado River Region,

USA. The STARFM model is then validated by running it for an entire calendar year and using a data denial method to exclude Landsat images from input and then compare them to model output from that date. With these results, model performance can be evaluated temporally across the year and spatially by topographic and vegetative features known to influence snow cover. Additionally, data from an Airborne Snow Observatory (ASO) overflight during the year of analysis is used for further model validation.

Chapter 3 describes the creation of the final multi-sensor, high spatiotemporal resolution snow cover dataset and provides examples of analyses of landscape-level controls on snow cover dynamics that can be accomplished with these data. We achieve this through an in-depth look into the snow cover data of three separate water years, one average water year, one dry water year, and one wet water year as determined by peak annual SWE values over an almost 20-year period. Finally, we provide a use case for relating snow cover trends observed with STARFM back to water delivery to the CZ by examining snow covered area trends with stream discharge behavior.

References

Akob, D. M. & Küsel, K. (2011). Where microorganisms meet rocks in the Earth's Critical Zone. *Biogeosciences, 8,* 3531-3543. doi:10.5194/bg-8-3531-2011

Anderson, S. P., von Blanckenburg, F., & White, A. F. (2007). Physical and Chemical Controls on the Critical Zone. *Elements, 3*(5), 315–319.

Banwart, S. A., Chorover, J., Gaillardet, J., Sparks, D., White, T., Anderson, S., Aufdenkampe, A., Bernasconi, S., Brantley, S. L., Chadwick, O., Dietrich, W. E., Duffy, C., Goldhaber, M., & Ragnarsdottir, K. V. (2013). Sustaining Earth's Critical Zone Basic Science and Interdisciplinary Solutions for Global Challenges. *The University of Sheffield*.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature, 438,* 303-309. doi:10.1038/nature04141

Bernier, M., Fortin, J. P., Gauthier, Y., Gauthier, R., Roy, R., & Vincent, P. (1999). Determination of Snow Water Equivalent using RADARSAT SAR data in eastern Canada. *Hydrological Processes, 13,* 3041-3051. doi:10.1002/(SICI)1099-1085(19991230)13:18<3041::AID-HYP14>3.0.CO;2-E

Brantley, S. L., Goldhaber, M. B., & Ragnarsdottir, K. V. (2007). Crossing Disciplines and Scales to Understand the Critical Zone. *Elements, 3*(5), 307-314.

Buhler, Y., Adams, M. S., Bosch, R., & Stoffel, A. (2016). Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations. *The Cryosphere, 10,* 1075-1088. doi:10.5194/tc-10-1075-2016

Carroll, R. W. H., Bearup, L. A., Brown, W., Dong, W., Bill, M., & Williams, K. H. (2018). "Factors controlling seasonal groundwater and solute flux from snow-dominated basins". *Hydrological Processes, 32*, 2187-2202. doi:10.1002/hyp.13151

Che, T., Li, X., Jin, R., Armstrong, R., & Zhang, F. (2008). Snow depth derived from passive microwave remote-sensing data in China. *Annals of Glaciology 29,* 145-154. doi:10.3189/172756408787814690

Daly, S. F., Vuyovich, C. M., Deeb, E. J., Newman, S. D., Baldwin, T. B., & Gagnon, J. J. (2012). Assessment of the snow conditions in the major watersheds of Afghanistan using multispectral and passive microwave remote sensing. *Hydrological Presses, 26,* 2631-2642. doi:10.1002/hyp.9367

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479. doi:10.3189/2013JoG12J154

Emelyanova, I. V., McVicar, T. R., Van Niel, T. G., Li, L. T., & van Dijk, A. I. J. M. (2012). On blending Landsat-MODIS surface reflectances in two landscapes with contrasting spectral, spatial and temporal dynamics. WIRADA Project 3.4: Technical Report. CSIRO Water for a Healthy Country Flagship, Australia. 72 pp.

Eswaran, H., Van Den Berg, E., & Reich, P. (1993). Division S-5 Notes. *Soil Sci. Soc. Am. J, 57,* 192–194.

Falloon, P., Jones, C. D., Ades, M., & Paul, K. (2011). Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty. *Global Biogeochemical Cycles, 25,* GB3010, 1-14. doi:10.1029/2010GB003938

Fernandes, R., Prevost, C., Canisius, F., Leblanc, S. G., Maloley, M., Oakes, S., Holman, K., & Knudby, A. (2018). Monitoring snow depth change across a range of landscapes with ephemeral snowpacks using structure from motion applied to lightweight unmanned aerial vehicle videos. *The Cryosphere 12,* 3535-3550. doi:10.5194/tc-12-3535-2018

Garfin, G., Jardine, A., Merideth, R., Black, M., & LeRoy, S. (2013). Natural ecosystems. Assessment of Climate Change in the Southwest United States: A Report Prepared for the National Climate Assessment. https://doi.org/10.5822/978-1-61091-484-0_8

Gaskill, D. L., Mutschler, F. E., Kramer, J. H., Thomas, J. A., & Zahony, S. G. (1991). Geologic map of the Gothic Quadrangle. Colorado: Gunnison County.

Gevaert, C. M., & García-Haro, F. J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment, 156,* 34–44. doi:10.1016/j.rse.2014.09.012

Hahm, W. J., Riebe, C. S., Lukens, C. E., & Araki, S. (2014). Bedrock composition regulates mountain ecosystems and landscape evolution. *Proc. Natl. Acad. Sci. {U.S.A.}, 111*(9), 3338–3343. doi:10.1073/pnas.1315667111

Hall, D. K., V. V. Salomonson, & G. A. Riggs. 2016. MODIS/Terra Snow Cover Daily L3 Global 500m Grid. Version 6. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center.

Hubbard, S. S., Williams, K. H., Agarwal, D., Banfield, J., Beller, H., Bouskill, N., Brodie, E., Carroll, R., Dafflon, B., Dwivedi, D., Falco, N., Faybishenko, B., Maxwell, R., Nico, P., Steefel, C., Steltzer, H., Tokunaga, T., Tran, P. A., Wainwright, H., & Varadharajan, C. (2018). The East River, Colorado, Watershed: A Mountainous Community Testbed for Improving Predictive Understanding of Multiscale Hydrological-Biogeochemical Dynamics. *Vadose Zone Journal*, 17:180061. doi:10.2136/vzj2018.03.0061

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kim, H., Dietrich, W. E., Thurnhoffer, B. M., Bishop, J. K. B., & Fung, I. Y. (2017). Controls on solute concentration-discharge relationships revealed by simultaneous hydrochemistry observations of hillslope runoff and stream flow: The importance of critical zone structure. *Water Resources Research, 53*, 1424-1443. doi:10.1002/2016WR019722

Kirschbaum, M. U. F. (2000). Will changes in soil organic carbon act as a positive or negative feedback on global warming? *Biogeochemistry, 48*(1), 21-51. doi:10.1023/A:1006238902976

Klos, P. Z., Link, T. E., & Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate. *Geophysical Research Letters, 41,* 4560–4568. doi:10.1002/2014GL060500

Knowles, J. F., Harpold, A. A., Cowie, R., Zeliff, M., Barnard, H. R., Burns, S. P., Blanken, P. D., Morse, J. F., & Williams, M. W. (2015). The relative contributions of alpine and subalpine ecosystems to the water balance of a mountainous, headwater catchment. *Hydrological Processes, 29,* 4794–4808. doi:10.1002/hyp.10526

Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma, 123*, 1-22. doi:10.1016/j.geoderma.2004.01.032

Lievens, H., Demuzere, M., Marshall, H. P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim , E. J., Koch, I., Marty, C., Saloranta, T., Schober, J., & De Lannoy, G. J. M. (2019). Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications, 10,* 1-12. doi:10.1038/s41467-019-12566-y

Lin, H. (2009). Earth's Critical Zone and hydropedology: concepts, characteristics, and advances. *Hydrology and Earth System Sciences, 14,* 25-45. https://www.hydrol-earth-syst-sci.net/14/25/2010/hess-14-25-2010.pdf

Marks, D., Winstral, A., Reba, M., Pomeroy, J., & Kumar, M. (2013). An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Advances in Water Resources Research, 55,* 98-110. dx.doi:10.1016/j.advwatres.2012.11.012

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). Declining Mountain Snowpack in Western North America. *American Meteorological Society, January 2005,* 39-49. doi:10.1 175/BAMS-86-1 -39

Nolin, A. W. (2010). Recent advances in remote sensing of seasonal snow. *Journal of Glaciology, 56*(200), 1141-1150.

NRC (National Research Council). (2001). Basic Research Opportunities in Earth Science. *National Academy Press*, Washington, 154 pp

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment, 113*, 868-879. doi:10.1016/j.rse.2009.01.001

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C., & Winstral, A. (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment, 184*, 139–152. doi:10.1016/j.rse.2016.06.018

Rempe, D. M., & Dietrich, W. E. (2018). Direct observations of rock moisture, a hidden component of the hydrologic cycle. *Proc. Natl. Acad. Sci. {U.S.A.}, 115*(11), 2664–2669. doi:10.1073/pnas.1800141115

Riebe, C. S., Hahm, W. J., & Brantley, S. L. (2017). Controls on deep critical zone architecture: a historical review and four testable hypotheses. *Earth Surface Processes and Landforms, 42*, 128–156. doi:/10.1002/esp.4052

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment, 65*, 267–279. doi:10.1016/S0034-4257(98) 00037-6

Roering J. J., Marshall, J., Booth, A. M., Mort, M., & Jin, Q. (2010). Evidence for biotic controls on topography and soil production. *Earth and Planetary Science Letters, 298*, 183-190. doi:10.1016/j.epsl.2010.07.040

Salomonson, V. V., & Appel, I. (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment, 89*, 351-360. doi:10.1016/j.rse.2003.10.016

Selkowitz, D. J., Painter, T. H., Rittger, K., Schmidt, G., & Forster, R. (2017). The USGS Landsat Snow Covered Area Products: Methods and Preliminary Validation CHAPTER 5. In Automated Approaches for Snow and Ice Cover Monitoring Using Optical Remote Sensing (pp. 76). The University of Utah.

Skiles, S. M., & Painter, T. (2017). Daily evolution in dust and black carbon content, snow grain size, and snow albedo during snowmelt, Rocky Mountains, Colorado. *Journal of Glaciology, 63*(237), 118-132. doi:/10.1017/jog.2016.125

Smith P., M. Bustamante, H. Ahammad, H. Clark, H. Dong, E.A. Elsiddig, H. Haberl, R. Harper, J. House, M. Jafari, O. Masera, C. Mbow, N.H. Ravindranath, C.W. Rice, C. Robledo Abad, A. Romanovskaya, F. Sperling, & F. Tubiello, 2014: Agriculture, Forestry and Other Land Use (AFOLU). In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel & J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sprenger, M., Stumpp, C., Weiler, M., Aeschbach, W., Allen, S. T., Benettin, P., Dubbert, M., Hartmann, A., Hrachowitz, M., Kirchner, J. W., McDonnell, J. J., Orlowski, N., Penna, D., Pfahl, S., Rinderer, M., Rodriguez, N., Schmidt, M., & Werner, C. (2019). The Demographics of Water: A Review of Water Ages in the Critical Zone. *Reviews of Geophysics, 57*, 800–834. doi:10.1029/2018RG000633

St. Clair, J., Moon, S., Holbrook, W. S., Perron, J. T., Riebe, C. S., Martel, S. J., Carr, B., Harman, C., Singha, K., & Richter, D. deB. (2015). Geophysical imaging reveals topographic stress control of bedrock weathering. *Science, 350*(6260), 534-538. doi:10.1126/science.aab2210

Storvold, R., Malnes, E., Larsen, Y., Hogda, K. A., Hamran, S. E., Muller, K., & Langley, K. A. (2012). Sar Remote Sensing of Snow Parameters in Norwegian Areas – Current Status and Future Perspective. *Journal of Electromagnetic Waves and Applications, 20*(13), 1751-1759, doi:10.1163/156939306779292192

Tennant, C. J., Crosby, B. T., & Godsey, S. E. (2015). Elevation-dependent responses of streamflow to climate warming. *Hydrological Processes*, *29*, 991–1001. doi:10.1002/hyp.10203

Vicars, W. C., & Sickman, J. O. (2011) Mineral dust transport to the Sierra Nevada, California: Loading rates and potential source areas, *Journal of Geophysical Research, 116,* G01018. doi:10.1029/2010JG001394

Vikhamar, D., & Solberg, R. (2002). Subpixel mapping of snow cover in forests by optical remote sensing. *Remote Sensing of Environment, 84*, 69–82. doi:10.1016/S0034-4257(02)00098-6

Vuyovich, C. M., Jacobs, J. M., & Daly, S. F. (2014). Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States. *Water Resources Research, 50*, 9088-9012. doi:10.1002/2013WR014734

Walters, R. D., Watson, K. A., Marshall, H. P., McNamara, J. P., & Flores, A. N. (2014). A physiographic approach to downscaling fractional snow cover data in mountainous regions. *Remote Sensing of Environment, 152*, 413–425. doi:10.1016/j.rse.2014.07.001

Winnick, M. J., Carroll, R. W. H., Williams, K. H., Maxwell, R. M., Dong, W., & Maher, K. (2017). Snowmelt controls on concentration-discharge relationships and the balance of oxidative and acid-base weathering fluxes in an alpine catchment, East River, Colorado. *Water Resources Research, 53,* 2507-2523. doi:10.1002/2016WR019724

Zhang, H., Zhang, F., Che, T., & Wang, S. (2020). Comparative evaluation of VIIRS daily snow cover product with MODIS for snow detection in China based on ground observations. *Science of the Total Environment, 724*(138156), 1- 11. doi:10.1016/j.scitotenv.2020.138156

Zhou, H., Aizen, E., & Aizen, V. (2013). Deriving long term snow cover extent dataset from AVHRR and MODIS data: Central Asia case study. *Remote Sensing of Environment, 136,* 146–162. doi:10.1016/j.rse.2013.04.015

CHAPTER TWO: VALIDATION OF THE STARFM MODEL FOR USE IN TRACKING SNOW-COVERED AREA CHANGES IN A MOUNTAIN WATERSHED

2.1 Introduction

Understanding the distribution of snow in a landscape, referred to as snow-covered area (SCA) or fractional SCA (*f*SCA), is a fundamental starting point for answering many key questions related to the hydrologic behaviors of watersheds with snow-dominated precipitation regimes. For this reason, SCA/*f*SCA is an input in many hydrological models, largely used to determine changes in the amount of water stored in the landscape in the form of snow water equivalent (SWE) (Elder et al., 1998; Homan et al., 2011). Data assimilation methods incorporating SCA/*f*SCA into hydrological models where it was absent before has been shown to improve many models' ability to replicate basin wide SWE measurements (Andreadis & Lettenmaier, 2006) and timing and volume of spring runoff (Roy et al., 2010). Franz & Karsten (2013) even went as far as to successfully use only SCA data to optimize parameters in snow and discharge models that had traditionally been calibrated using observed streamflow.

SCA has been measured with a variety of well-established remote sensing methods (Hall et al., 1995; Maurer et al., 2003; Nolin, 2010), but historically, the use of these data in hydrologic modeling has often been constrained by the data's coarse spatial resolution. NASA's Moderate Resolution Imaging Spectroradiometer (MODIS), commonly used to determine SCA in a landscape (Roy et al., 2010; Homan et al., 2011; Franz & Karsten, 2017), captures data at spatial resolutions of 250-1000 m, much greater

than the length scales that differences in SCA have been shown to occur (Anderson et al., 2014). The inability to capture these small-scale changes in a landscape with remote sensing can also create other difficulties such as attempting to validate snow remote sensing products with in situ data (Roy et al., 2010).

Even when fine spatial resolution SCA data does exist, its usefulness is often limited by infrequent acquisitions and long revisit times, a problem that can be exacerbated by winter cloud cover obscuring available observations (Selkowitz et al., 2017). The launch of recent satellite missions such as Sentinel-2, which acquires optical data at high spatial resolutions of 10, 20, and 60-m and has a shorter revisit time of 5 days, are greatly enhancing our ability to monitor these quickly occurring, small-scale changes of SCA in a landscape. Any comparable SCA analysis using prior data, however, must rely on methods of downscaling high temporal, low spatial resolution data to observe snow conditions at these scales.

Various methods of downscaling coarse optical spatial resolution data with shorter revisit times have been utilized in previous studies. These include low-computing cost linear interpolation methods (Emelyanova et al., 2012), Bayesian unmixing methods (Gevaert & García-Haro, 2015), and more recently, advanced methods such as deep learning convolutional neural networks (Tan et al., 2018). Until approximately the last decade, however, methods such as these had largely been untested for applications of high-resolution snow cover monitoring (Walters et al., 2014; Berman et al., 2018; Margulis et al., 2019), and thus is still a developing area of study.

Our study utilized the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) with the purpose of downscaling daily MODIS data to a finer spatial

resolution for the specific purpose of SCA monitoring. STARFM is a multi-sensor data fusion model that was created to address the absence of high spatial resolution, frequent coverage datasets (Gao et al., 2006). STARFM was designed to preserve reflectance data during downscaling and was developed using Land Remote-Sensing Satellite System (Landsat) 7 Enhanced Thematic Mapper Plus (ETM+) and MODIS data due to the large amount of overlap in their spectral bands (Table 2.1) (Gao et al., 2006). Although it has historically been employed to capture rapid phenological changes at the landscape scale by generating synthetic data from spectral wavelengths used to monitor vegetation patterns (Hilker et al., 2009; Singh, 2011; Wang et al., 2017; Gallagher, 2018), Olsoy et al. (2017) found that STARFM also performed well when synthetic data was generated from spectral wavelengths traditionally used for identifying snow cover. Due to the promising results of Olsoy et al. (2017), it is reasonable to expect that STARFM could be used to successfully monitor daily changes in snow cover, which our study was first to do.

Table 2.1: Band numbers and bandwidth for the Landsat 7 ETM+ and MODIS sensors

Landsat EMT+ Band	ETM+ Bandwidth (nm)	MODIS Land Band	MODIS Bandwidth (nm)
1	450-520	3	459-479
2	530-610	4	545-565
3	630-690	1	620-670
4	780-900	2	841-876
5	1550-1750	6	1628-1652
7	2090-2350	7	2105-2155

A full review of the STARFM algorithm can be found in Gao et al. (2006). The model is available for download at no cost from the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) website (www.ars.usda.gov).

We implement STARFM in this study with an R-shell. The data pre-processing and STARFM methods outlined in the following sections are modified from work by Peters (2016) and Gallagher (2018). In brief, STARFM uses one or more Landsat images and daily MODIS imagery to predict the surface reflectance at 30-m spatial resolution for the days when no Landsat data were obtained (Figure 2.1). STARFM uses a moving window centered around a single Landsat pixel and finds neighboring Landsat pixels from the same image that are spectrally similar to the central pixel. Next, the spectrally similar neighboring Landsat pixels are assigned a weight based on the spectral differences between them and their corresponding MODIS pixels (which are required to have been resampled to 30-m), the temporal differences (from one MODIS image to the next) of the pixels' values, and the actual spatial distance between the central and neighboring pixels. Finally, the surface reflectance value of the central pixel is calculated for the synthetic image. This process repeats for every MODIS image between the Landsat-MODIS pairs and can be characterized by Equation 2.1

$$L\left(x\frac{\omega}{2}, y\frac{\omega}{2}, t\right) = \sum_{i=1}^{\omega}\sum_{j=1}^{\omega}\sum_{k=1}^{n} W_{ijk} * \left(M\left(x_i, y_i, t_k\right) + L(x_i, y_i, t_0) - M(x_i, y_i, t_0)\right) \quad (2.1).$$

The value of the Landsat, $L(x_i, y_i, t_0)$ and MODIS, $M(x_i, y_i, t_0)$ base pair pixels are considered at time t_0 along with the value of the MODIS pixel, $M(x_i, y_i, t_k)$ at the prediction time t to find the value of the central pixel $L\left(x\frac{\omega}{2}, y\frac{\omega}{2}, t\right)$ at time t of the synthetic image with a moving window size ω. W_{ijk} is the weight assigned to each

neighboring pixel that is determined to be similar to the central pixel based on the spatial, temporal, and spectral differences described above.

Figure 2.1: A schematic of the STARFM model. High resolution Landsat data from days 1 and 17 along with daily coarse resolution MODIS data are inputs to the model. Outputs are high resolution imagery predictions for days 2-16. Figure modified from Cammalleri et al. (2014).

In this chapter we outline our validation process for determining the suitability of STARFM to accurately predict snow cover in heterogeneous mountain areas on days when no fine spatial resolution remote sensing data were acquired. We compare our STARFM output with corresponding Landsat acquisitions from the same dates, as well as data from the Airborne Snow Observatory (ASO)'s Light Detection and Ranging (lidar) overflight of our study area. We analyze STARFM's accuracy temporally over an entire calendar year, and spatially over a variety of landscape features known to influence snow cover.

2.2 Methods

2.2.1 Study Area

The East River watershed, located in the Southern Rocky Mountains of Colorado, USA and in the headwaters of the Upper Colorado River Region, (Figure 2.2) drains approximately 750 km², has a mean elevation of 3137-m, and a topographic relief of 1869-m. The large topographic relief enables the catchment to sustain alpine, subalpine, montane, and riparian ecosystems, with the majority of the area covered by conifer and aspen forests (Carroll et al., 2018). The region has a continental climate with long, cold winters and short, cool summers. Two Snow Telemetry (SNOTEL) (wcc.nrcs.usda.gov/snow/) sites are included within the watershed area. The Schofield station (site number 737) and the Butte station (site number 380) are located at elevations of 3,261-m and 3,097-m, respectively. For the almost 40-year period of record, mean annual temperature has been recorded as ~0.5°C at the Schofield station and ~2.4°C at the Butte station. Mean annual precipitation is 1222 mm at Schofield and 663 mm at Butte, the majority of which falls during the winter months as snow in both locations (Carroll et al., 2018). The geology of the SFA has been mapped and consists mainly of Paleozoic and Mesozoic sedimentary layers, including the large Mancos Shale layer, with multiple Cenozoic igneous intrusions (Gaskill et al., 1991).

Figure 2.2: The East River watershed, located in western Colorado, USA in the headwaters of the Upper Colorado River Region. The Butte and Schofield SNOTEL sites located within the watershed are identified in green. The East River at Almont USGS gauging station is identified in blue.

A unique characteristic of the East River Watershed from a remote sensing perspective is that it is located in an area where Landsat Path 34 Row 33 and Path 35 Row 33 overlap (Figure 2.3), making images available every 7-9 days as opposed to every 16 days as is the case for most locations around the globe. This overlap in coverage is convenient for this study, as it provides additional dates for model input and validation.

Figure 2.3: Landsat 8 OLI Path 35 Row 33 and Path 34 Row 33 areas with the East River watershed study site identified by the blue marker in the overlapping area. Source imagery: Landsat Acquisition Tool (https://landsat.usgs.gov/landsat_acq)

2.2.2 Data

The time frame chosen for the following sensitivity analysis and validation of the model is the water year (WY) 2016, although our validation analysis begins in July 2015 and concludes in July 2016 (Table 2.2). WY 2016 was chosen after reviewing maximum snow water equivalent (SWE) data (Figure 2.4) from the Butte Snow Telemetry (SNOTEL) station. The station is within the East River watershed and is located approximately 3 km from the town of Crested Butte, CO at an elevation of 3,097-m. The mean maximum SWE as shown in Figure 2.4 for WYs 2000-2018 is 14.39 in. The maximum SWE value for WY 2016 is 13.3 in, which is close to, but not above, the mean

Table 1.2 Landsat test periods and dates removed for STARFM validation analysis.

Date Before:	Date Removed:	Date After:
2015:		
June 26	July 12	July 28
July 21	July 28	August 7
July 28	August 6	August 13
August 6	August 13	August 22
August 13	August 22	August 29
August 22	August 29	September 7
August 29	September 23	October 9
September 30	October 9	October 16
October 9	October 16	October 25
October 25	November 1	November 10
November 1	November 10	November 17
November 10	November 17	November 26
November 17	December 3	December 28
December 19	December 28	January 4, 2016
2016:		
December 28, 2015	January 4	January 13
January 4	January 13	January 20
January 20	January 29	February 6
March 1	March 8	March 17
March 8	March 17	March 24
March 17	March 24	April 2
March 24	April 2	April 9
April 18	April 25	May 4
April 25	May 4	May 11
May 4	May 11	May 20
May 27	June 5	June 12
June 5	June 12	June 21
June 12	June 21	June 28
June 21	June 28	July 7
June 28	July 7	July 14

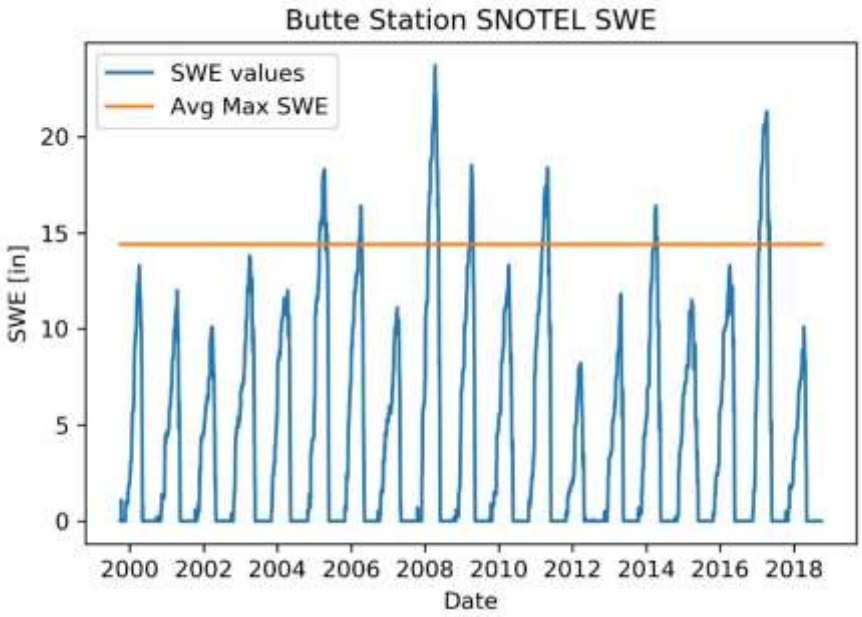


Figure 2.4: SWE values by WY from the Butte SNOTEL station located within the East River watershed. The orange bar represents the mean maximum SWE value from the 2000-2018 WY period.

maximum SWE for the entire time period, and therefore makes a fairly representative data point for a typical WY for analysis.

The fine temporal resolution dataset used in this study was the MODIS product MOD09GA.006 (Terra Surface Reflectance Daily L2G Global 1 km and 500-m) of daily surface reflectance values. This dataset provides daily 500-m reflectance values for Bands 1-7, which have been gridded and corrected for atmospheric gases and aerosols (Vermote et al., 2015), and has imagery available from 24 February 2000 to present.

This choice of MODIS dataset is in contrast to Walker et al. (2012)'s findings that MODIS products with Nadir Bidirectional Reflectance Distribution Function-Adjusted Reflectance (NBAR) produce fewer errors due to Landsat viewing angle differences when used as the coarse resolution dataset for STARFM. Being aware of this, we first attempted to use the MODIS NBAR product MCD43A4, which is a daily 500-m

composite of 16 days of imagery from both Terra and Aqua. We found these data unusable for our study area, however, due to large, consistent data gaps that occurred in our spectral bands of interest and excluded large amounts of area from analysis.

For the fine spatial resolution dataset, United States Geological Survey (USGS) Landsat 8 Surface Reflectance Tier 1 (LC08/C01/T1_SR) was used. These data are processed from the Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) to the Level-1 Precision and Terrain Correction standards (L1TP). These data have been atmospherically corrected using LaSRC. They include cloud, shadow, and snow masks produced with CFMASK, a per-pixel saturation mask, and are projected onto a standardized reference grid (USGS, 2020). Landsat 8 OLI imagery is available from 11 April 2013 to present.

2.2.3 Pre-Processing of Remote Sensing Data

Google Earth Engine (GEE) was utilized for its ability to quickly access, pre-process, and export the large amount of data required for this study. GEE is an internet-based, analysis-ready data catalogue with an application programming interface (API) in a high-performance computing environment (Gorelick et al., 2017). One of the main advantages of GEE is that it allows the user to sort, filter, and perform analyses on large datasets on a cloud-based platform, thereby eliminating the need to download large image files to the user's local machine.

Within GEE's web interface, we filtered our MODIS and Landsat data by dates of interest, area of interest, and set them to a common projection. MODIS data were then filtered via the State QA Scientific Data Set, whose information about the characteristics of each pixel is contained in the "state_1km" band, which is then used to select MODIS

pixels based on their binary number values. In this manner, pixels that are flagged as "clear", "snow and ice", and "assumed clear" are selected, and the remaining pixels not matching these criteria are masked out. Landsat data were filtered similarly using the "pixel_qa" band to select clear and snow scenes and mask out the remaining pixels.

After filtering MODIS and Landsat pixels for cloud cover, all remaining pixels were assigned a Normalized Difference Snow Index (NDSI) value. NDSI is a spectral band ratio that is calculated from the green and short-wave infrared (SWIR) bands (Equation 2.2) calculated as

$$NDSI = \frac{(R_{green} - R_{SWIR})}{(R_{green} + R_{SWIR})} \quad (2.2).$$

NDSI is useful for identifying snow, which has high reflectance in the visible spectrum and low reflectance in the SWIR bands, from other features, particularly clouds, which give stronger signals in the SWIR spectrum, but can be difficult to identify from snow in the visible spectrum. NDSI values resulting from Equation 2.2 range between -1 and 1.

Prior to export from GEE, MODIS imagery was rescaled to 30-m/pixel to match the pixel size of the Landsat imagery, a requirement for STARFM. The final outputs were two image collections of NDSI values that had the same number of bands as the number of days in the date range specified, with matching dates aligned between both raster stacks. For the Landsat image collection, the bands for the dates of Landsat data acquisition contained actual data, and the bands with no Landsat data remained empty. The image collections were then exported from GEE onto Google Drive as Georeferenced Tiff (GeoTiff) files, along with a separate .csv file containing the dates of analysis for reference.

2.2.4 STARFM

Inputs to STARFM were the GeoTiff files of NDSI values calculated from Landsat and MODIS imagery in GEE. Because STARFM had never before been used to identify and predict snow cover in a mountain watershed, a sensitivity analysis was first performed to identify the spatial scale at which the model performed the best, as well as to tune the parameters of the model to the study area.

To determine how STARFM handled increasing levels of heterogeneity in the mountain landscape, we developed a four-tiered model domain approach. Landsat NDSI data were used as "ground-truth" data to evaluate model performance for three separate dates during the 2016 spring snowmelt season. We chose this time of year as a representative timeframe for when heterogeneity is likely to be highest in the landscape as snow cover is decreasing but large patches of snow still remain in the landscape. The imagery dates used in analysis and the corresponding SCA for both Landsat and STARFM images are shown in Tables 2.3 and 2.4. For each of the four STARFM runs, the size of the modeled area, or area of interest (AOI), was progressively increased (Figure 2.5). AOIs tested were approximately 12 km^2, 405 km^2, and 990 km^2, encompassing approximately the area of a single mountain-top, one-half, and two-thirds of the watershed area, respectively. The largest AOI measured around 2200 km^2 and was slightly larger than the entire watershed area. Predicted NDSI values from each model run were then plotted against Landsat NDSI values from the same AOI, and the coefficient of determination was calculated for each pair. The optimal AOI was determined to be 2200 km^2, as it was the domain area with the highest consistent R^2

values across all test dates. For the remainder of the study, and AOI of at least 2200 km² was used for analysis.

Table 2.3: Fraction of AOI with snow cover for each date of Landsat images used in sensitivity tests

Size of AOI	12 km²	405 km²	990 km²	2200 km²
24 March 2016	0.82	0.76	0.60	0.40
2 April 2016	0.93	0.90	0.90	0.75
25 April 2016	0.65	0.38	0.52	0.38

Table 2.4: Fraction of AOI with snow cover for each date of STARFM images used in sensitivity tests

Size of AOI	12 km²	405 km²	990 km²	2200 km²
24 March 2016	0.94	0.90	0.87	0.31
2 April 2016	0.83	0.72	0.56	0.21
25 April 2016	0.69	0.42	0.53	0.15

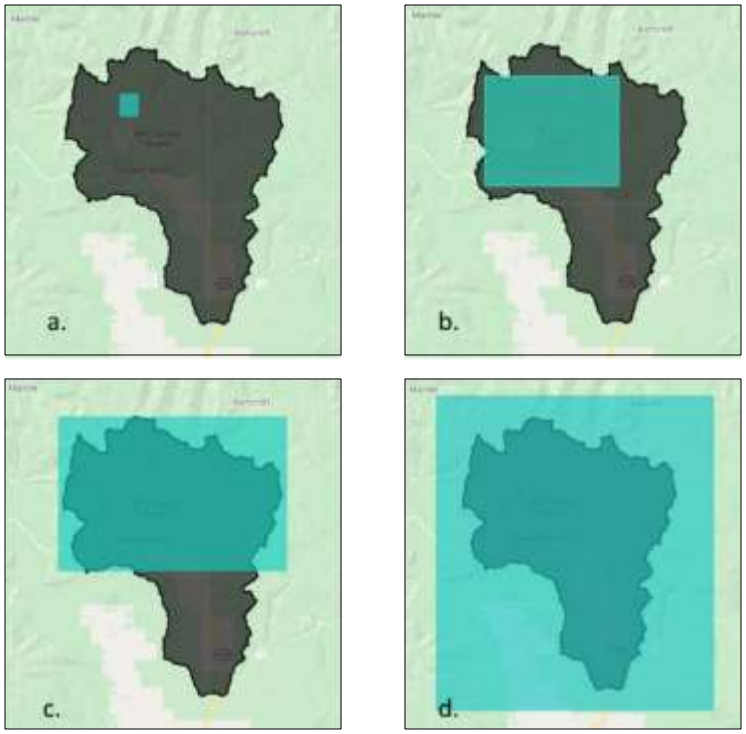


Figure 2.5: Modeled AOIs for the STARFM sensitivity test, a) 12 km², b) 405 km², c) 990 km², and d) 2200 km². The East River watershed is outlined in black in each image. Source imagery: Google Earth Engine.

With the optimal domain for the study area determined, parameter tuning was used to find the ideal specifications for the model. According to the theoretical basis of the moving window method of STARFM, a highly heterogeneous landscape such as our study area should require a large weighted moving search window size (tens to hundreds of meters) for optimal model performance (Zhu et al., 2010). During our parameter tuning, however, we found that STARFM performed better with a smaller search window size range of 25-15 m, much smaller than what is typically used when running STARFM in other, more homogenous landscapes (Gao et al., 2006; Gallagher, 2018). 15-m was the search window size selected for the remainder of the study. We also tested the spectral slice value parameter, which tells STARFM the number of land cover classes expected in

the image. A larger number represents stricter conditions for selecting spectrally similar pixels from the fine-resolution pixels within the search window distance (Zhu et al., 2010). We did not notice a difference in model performance, however, when we experimented with changing the values of this parameter. We left the value at 40 spectral slices, which is the median value of the range used in other studies (Gevaert and García-Haro, 2015). The final parameter, the Landsat and MODIS uncertainty values, were kept at their default values.

The STARFM output is a stack of GeoTiff files (one for each day of analysis) of the interpolated, synthetic NDSI values.

2.2.5 Temporal Analysis of STARFM Performance

To validate the results of the STARFM model and understand how the model performs at different times of the year as snow cover changes, we used a data denial method where single Landsat images were excluded from the input data one at a time. In each instance, the model was run as if that date did not have a Landsat acquisition. The synthetic image produced by STARFM was then compared to the Landsat data from that date to determine how well the model performed at the prediction date. This was done for every date of Landsat acquisition from July 2015 to July 2016 (see Table 2.2) where data were available for an AOI of approximately 2,600 km², which included the East River watershed and surrounding area. Dates that were more than 80% cloud covered were not used for analysis. Although this is a high threshold for cloud covered area (Andreadis & Lettenmaier, 2006; Roy et al., 2010; Walters et al., 2014), the large AOI meant there was enough data to calculate meaningful performance metrics even with only 20% of data available (approximately 520 km²) for a single date. Above 80% cloud coverage,

however, too little data was left to calculate performance metrics for any meaningful comparison. Additionally, a data gap exists in our analysis in February and early March 2016 when the MODIS sensor went offline from 19-27 February 2016. Even with these data restrictions, we still retained 29 dates with varying cloud cover and SCA for analysis (Figure 2.6).

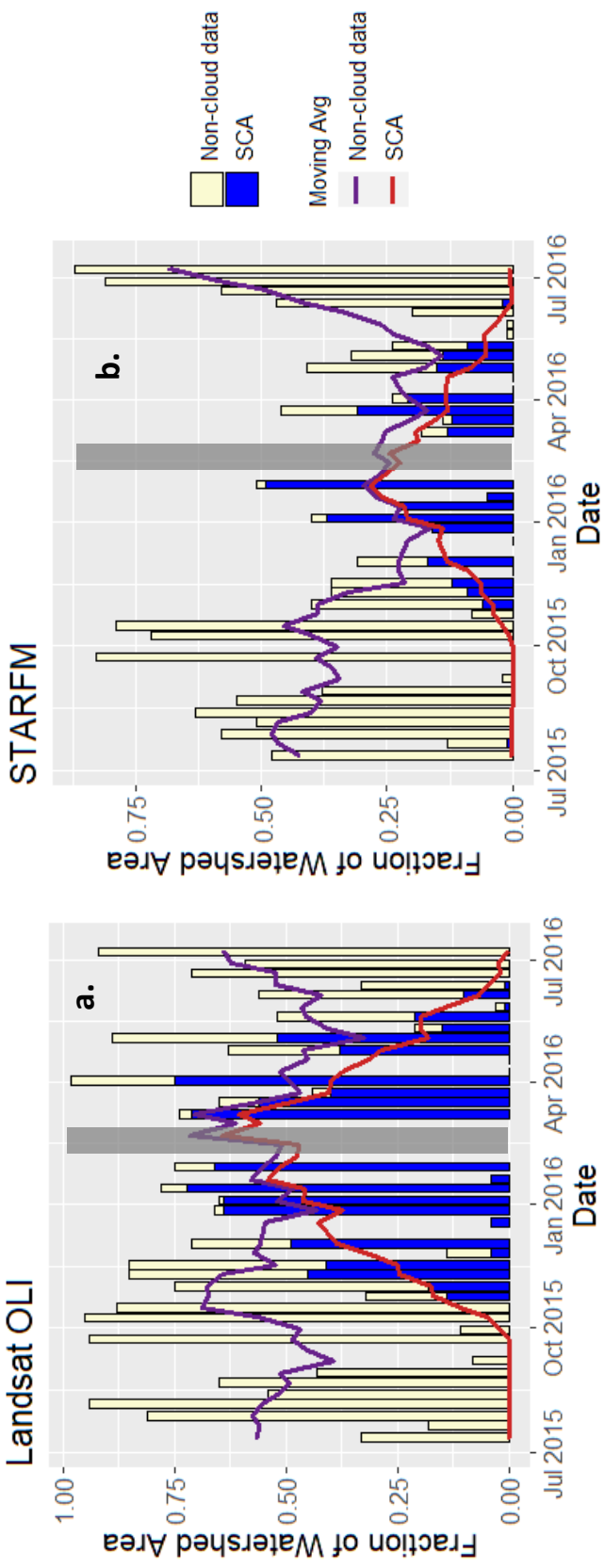


Figure 2.6: Fraction of watershed area of non-cloud covered data available (yellow) for each analysis date and the SCA of that data (blue) for a) Landsat 8 OLI data, and b) STARFM synthetic data. The gap area highlighted in gray is due to MODIS being offline from 19-27 February. All other gaps are due to cloud cover. The lines on the plots represent the 7-day centered moving average for the non-cloud covered data (purple) and SCA data (red).

The NDSI was the metric used to determine SCA for both Landsat acquisitions and synthetic modeled data (Figure 2.7). We used a threshold of NDSI = 0.4 to indicate the presence of snow, with any pixel value at or above this threshold being classified as "snow", similar to the SNOWMAP algorithm classification by Hall et al. (1995). Although we recognize recent findings have shown that optimal NDSI threshold values can be dependent on many factors, including but not limited to study area, elevation, time of year, land cover, and snow depth (Tong et al., 2020; Zhang et al., 2020), an extensive evaluation of the optimal NDSI threshold for our study area was beyond the scope of this study.

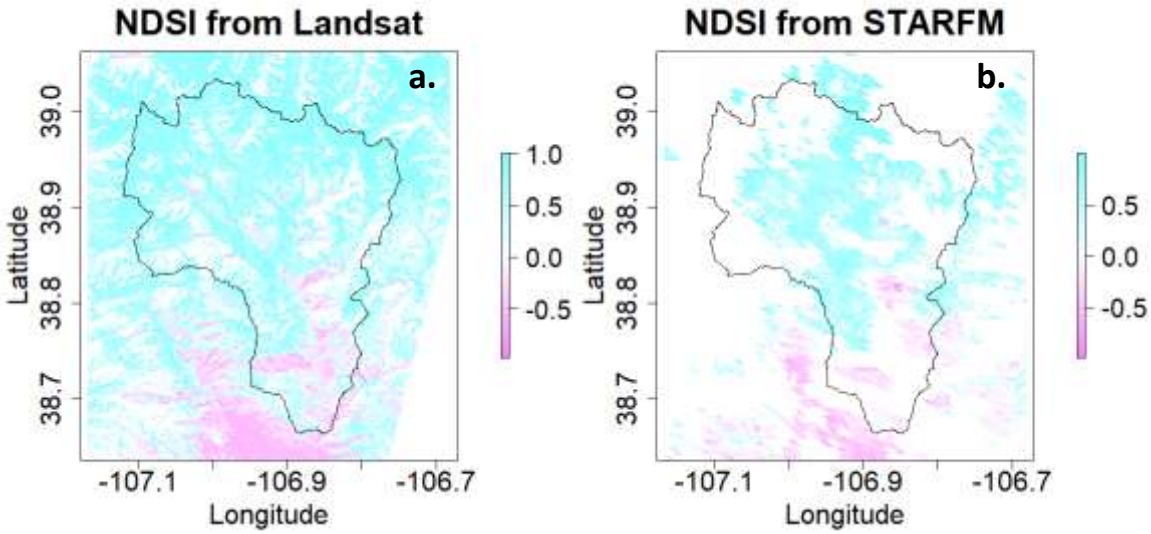


Figure 2.7: Distribution of calculated NDSI values for 3 December 2015 from a) Landsat data and b) STARFM synthetic data.

All pixels in both images were classified into binary "snow" and "no-snow" categories based on their NDSI value (Figure 2.8). Once all pixels had been classified, a confusion matrix was generated using "snow" as the positive class. True positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) occurrences were

obtained by comparing a Landsat pixel value to its corresponding STARFM output pixel value. If the NDSI value of the model pixel classifies as “snow”, but the NDSI value of the Landsat pixel classifies as “no snow” at that location, this results in a false positive. Conversely, if the model NDSI value classifies as “no-snow”, but Landsat classifies as

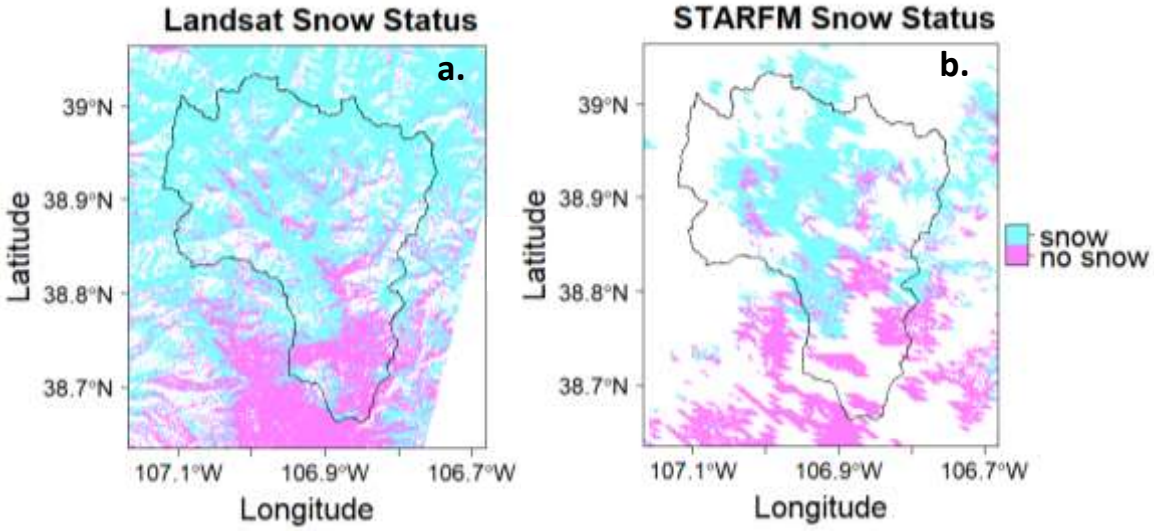


Figure 2.8: Binary snow distribution for 3 December 2015 from a) Landsat data and b) STARFM synthetic data.

“snow”, this results in a false negative. If both the model pixel value and the Landsat pixel value both result in a classification of “snow” or “no-snow”, then this results in a true positive or true negative, respectively.

The results from the confusion matrix can be used to calculate the accuracy, precision, recall, and F-score of the model. Accuracy (a) is the value of the fraction of pixels (both “snow” and “no-snow”) that the model classified correctly, and is calculated in Equation 2.3 by

$$a = \frac{(TP+TN)}{(TN+TP+FP+FN)} \quad (2.3).$$

This metric alone, however, cannot be used to evaluate model performance. As snowmelt progresses throughout the spring season, the "no-snow" category becomes the majority classifier, which can falsely inflate the accuracy value. In other words, the number of "no-snow" pixels become so numerous that even an unskilled model predicting all pixels as "no-snow" would have a high accuracy value. Because of this, it is important to consider other classification metrics that directly measure the model's performance to correctly identify the positive class, even when its total instances are low. For this reason, we calculate the values of precision (*p*), recall (*r*), and F-score (*F*) in a manner similar to Rittger et al. (2013) and Walters et al. (2014). Precision (Equation 2.4) is the fraction of model pixels classified as "snow" that actually were "snow",

$$p = \frac{TP}{(TP+FP)} \qquad (2.4).$$

Recall (Equation 2.5) is the fraction of total observed "snow" pixels that are correctly modeled, or the probability of detection of a snow-covered pixel,

$$r = \frac{TP}{(TP+FN)} \qquad (2.5).$$

These metrics, in other words, tell us how well the model does at predicting snow. Additionally, the F-score, which is the harmonic mean of *p* and *r* and provides a robust statistical balance for binary testing, can then be calculated by Equation 2.6,

$$F = 2\left(\frac{p*r}{p+r}\right) \qquad (2.6)$$

to evaluate the overall performance of the model in identifying the positive ("snow") class.

2.2.5.1 ASO Data

A secondary "ground-truth" dataset used to validate STARFM output came from the Airborne Snow Observatory (ASO), a coupled scanning lidar system and imaging spectrometer with the objective of generating comprehensive time-series datasets of SWE and snow albedo in select mountain basins in the western U.S. (Painter et al., 2016). At present, airborne ASO snow depth data for the East River watershed is available for the dates of 4 April 2016, 30-31 March 2018, 24 May 2018, 7 April 2019, and 10 June 2019 (Painter, 2018). We used ASO 50-m resolution snow depth data from the 4 April 2016 sample date (Figure 2.9) Due to uncertainty in snow depth values below 20 cm (H.P. Marshall, personal communication, 3 June 2019), all depth data below 20 cm was classified as "no snow", and all remaining data as "snow". These data were then resampled bilinearly to the 30-m STARFM grid, a confusion matrix was generated with

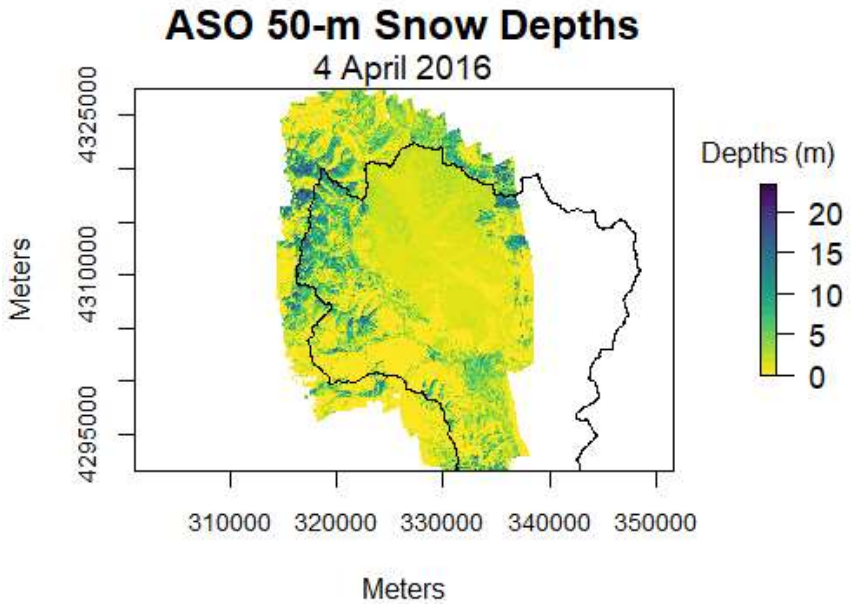


Figure 2.9: Distribution of ASO snow depth data at 50-m spatial resolution for 4 April 2016.

STARFM data from the same date, and *a*, *p*, *r*, and *F* calculated for all pixels with overlapping data (Figure 2.10).

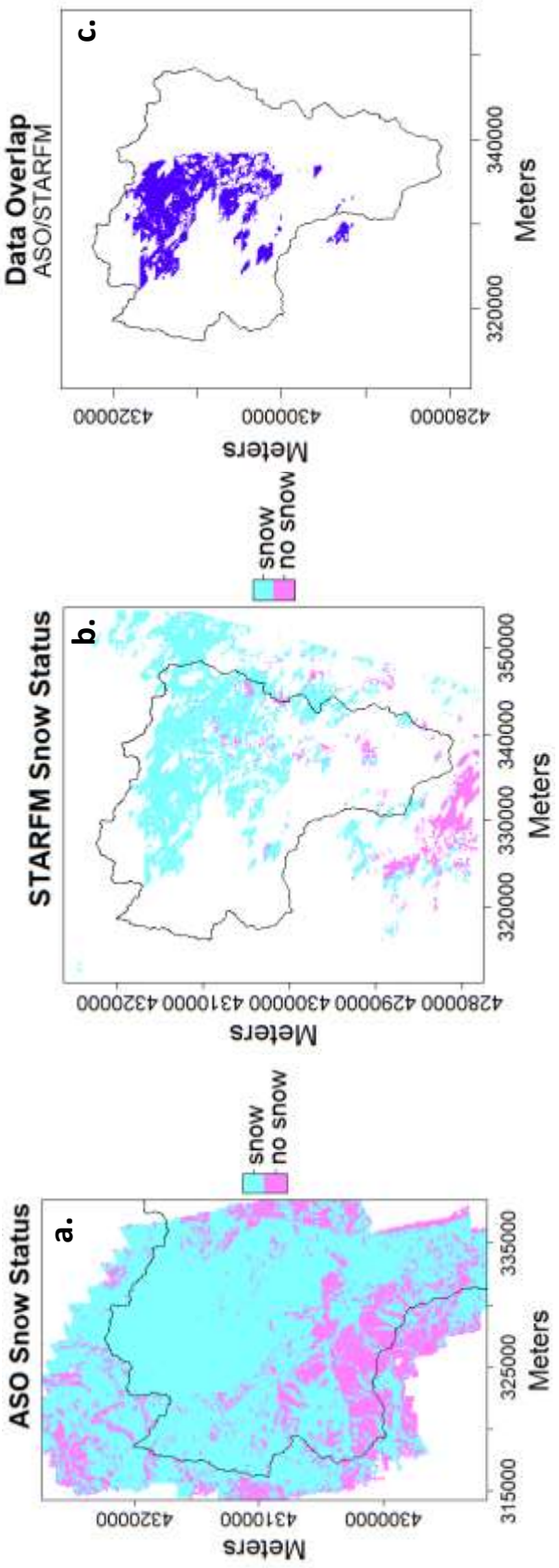


Figure 2.10: a) Binary snow distribution for resampled, 30-m ASO data; b) binary snow distribution for STARFM data; and c) area of overlap between the two datasets.

2.2.5.2 Random Raster Test

The final STARFM temporal validation method was a test to confirm that the STARFM model was better at identifying the correct location of snow on the landscape than a random algorithm placing snow on the landscape. We expected the accuracy of both the model results and results generated by a random algorithm to be high during times when the domain was majority snow covered or majority snow free. During times when SCA was between the two extremes, however, we expected that if our model was correctly predicting snow cover it would outperform a raster with randomly placed snow pixels.

We generated random rasters with the same SCA value for each date in the data denial analysis by randomly resampling the model data 100 times, which preserved the same number of positive ("snow"), negative ("no-snow"), and no data pixels as the model results (Figure 2.11). A confusion matrix was then generated for each of the 100 randomized rasters to compare pixel classifications between the two. In the same manner that the pixel classification of the STARFM data was compared to the Landsat data, the pixel classification of the randomized raster data was compared to the STARFM data. *A*, *p*, *r*, and *F* were computed and stored for each of the 100 confusion matrices for each date. Once this was complete, the mean and standard deviation of the classification metrics for the 100 randomized rasters were calculated for comparison against the same metrics from the STARFM/Landsat validation analysis.

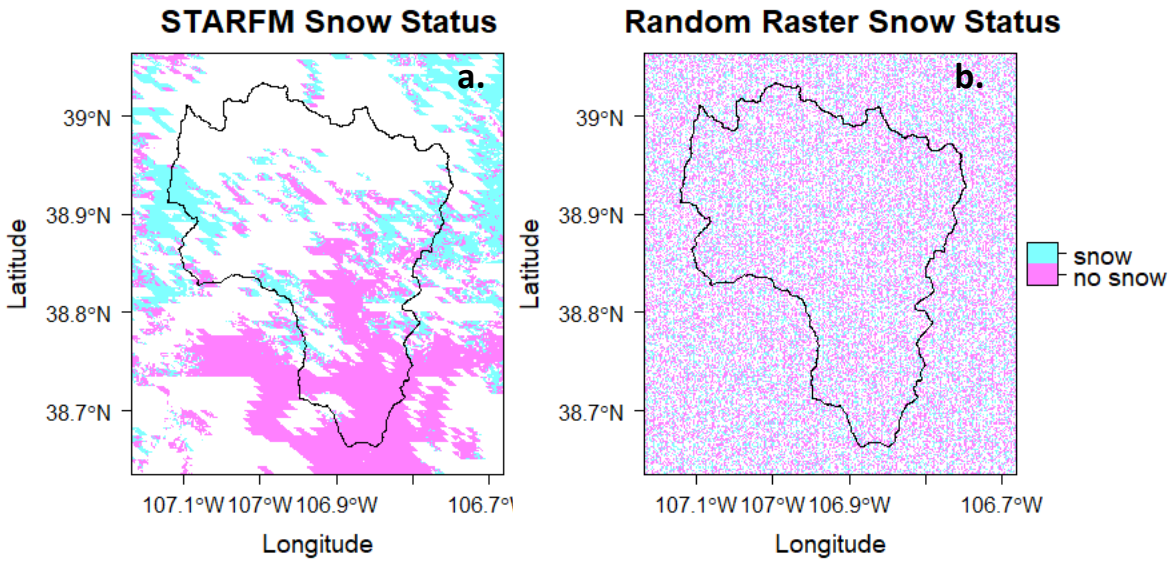


Figure 2.11: a) STARFM binary snow distribution for 25 April 2016 with a SCA = 0.15; b) randomized raster of a.

2.2.6 Spatial Analysis of STARFM Performance

In addition to understanding how STARFM performs at different times of the year, it is also important to determine if the large heterogeneities in the mountain landscape affect the ability of STARFM to accurately determine snow cover. A spatial analysis of STARFM results as they relate to landscape features can assist in answering this question. We decided to examine the relationship between the accuracy of each pixel of STARFM results over the entire one-year validation timeframe and a variety of landscape features that often influence snow cover, which included elevation, slope, aspect, and vegetation class (Anderton et al., 2004; Anderson et al., 2014) .

To determine per-pixel accuracy for model results over the entire year, we counted the number of instances where each pixel was classified correctly (Figure 2.12b), as well as the number of times a prediction was made for that pixel, regardless of correctness (Figure 2.12a), and the difference between the two (Figure 2.12c). As stated

previously, of the 45 dates where a Landsat acquisition was acquired between July 2015 and July 2016, a comparison between the Landsat and STARFM data could only be made for 29 of those dates due to cloud cover, missing data, or data overlap issues. Of those dates, however, not every Landsat pixel had data all 29 times to be compared to the STARFM output, and vice versa. Therefore, the data were normalized by calculating per-pixel accuracy. This metric was determined by the number of times a pixel was correctly identified as "snow" or "no-snow" by STARFM divided by the number of times the pixel had data in both the Landsat and STARFM datasets. Once per-pixel accuracy had been determined where possible for the entire study area (Figure 2.13), smaller subset areas defined by landscape features could be examined for relationships with model performance.

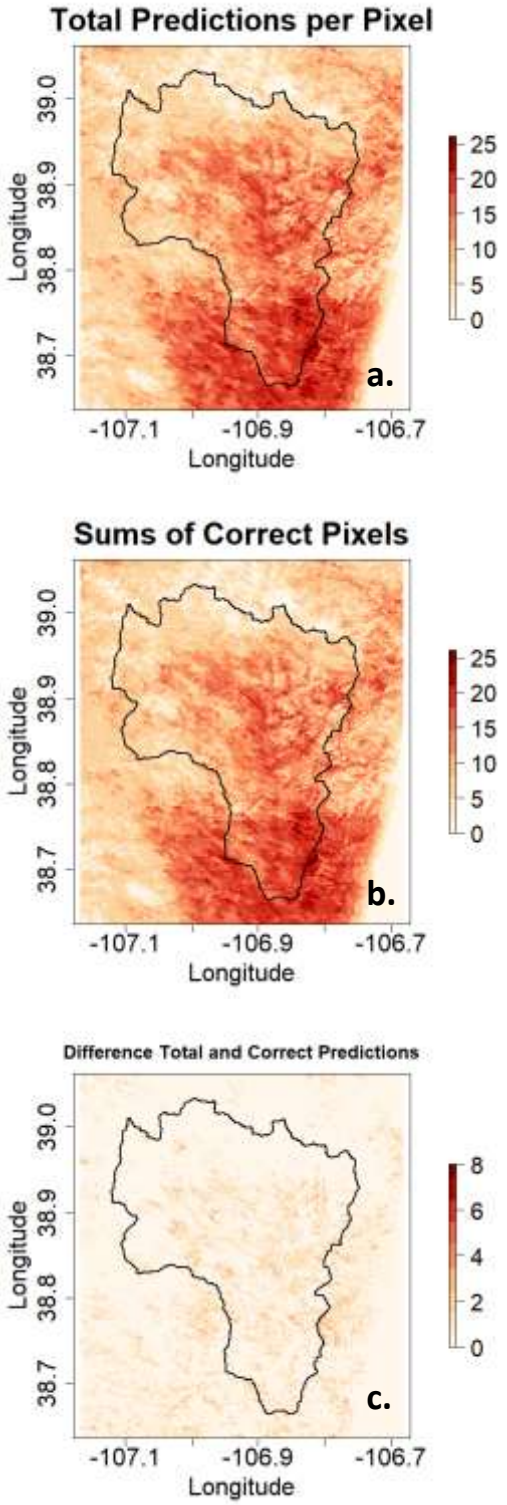


Figure 2.12: a) The distribution of the total number of model predictions per pixel, regardless of whether the prediction was correct or not; b) the distribution of the sum of the instances per pixel in which that pixel was identified correctly; c) the distribution of the differences between a and b.

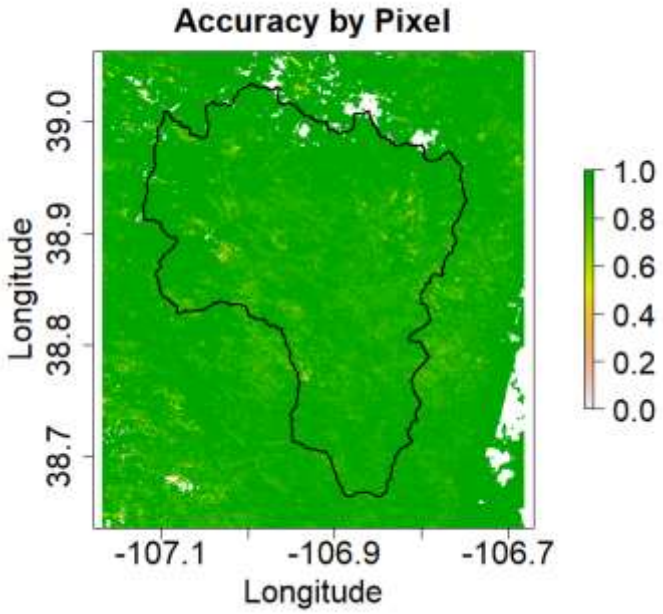


Figure 2.13: Distribution of per-pixel accuracy values calculated with results from all 29 data denial instances.

2.2.6.1 Elevation

We first considered the relationship between model performance and elevation. While there are strong demonstrated relationships between increasing elevation and increasing snow cover duration (Daly et al., 2012; Che et al., 2016), high elevation areas can also contain large variabilities in terms of snow cover. High elevation areas tend to have sparse or no vegetation, so wind scour can leave some areas devoid of snow while depositing that snow in nearby locations (Buhler et al., 2016). Due to these variations and inconsistencies, we divided our watershed into three elevation bands to examine how well our model captures these SCA patterns.

We used NASA Shuttle Radar Topography Mission (SRTM) V3 Digital Elevation 30-m data (Farr et al., 2007) to classify each pixel of our study area into either a Low, Medium, or High elevation band. These classifications were determined by using an

Empirical Cumulative Distribution Function (ECDF) of elevation values (Figure 2.14). Elevation values that fall into the lower one-third of the distribution (2399-2963 m) comprise the Low band, those that are within the top one-third of the distribution (>3314-4332 m) make up the High band, and everything else in between (>2963-3314 m) falls into the Medium band. The mean and standard deviations of accuracy values of the pixels that fall into each of these categories could then be analyzed separately to evaluate model performance by elevation.

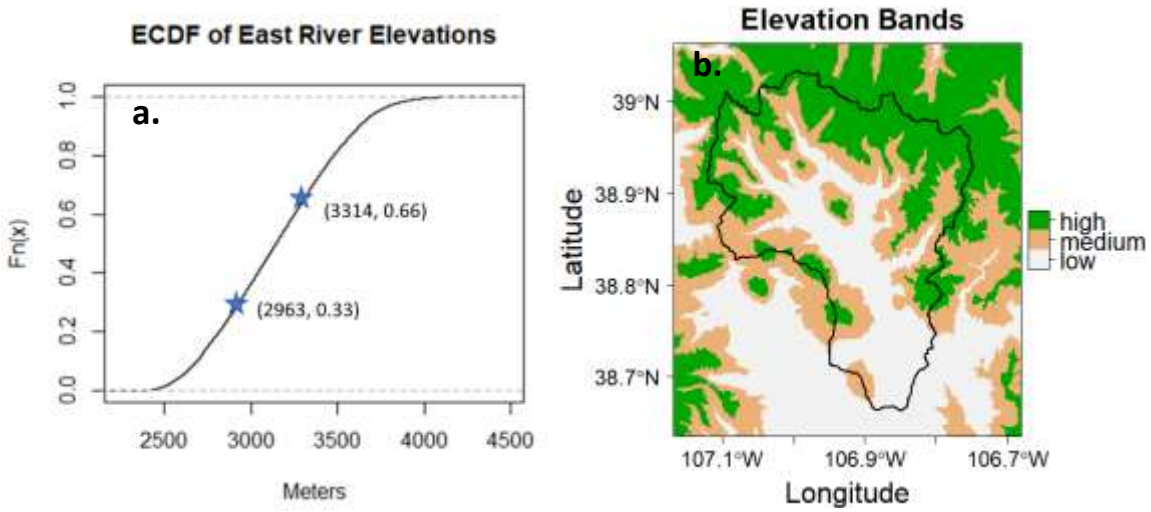


Figure 2.14: a) The empirical cumulative distribution of elevation values for the study area. The elevation values marking the first and second one-thirds of the distribution (y = 0.33, 0.66) are labeled; b) the spatial distribution of the elevation bands.

2.2.6.2 Slope

Slope, or the change in vertical distance over the change in horizontal distance, can also have an influence on snow cover. Areas in mountain watersheds with very steep slopes tend to accumulate the least amount of snow, and due to this often melt earlier when compared to more gently sloped areas (Elder et al., 1998; Kerr et al., 2013). Slope

degree values were first calculated from the STRM DEM data using the Horn (1981) method with 8 neighbors. Slope values were classified similar to elevation values into Low (0-11°), Medium (>11-20°), and High (>20-74°) bands using an ECDF and the distribution divided into thirds (Figure 2.15). Due to the way slope values are distributed in our watershed, any slope measuring greater than 20° is considered a high slope area. Much of the high slope area is well above 25°, the threshold above where movement of snow due to avalanching and or sluffing is likely to occur (McClung & Schaerer, 2006).

The same method used to analyze model performance by elevation was used to analyze model performance by slope. Each pixel was placed into a slope band based on its slope degree. The mean and standard deviation of the accuracy values of pixels located in each slope band were then calculated.

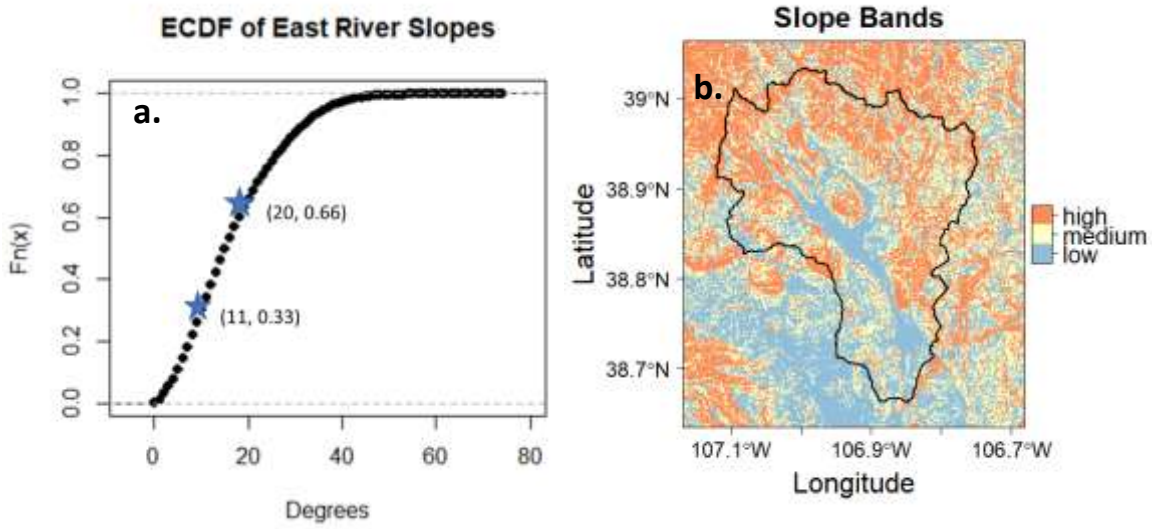


Figure 2.15: a) The empirical cumulative distribution of slope values for the study area. The slope values marking the first and second one-thirds of the distribution (y = 0.33, 0.66) are labeled; b) the spatial distribution of the slope bands.

2.2.6.3 Aspect

Aspect, or the direction a slope faces, can influence snow cover in a variety of ways that often change throughout a winter season. Aspect influences snow accumulation patterns through primary wind direction and storm events and also snow melting patterns due to variability in solar radiation (Anderton et al., 2004; Anderson et al., 2014; Buhler et al., 2016).

Aspect was calculated in degrees from the SRTM DEM data also using the Horn (1981) method with 8 neighbors. Each pixel was then classified by their degree value as either north, east, west, or south-facing. All pixels with values between 315° and 45° were classified as north-facing, 45° and 135° as east-facing, 135° and 225° as south-facing, and 225° and 315° as west-facing (Figure 2.16). Mean and standard deviation of accuracy values for pixels classified as having the same aspect were analyzed to evaluate model performance by aspect.

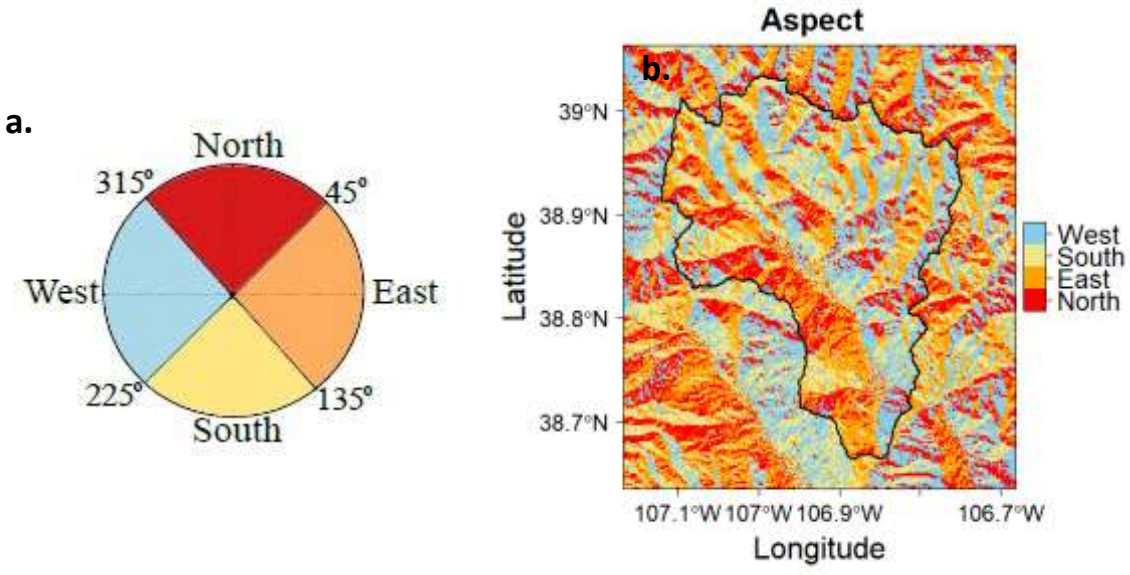


Figure 2.16: a) The range of degree values used for classifying aspect; b) spatial distribution of hillslope aspects identified using this classification. Figure a from Boz et al. (2015).

2.2.6.4 Land Cover Type

The final metric used to analyze model performance by landscape feature was land cover type. Snow patterns and land cover type can be difficult to quantify, and this is especially true when it comes to trees. Trees influence both snow accumulation and melt rates by intercepting and sublimating falling snow and shading snow on the ground from incoming solar radiation (Veatch et al., 2009). Vegetation also shades snow underneath it from overhead view, and thus can cause difficulties in identifying and predicting snow cover with optical data. We wanted to examine whether STARFM's performance predicting snow cover was influenced by land cover type within the watershed.

The USGS Landfire 30-m raster maps of Existing Vegetation Type data (Landfire, 2016) were used to classify land cover by five dominant features in a manner similar to Carroll et al. (2018). The dominant features included trees (60%), shrubs (13%), grass (15%, which included agricultural areas), clear/unvegetated (11%), and other (1%) (Figure 2.17). The "other" land cover feature class included areas of human development, water, and permanent snow/ice, all of which individually occupied too small of an area to perform any meaningful analysis with model performance.

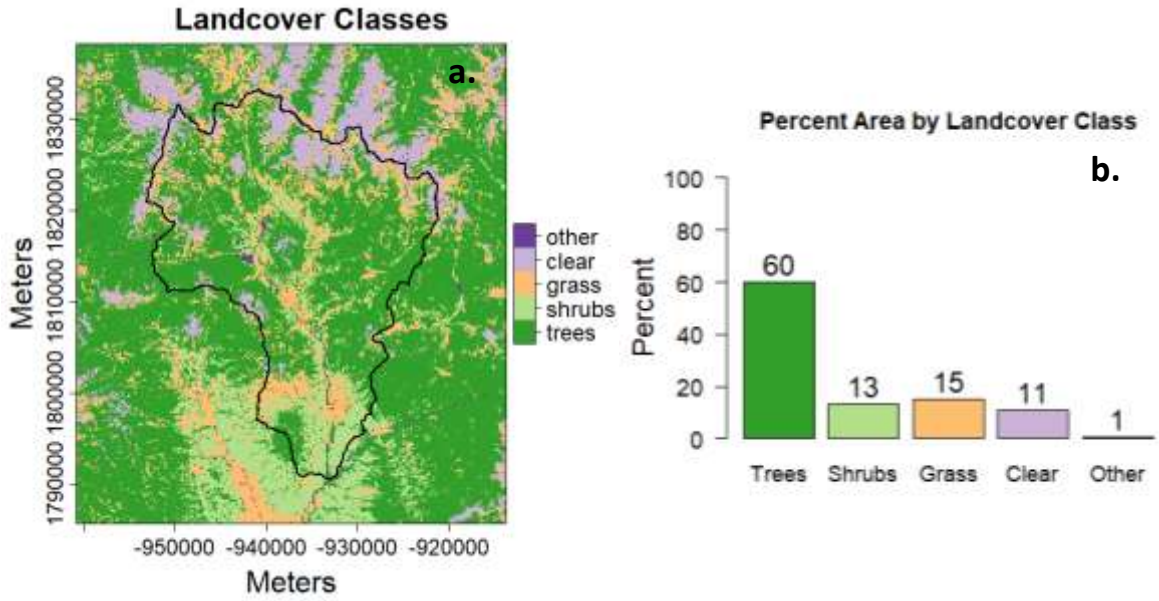


Figure 2.17: a) Spatial distribution of landcover classes; b) percent area of each class from a.

2.3 Results

2.3.1 Sensitivity Analysis

From our sensitivity analysis, conducted by running STARFM at three spring dates with four AOI sizes and evaluating model performance, we determined that the optimal region size for the model was the largest AOI, which encompassed the entire East River watershed and surrounding area at 2200 km². The coefficient of determination values steadily increased with each increase in area, with the exception of two instances at 990 km² for the 24 March 2016 and 25 April 2016 dates, where a slight decrease in the R^2 values (-0.3) occurred with the increase in AOI (Table 2.5). The highest R^2 values for all dates occurred for the model runs covering the largest area.

Table 2.5: R^2 values for each date and AOI tested in the sensitivity analysis

Size of AOI	12 km²	405 km²	990 km²	2200 km²
24 March 2016	0.44	0.52	0.49	0.69
2 April 2016	0.37	0.55	0.58	0.76
25 April 2016	0.32	0.63	0.60	0.79

2.3.2 Temporal Analysis of STARFM Performance

STARFM performance varied by season over the WY 2016 analysis. The accuracy of the model ranged from a minimum value of 0.72 for 17 November 2015, to a maximum of 1.00 for multiple dates in the summer months of both 2015 and 2016 and some early fall dates of 2016 (Figure 2.18a). All dates with perfect accuracy values occur when the SCA for the entire Landsat scene is <= 0.1, indicating that the model can easily predict the majority class of "no snow" in these cases. High accuracy values also occurred during the winter months, which is also a time when the majority classifier is highly skewed towards "snow", with values of 0.98 for 28 December 2015, and 4 and 13 January 2016. The distribution of accuracy values over the entire year, however, had a mean, a_μ, of 0.95 and a standard deviation, a_σ, of 0.07, indicating the model is highly effective at correctly identifying the snow state of the pixels regardless of the season and the heterogeneity in the landscape.

Precision and recall, the metrics that inform how capable the model is at detecting the positive class of "snow", exhibited wide ranges with very low values in the summer and very high values in the winter (Figures 2.18b and 2.18c). Precision ranged from a minimum value of 0.02 on 29 August 2015 to a maximum value of 1.00 on 28 December 2015, and recall had a minimum of 0.17 on 6 August 2015 and a maximum of 1.00 on 22

August 2015. Precision had a mean, p_{μ}, of 0.72 and a standard deviation, p_{σ}, of 0.31, and recall had a mean, r_{μ}, of 0.75 and a standard deviation, r_{σ}, of 0.23. F-score, the harmonic mean of precision and recall, ranged from a minimum value of 0.04 on 29 August 2015 to a maximum of 0.99 on 28 December 2016, 4 January 2016, and 13 January 2016 (Figure 2.18d). The mean f-score, F_{μ}, was 0.69 and the standard deviation, F_{σ}, was 0.28. These metrics tell us that STARFM is generally identifying snow cover correctly where it exists and not misidentifying non-snow pixels as snow.

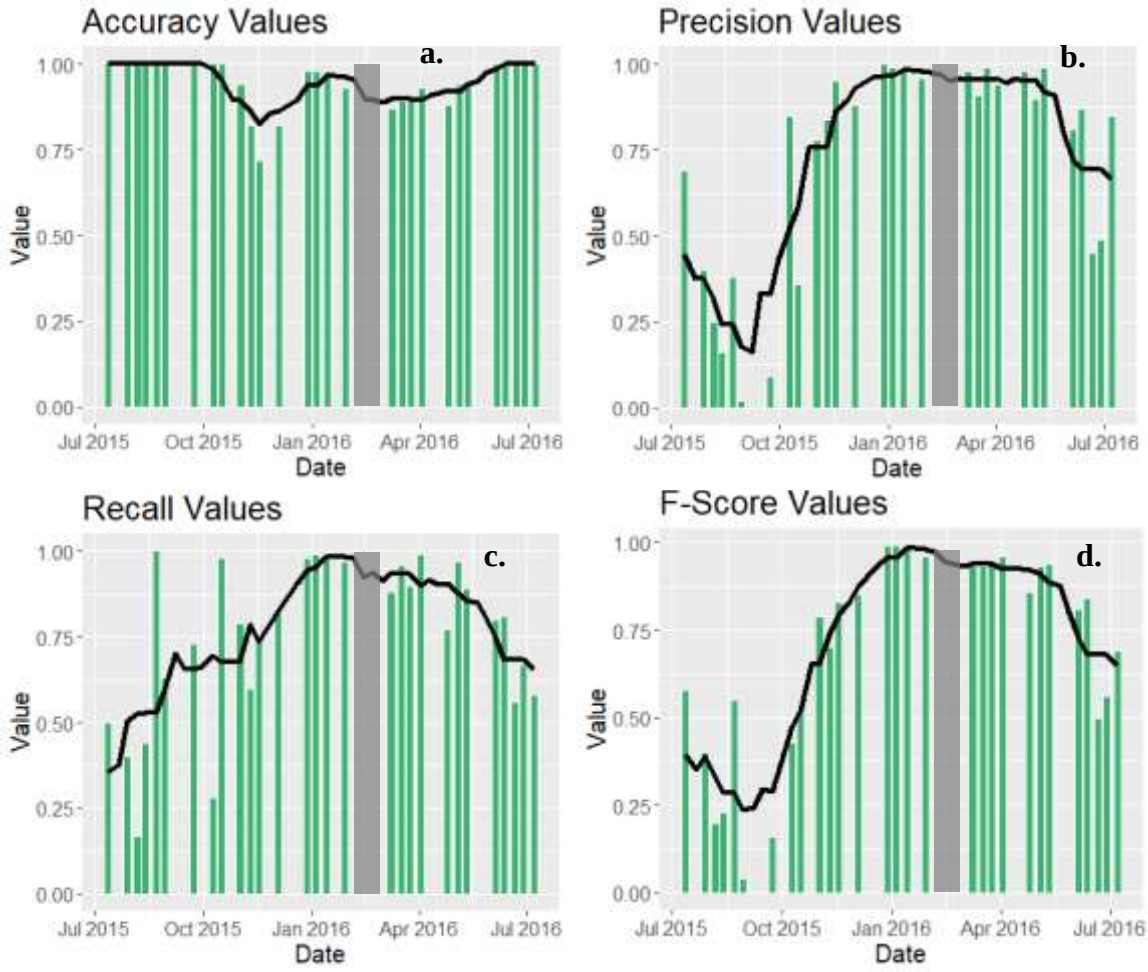


Figure 2.18: The green bars are the values of a) accuracy, b) precision, c) recall, and d) F-score for the year-long model validation runs. The black lines in all figures are the 7-day centered moving averages. The gap areas highlighted in gray are due to MODIS being offline from 19-27 February. All other gaps are due to cloud cover.

2.3.2.1 ASO Data

ASO data is only available for a single date in the WY 2016 timeframe of our analysis, but it is still a useful data point to test the validity of STARFM predictions for that day. While there are issues with clouds obscuring large areas of our satellite data on this date, there is still a 17%, or approximately 360 km², overlap of pixels for the area where ASO data were obtained. When these two datasets were compared, STARFM

performed with an accuracy (a) of 0.89, precision (p) of 0.90, recall (r) of 0.99, and an F-score (F) of 0.94.

2.3.2.2 Random Raster Test

If STARFM can better predict the location of snow-covered pixels in the landscape than a random placement of snow-covered pixels over the same area, then the classification metrics for the former should have higher values than the latter. To compare these scenarios, the mean accuracy (a_2), precision (p_2), recall (r_2), and F-score (F_2) for the 100 random raster tests were compared against the same metrics from the STARFM/Landsat validation analysis.

The a_2 for the random raster tests varied within a range of 0.44 to 1.00 (Figure 2.19a). Very high accuracies, 0.99 or 1.00, occurred for dates when the STARFM synthetic image had no snow cover, so randomly resampling the data here did not provide us with any new information. The mean of a_2 was 0.82 and the standard deviation was 0.20.

Mean precision, p_2, ranged from 0 to 0.96 (Figure 2.19b). We expected high precision values (0.90 or above) when the watershed was fully snow covered due to random resampling being able to easily predict the correct classification. When precision values were high, the SCA of the available data for the region was 92% snow covered or higher. As shown in Figure 2.19a, however, the high precision values in the winter months were still consistently below the precision values calculated from the STARFM/Landsat comparison. Zero values in precision occurred during the summer months when there was no snow or too little snow to compute this value. The mean of p_2 was 0.34 and the standard deviation was 0.38.

Mean recall, r_2, ranged from 0 to 0.98 (Figure 2.19c). Similar to precision, we also expected high recall values when the watershed snow cover was high. Recall values of 0.90 and above occurred on dates when the SCA of available data for the watershed was 72% or higher. There were two instances in which the mean of the 100 recall values from the random rasters was greater than the recall value from the STARFM/Landsat evaluation on the same date, and one date in which the two values were the same. In the same manner as precision, recall values were zero in the summer months when there was too little snow for them to be calculated. The mean of r_2 was 0.41 and the standard deviation was 0.41.

Mean F-score, F_2, varied within a range of 0 to 0.97 (Figure 2.19d), which was also reflective of the trends in maximum and minimum values seen with mean precision and mean recall. Because F-score is the harmonic mean of precision and recall, instances in which one of these values equals zero meant that the F-score could not be calculated for that date, which is why some data are missing in Figure 2.19d. The mean of F_2 was 0.56 and the standard deviation was 0.36.

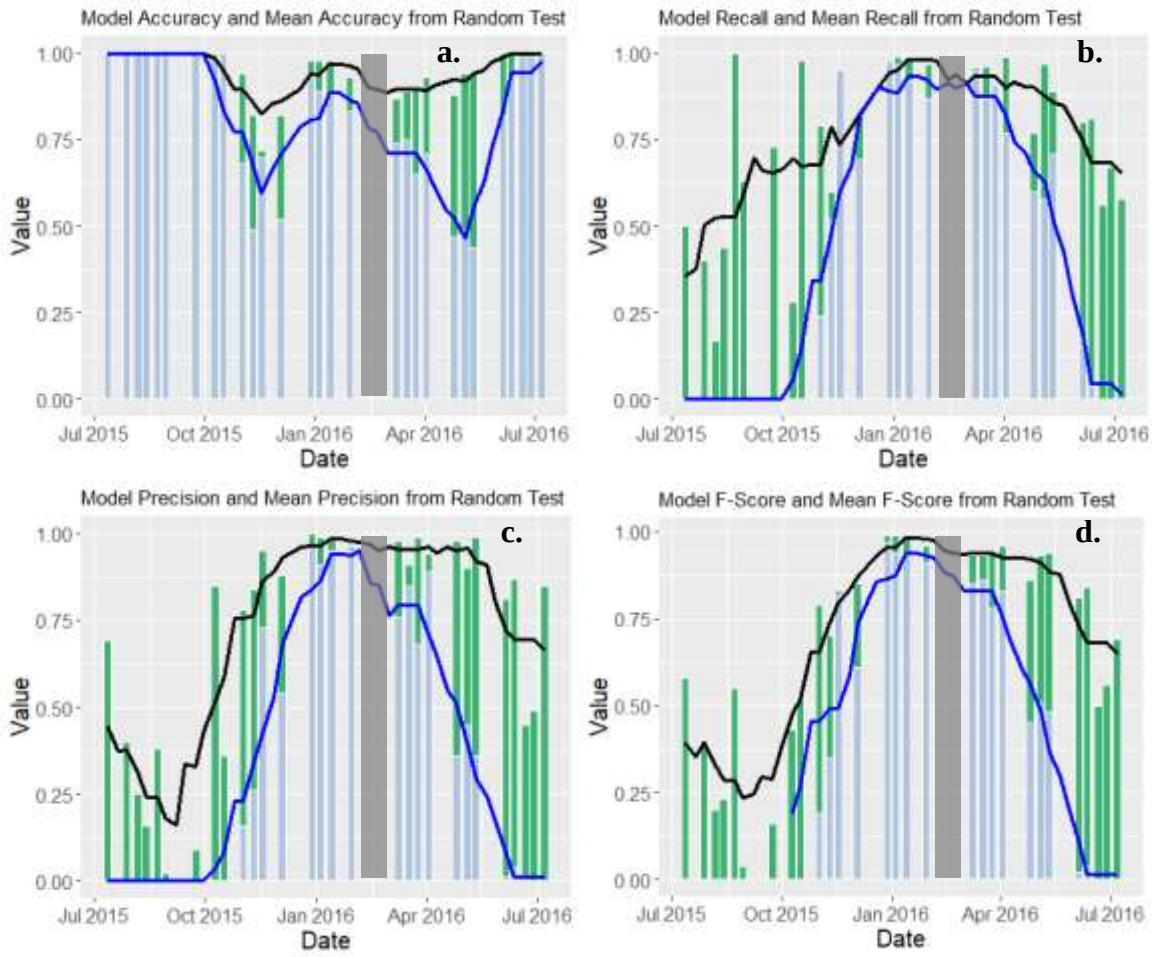


Figure 2.19: The green bars are the values of a) accuracy, b) precision, c) recall, and d) F-score for the year-long model validation runs with 7-day centered moving averages (black) for all metrics. Blue bars are values of a-d for the random raster tests with 7-day centered moving averages in dark blue. The gap areas highlighted in gray are due to MODIS being offline from 19-27 February. All other gaps are due to cloud cover.

2.3.3 Spatial Analysis of STARFM Performance

2.3.3.1 Elevation

We did not find a significant relationship between STARFM performance and elevation band (Figure 2.20). STARFM was able to correctly predict snow for low elevations with an mean accuracy of 0.97 and standard deviation of 0.06, medium elevations with a mean accuracy of 0.96 and a standard deviation of 0.08, and high elevations with a mean accuracy of 0.98 and a standard deviation of 0.08 (Figure 2.21).

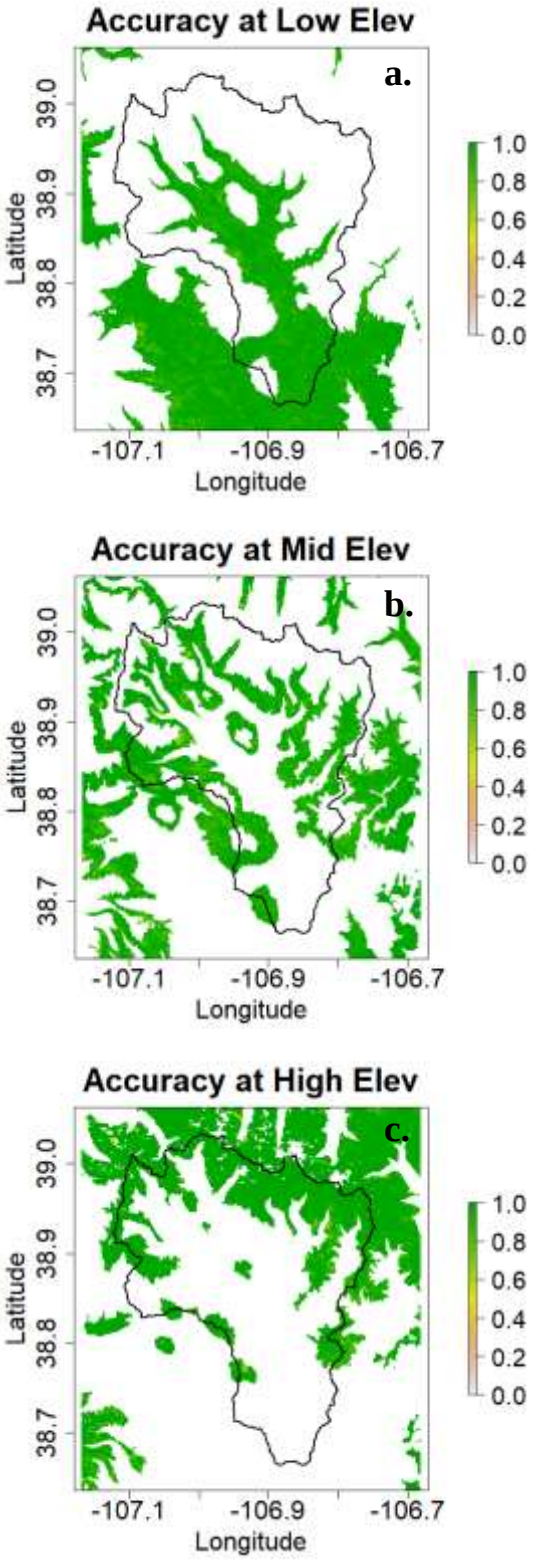


Figure 2.20: Spatial distribution of accuracy values by pixel for the a) low elevation band, b) medium elevation band, and c) high elevation band.

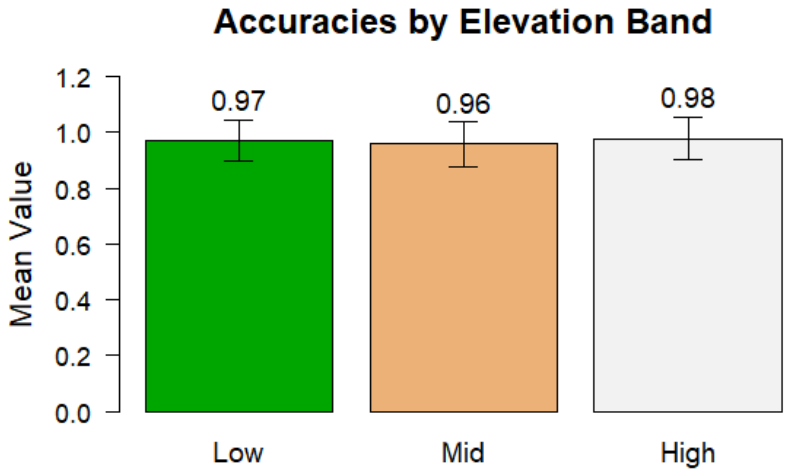


Figure 2.21: Mean accuracies by elevation band. Standard deviations represented by the error bars for low, medium, and high elevations are 0.06, 0.08, and 0.08, respectively.

When analyzing accuracy results, we noticed that there appeared to be fewer total model observations for pixels in the high elevation band. This is a known issue with remotely sensed satellite data that can be caused by greater instances of cloud cover at higher elevations (Hou et al., 2019) or issues with the Landsat and MODIS cloud masks that can misidentify snow as clouds at higher elevations (USGS, 2020). As a result, the input files to STARFM are often missing data at these pixels, and therefore the model cannot synthesize an image at that date and location. To better understand the availability of modeled data at each elevation band, we evaluated the distributions of the number of times all pixels in each elevation band had a model prediction, regardless of correctness. This analysis revealed that the distribution of the numbers of predictions per pixel did decrease with increasing elevation (Figure 2.22). When plotted numerically, the distribution of the number of predictions per pixel for those in the lower elevation band had a lower quartile (Q1) of 10, a median of 15, and an upper quartile (Q3) of 18 (Figure 2.23).

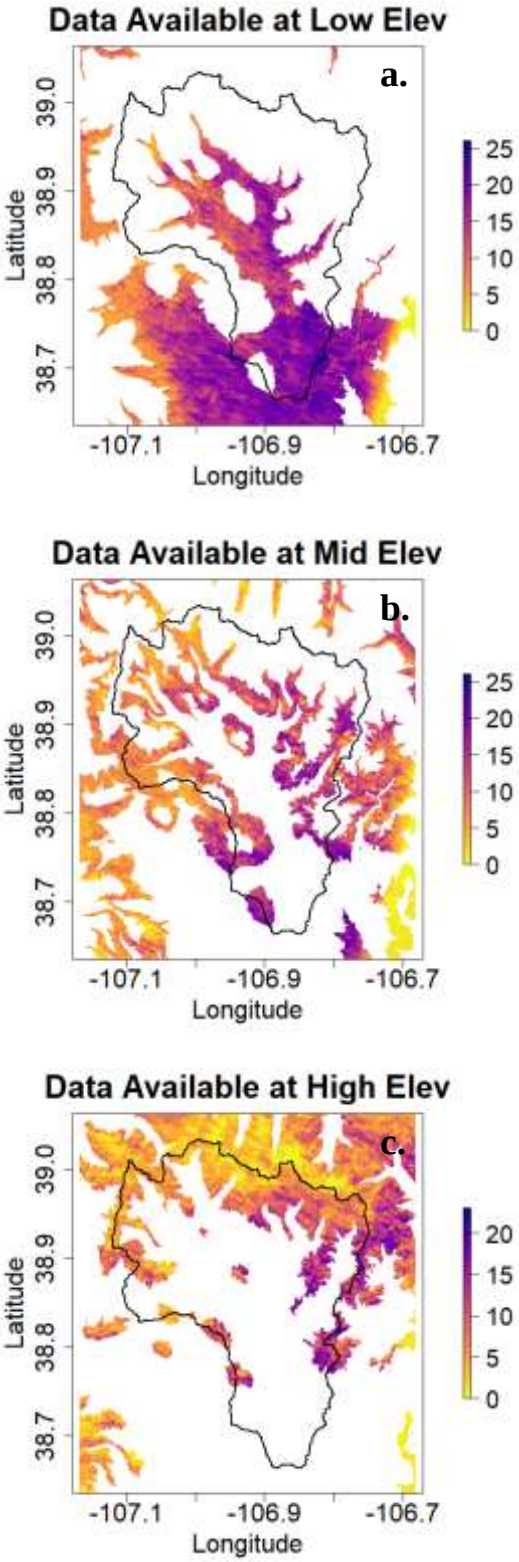


Figure 2.22: Spatial distribution of data availability per pixel for the a) low elevation band, b) medium elevation band, and c) high elevation band.

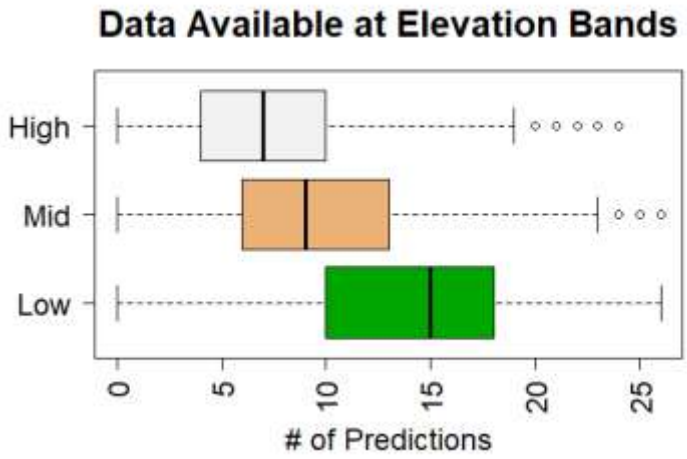


Figure 2.23: Distribution of the number of predictions per pixel, regardless of correctness, for each elevation band.

The pixels in the medium elevation band had a Q1 of 6, a median of 9, and a Q3 of 13, and pixels in the high elevation band had a Q1 of 4, a median of 7, and a Q3 of 10.

2.3.3.2. Slope

STARFM also performed equally well across all slope steepness bands (Figure 2.24). Low slope steepness had a mean accuracy of 0.97 and a standard deviation of 0.07. Medium slope steepness had a mean accuracy of 0.97 with a standard deviation of 0.08, and high slope steepness had a mean accuracy of 0.97 with a standard deviation of 0.08 (Figure 2.25).

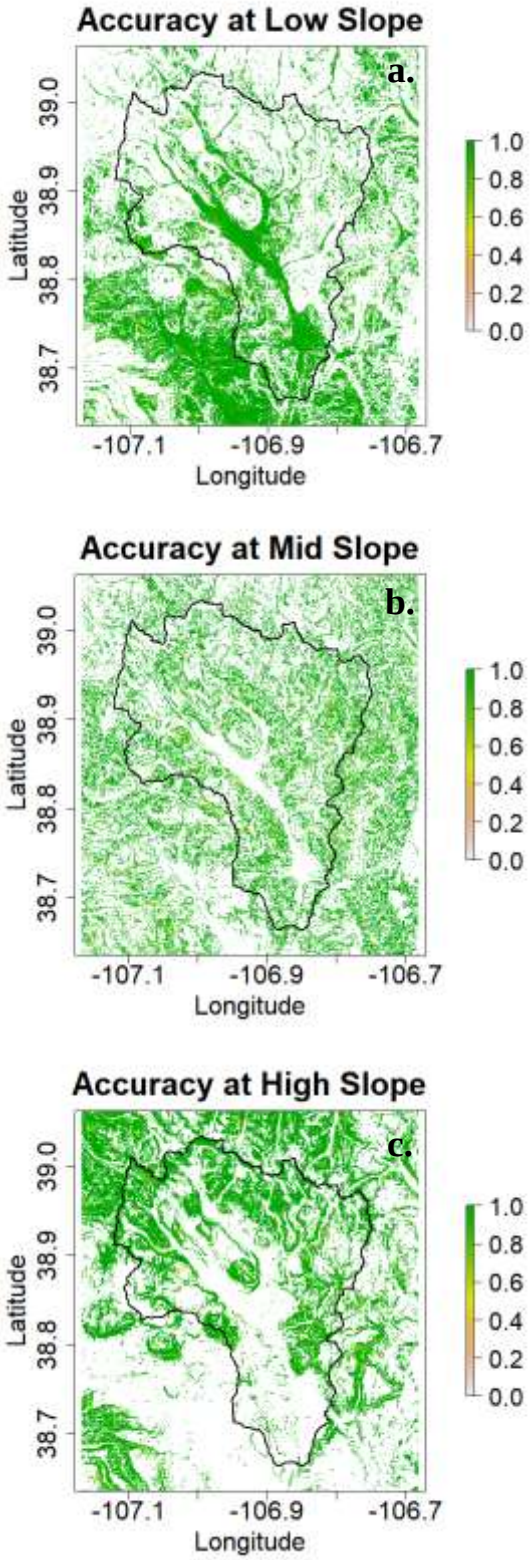


Figure 2.24: Spatial distribution of accuracy values by pixel for the a) low slope band, b) medium slope band, and c) high slope band.

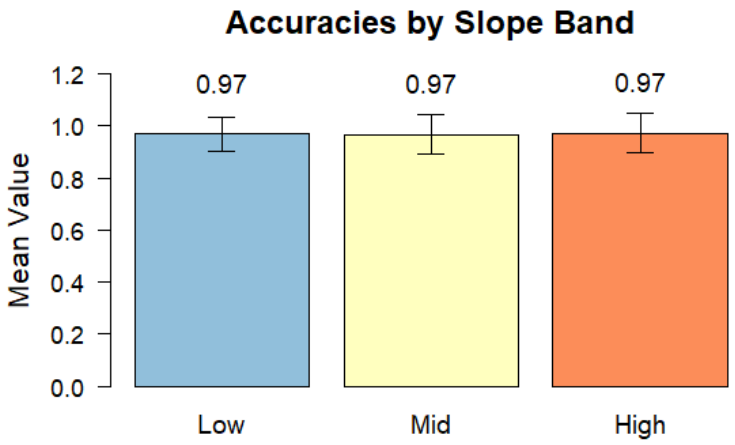


Figure 2.25: Mean accuracies by slope band. Standard deviations represented by the error bars for low, medium, and high slopes are 0.07, 0.08, and 0.08, respectively.

2.3.3.3 Aspect

Performance of STARFM across hillslope aspect was also consistent (Figure 2.26). Snow cover was correctly predicted for north aspects with a mean accuracy of 0.97 and a standard deviation of 0.08, for east aspects with a mean accuracy of 0.97 and a standard deviation of 0.07, for south aspects with a mean accuracy of 0.97 and a standard deviation of 0.07, and for west aspects with a mean accuracy of 0.97 and a standard deviation of 0.07 (Figure 2.27).

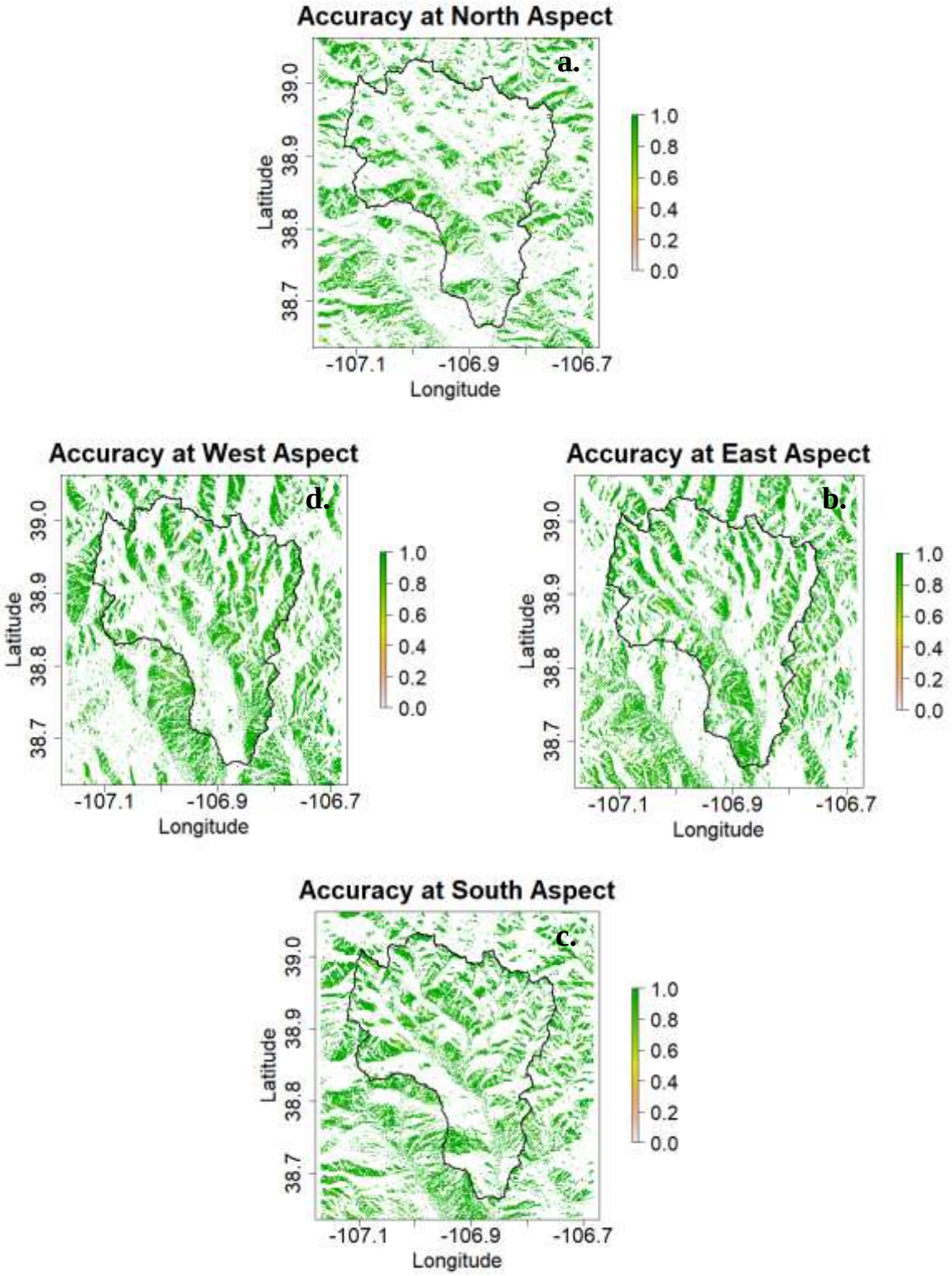


Figure 2.26: Spatial distribution of accuracy values by pixel for the a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

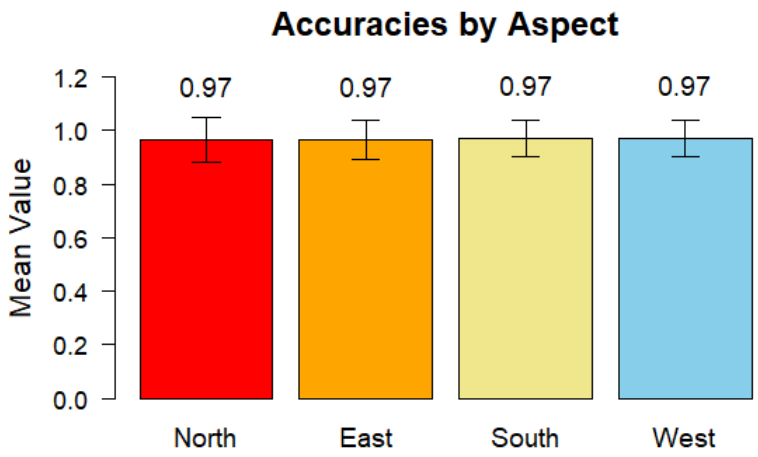


Figure 2.27: Mean accuracies by aspect direction. Standard deviations represented by the error bars for north, east, south, and west aspects are 0.08, 0.07, 0.07, and 0.07 respectively.

We also wanted to examine whether our cloud mask might be treating hillslopes differently, and thus affecting data availability, depending on aspect. Since snow tends to persist longer on north and east aspects (Anderson et al., 2014), it is possible those pixels were classified as clouds and masked out during the data pre-processing stage more frequently, leading to fewer numbers of predictions in those areas. This trend was visible in our data, with north aspects having the least number of model predictions (Figures 2.28 and 2.29). Numerical distribution of predictions for north aspects had a Q1 of 4, a median of 7, and a Q3 of 12. Pixels on east and west aspects had a similar distribution of numbers of observations, with east facing pixels having a Q1 of 6, a median of 10, and a Q3 of 14, and west facing pixels having a Q1 of 6, a median of 10, and a Q3 of 15. South facing pixels had the highest numbers of observations, with their distribution having a Q1 of 6, a median of 10, and a Q3 of 15.

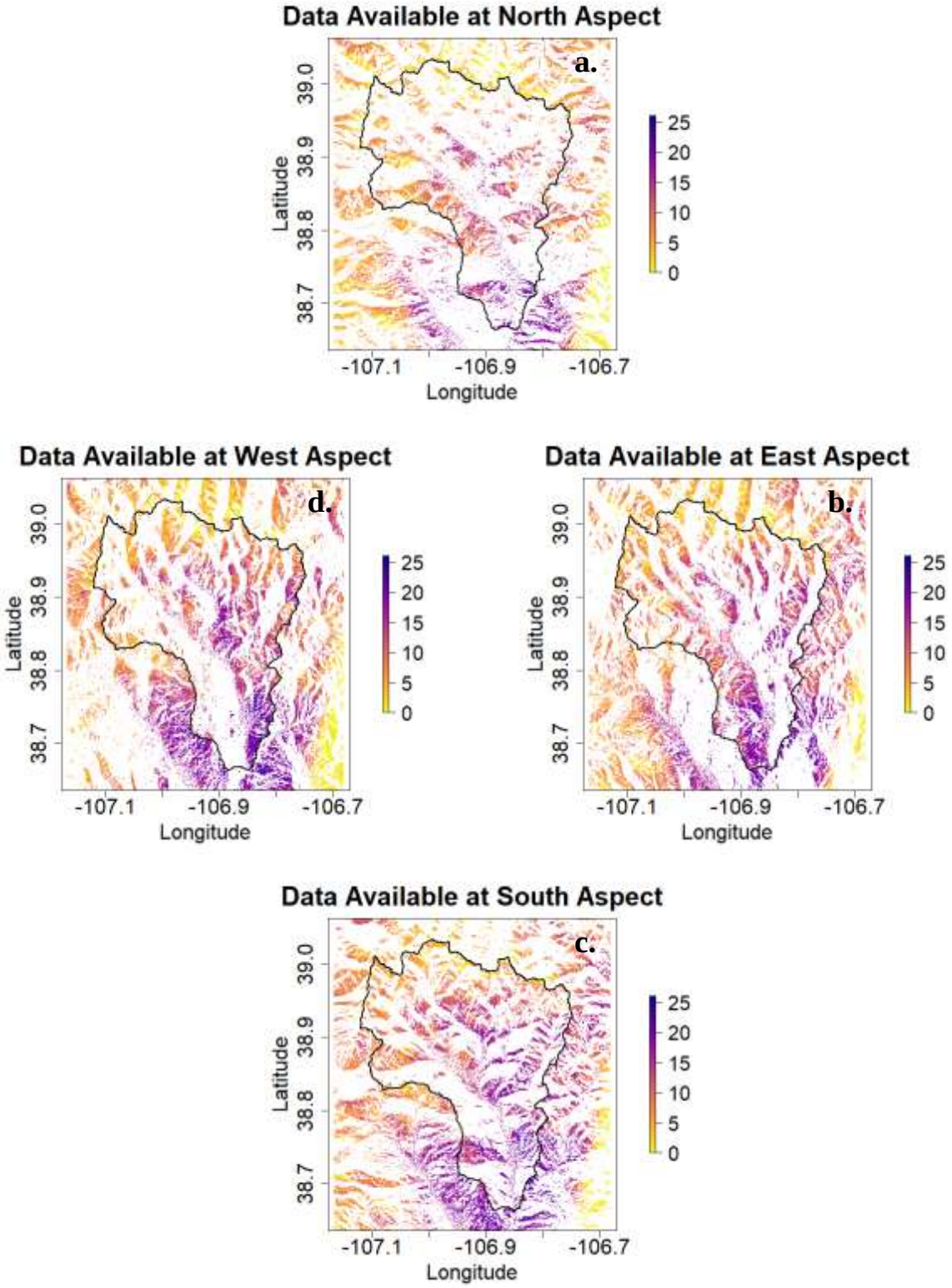


Figure 2.28: Spatial distribution of data availability per pixel for the a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

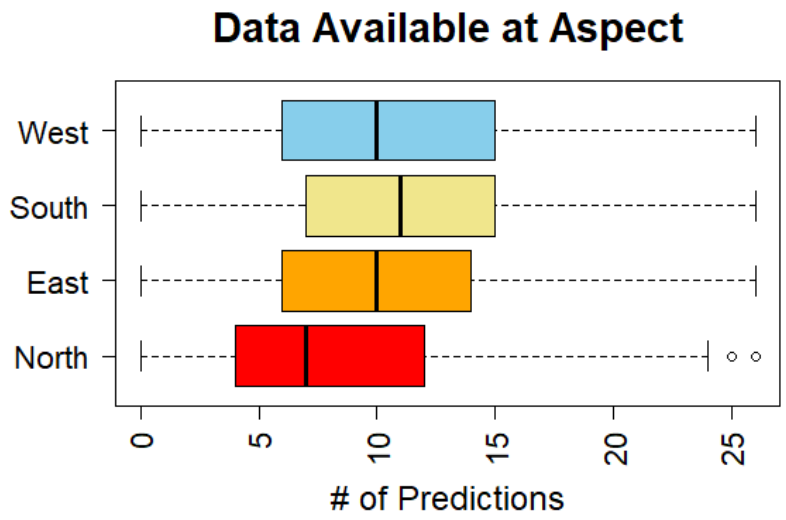


Figure 2.29: Distribution of the number of predictions per pixel, regardless of correctness, for each aspect direction.

2.3.3.4 Land Cover Type

STARFM performed well regardless of land cover class (Figure 2.30). Similar to the results from our other landscape characteristics, there appeared to be no relationship between model accuracy and data availability. In fact, the largest land cover class, trees, had the lowest mean accuracy and largest standard deviation values at 0.96 and 0.08, respectively. The next largest land cover class of grass had a mean accuracy of 0.98 and a standard deviation of 0.04. The shrubs class had a mean accuracy of 0.98 and a standard deviation of 0.05, clear/unvegetated areas had a mean accuracy of 0.99 and standard deviation of 0.06, and the class encompassing all other land cover classes had a mean accuracy of 0.97 and standard deviation of 0.08 (Figure 2.31).

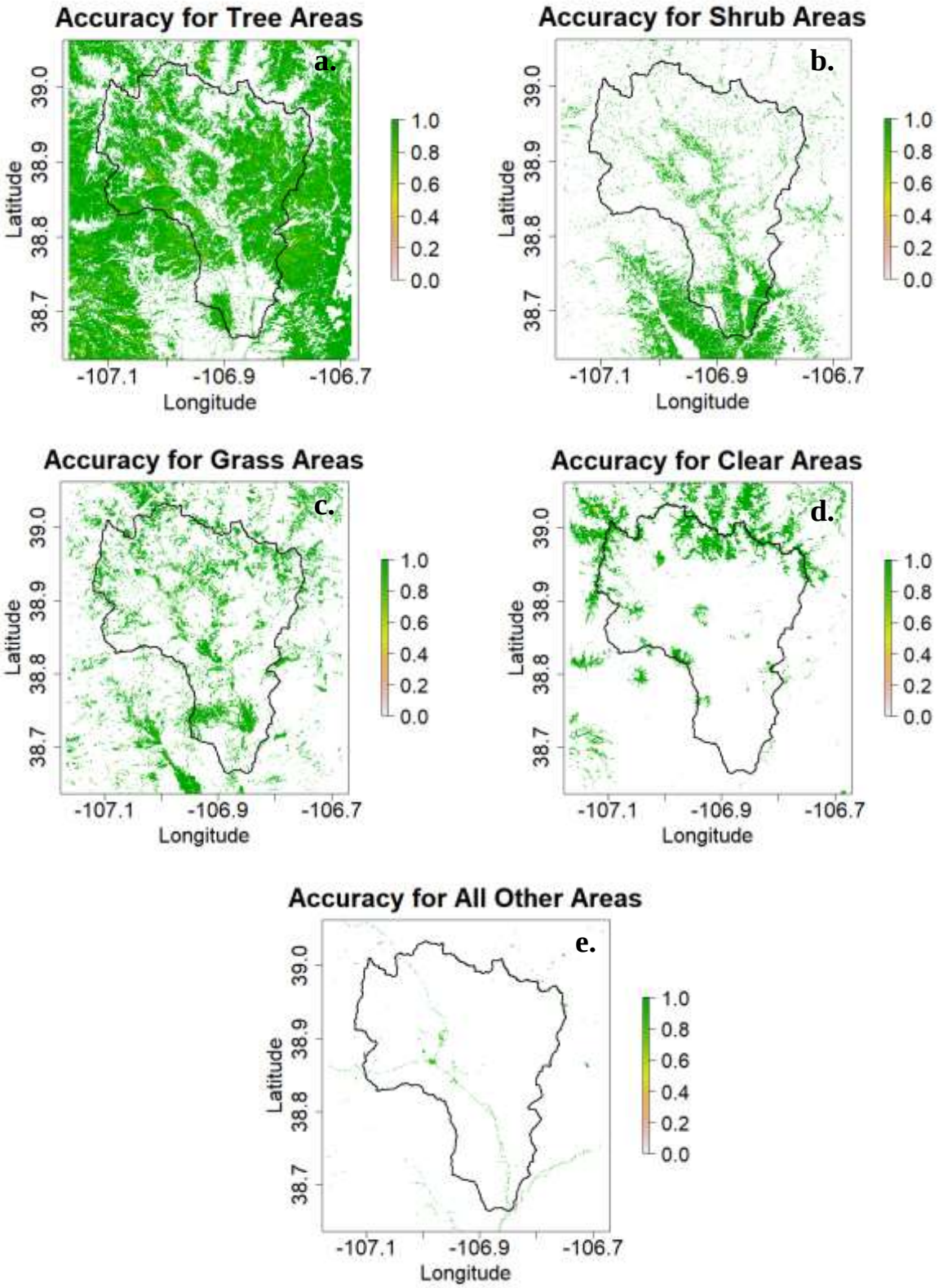


Figure 2.30: Spatial distribution of accuracy values by pixel for the a) tree areas, b) shrub areas, c) grass areas, d) clear areas, e) all other areas.

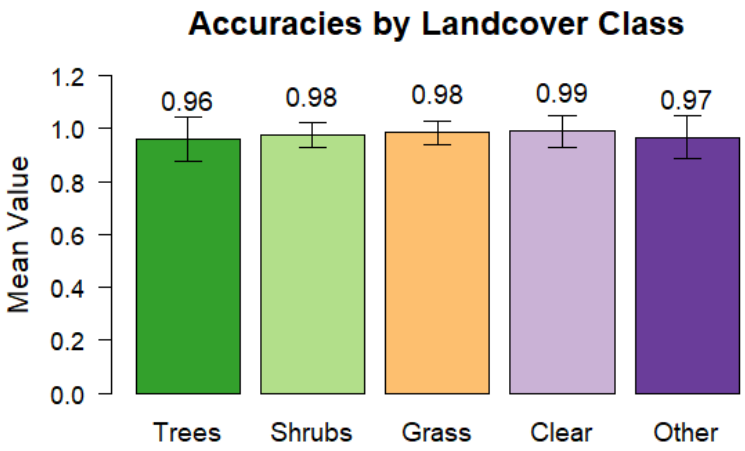


Figure 2.31: Mean accuracies by landcover class. Standard deviations represented by the error bars for tree, shrub, grass, clear, and all other areas are 0.08, 0.05, 0.04, 0.06, and 0.08, respectively.

2.4 Discussion

In this study we found that STARFM is effective at tracking snow-cover in a heterogeneous, mountain area regardless of time of year or landscape feature and is subject mainly to limitations related to data gaps and cloud cover. These findings have the potential to improve the resolution of snow data inputs to hydrologic models, including those that seek to better understand water related processes in the CZ.

The limitations of using surface reflectance data with minimal pre-processing to identify and monitor SCA have been well documented and will be discussed in the following paragraphs. While we were aware of many of these limitations throughout the course of this study, we chose not to do extensive pre-processing with our data, as that was beyond the scope of our goal of this study. Due to the fact that STARFM has never before been used to monitor SCA in mountain landscape, our primary goal was to assess whether or not the model was capable of being used in this manner with the understanding that the methods may need to be further tailored to other datasets and study

areas in future work. In other words, our intention was to provide a preliminary study of an as-is, "off-the-shelf", method with minimal pre-processing as a proof of concept for using STARFM to monitor SCA. With this perspective and the goals of this study in mind, we chose to calculate our own NDSI values using surface reflectance data rather than using data from other available snow cover products, such as the MODIS daily snow cover product, MODA110. Surface reflectance is a more "raw" form of data, and thus allowed us greater flexibility to choose our own data filtering methods consistent with the level of pre-processing we desired.

A significant limitation we found with this choice of data, however, was the large data gaps in our Landsat and MODIS input files resulting from overly robust cloud cover and cloud shadow filters. Cloud masking algorithms informed by QA pixel flags often have difficulties differentiating between snow and clouds, especially at high elevation areas. This can occur due to brightness or temperature differentials between the surface and clouds in these areas being too small for the cloud to be distinguished from snow (USGS, 2020). This is likely the reason why the number of per-pixel observations decreased with increasing elevation. While it is also likely that there were more cloud-covered days with increasing elevation (Hou et al., 2019), visual inspections of true-color images revealed that areas masked out as clouds were often actually snow-covered areas. In addition, when we look at Figure 2.12 we can see that locations exist in the study area where no model predictions were made over the entire year of analysis, many of which are located at the highest elevations at the north end of the watershed. The QA pixel flags also had trouble at times with areas in mountain shadows, flagging these areas as cloud shadows, or at the very least as not "clear" pixels, and thus masking them out.

Other data products have attempted to correct for misclassification errors in cloud masks with additional pre-processing. The MODIS Snow-Covered Area and Grain Size (MODSCAG) (Painter et al., 2009) and Thematic Mapper Snow-Covered Area and Grain Size (TMSCAG) (Selkowitz et al., 2017) datasets have been developed to address these issues. MODSCAG is created by applying a spectral mixture analysis to MODIS surface reflectance data (MOD09GA, the same MODIS dataset we use here in our study). A library of spectral endmembers for snow at various grain sizes, rock, soil, vegetation, and lake ice are used to estimate *f*SCA and snow albedo in the image. In addition to this more sophisticated method of identifying snow cover, MODSCAG data is processed with its own cloud mask that considers particle size. Above a certain threshold value, a particle is considered a cloud, even if the algorithm originally classifies as snow, and under the threshold value, a particle is considered snow, even if it is originally classified as cloud. The TMSCAG model is similar to MODSCAG, with the main difference being that the former is designed to handle radiometric saturation in bands 1-4 (Selkowitz et al., 2017). As STARFM is not able to make predictions for pixels on dates where input information is missing, future work with STARFM where a more selective cloud filtering methodology is preferred will benefit from using datasets such as those described above as input files to the model.

Finally, we must also consider the assumptions that are made and the limitations that result from validating STARFM's performance with Landsat data. Withholding Landsat data from STARFM and using it as "ground-truth" to evaluate model error is a method commonly used with STARFM and similar models (Hilker et al., 2009; Gevaert & Garcío-Haro, 2014; Wang et al., 2014), as the goal of the STARFM model is to

accurately replicate Landsat-like data (Gao et al., 2006). While we can be confident in STARFM's ability to replicate Landsat data for our study area as shown by our results, we did not validate STARFM's snow cover predictions against in situ snow cover data. We operate under the assumption that our algorithm, which identified snow cover by NDSI values, accurately located snow in the landscape for both STARFM output and our Landsat and MODIS input data. We are aware of instances in which binary identification has been shown to misidentify snow cover within both Landsat and MODIS data (Elder et al., 1998; Rittger et al., 2013), and thus we suspect that these same errors are present in our STARFM data, although we did not explicitly identify them in this study.

The incorporation of ancillary ASO data as part of the model validation process, however, does provide some insight into possible sources of error regarding STARFM data and the location of actual snow cover on the ground. Some specific potential sources of error are examined in greater detail below in Section 2.4.2.1. We also recognize that the inclusion of ASO data in our study provides us with only one data point by which to draw these conclusions. In addition, this single data point was collected as close as possible to the time of peak SWE, the time when snow cover is most homogeneous and remote sensing methods have the least difficulty in identifying SCA.

Future work quantifying STARFM errors can benefit from the use of lidar data acquired over multiple years, and if available, at different dates throughout the snow season. Other satellite remote sensing data sources with finer spatial and temporal resolutions, such as Sentinel-2 data with its 5-day revisit time and spatial resolution as fine as 10-m at some bands, can also be used as validation information to increase confidence in STARFM's predictions of snow cover.

2.4.1 Sensitivity Analysis

The results of our sensitivity analysis indicated that in snow-covered environments, STARFM was able to more accurately replicate the heterogeneous nature of the landscape when a larger model domain and a smaller search window is used. Our search window size of 15-m was significantly smaller than the search window sizes used in most other STARFM studies, which commonly used distances of 750-1500 m (Gao et al., 2006; Gevaert & García-Haro, 2014; Gallagher, 2018). Few studies have completed sensitivity analyses to optimize STARFM performance over highly heterogeneous landscapes, but those that have optimized model performance by increasing search window sizes, even though it resulted in increased computational costs (Gao et al., 2006; Gevaert & García-Haro, 2014). We suspect that our model's optimal performance with a smaller window size may be due to the large spectral contrasts between snow covered and non-snow covered pixels, which allows the model to easily differentiate between these areas without the need for increased information from greater search distances. This theory, however, seems to be in direct contrast with the better model performance we achieved by enlarging the model domain. Future work optimizing STARFM performance in a variety of snow-covered, mountainous landscapes may provide a better understanding of this inconsistency.

2.4.2 Temporal Analysis

STARFM produced high accuracy values in the winter and summer when the landscape was dominated by either "snow" or "no-snow" classes, respectively. During these time frames the model had little difficulty assigning NDSI values into the correct class. The 5-day centered moving average of accuracy (*a*) values displayed a slight

decrease in accuracy during the spring months, with the lowest accuracy values occurring in the fall months. We expect to see a decrease in model a during times of the year when landscape heterogeneity increases, such as during these shoulder seasons.

There are a few key differences in snow properties during the fall and spring that may contribute to differences in model performance during these times. Early snow precipitation that falls before the snowpack is established can be quickly melted away by sudden increases in temperature. New snow has lower densities and is much more easily blown around the landscape (Hiemstra et al., 2006). As snow persists longer in the landscape and accumulates more layers, field observations have shown that the entire snowpack becomes more dense overall, with the greatest densities in the middle of the snowpack as one moves away from the surface (Skiles & Painter, 2017). Therefore, we can reasonably assume that even if the top layers of new, less dense snow are blown around by late season storms, the rest of the snowpack will remain largely stationary on the landscape, resulting in more consistent snow patterns between observations in the spring than in the fall. These seasonal differences are important as quicker, smaller scale landscape changes are more difficult for STARFM to catch. Although increasing the spatial and temporal resolution is the goal of STARFM, the model is still limited by subtle changes in the landscape that occur between Landsat images but are unable to be seen by the coarse scale of the MODIS imagery. However, the lowest accuracy value of 0.72 indicated that the model is still able to perform reasonably well even with the above limitations.

Precision (p), recall (r), and F-score (F) values all follow similar trends throughout the year. The high recall values achieved in the winter months indicate that

STARFM was able to recognize and identify the majority of snow present in the landscape and assign NDSI values accordingly. The high precision values of this same time period indicated that the majority of the pixels to which STARFM is assigning 0.4 and above NDSI values are indeed snow covered. Similar to what we see in our accuracy plots, 5-day centered moving averages of precision and recall values decreased for the shoulder seasons of the year, and again were lower in the fall than in the spring. Precision values were especially low for the fall, indicating that the model was greatly overpredicting snow cover during these times. As the harmonic mean of precision and recall, F-score values are similar to precision and recall values for the spring, summer, and winter, but display the large decrease in value during the fall, as seen with precision.

Other reasons why STARFM performance was worse in the fall as opposed to the spring, especially in terms of overpredicting snow, may be due to differences in landscape characteristics unrelated to snow. Because the model was overpredicting snow, we know that the STARFM synthetic data had a higher NDSI than the "ground-truth" Landsat data. One possible explanation may be that our study area is 60% tree covered, many of which are deciduous aspen as identified by Carroll et al. (2018). The colors of these leaves change and become brighter in the fall, which may in turn artificially raise the NDSI values of the input data. This potential for misclassification from higher NDSI would disproportionately affect the coarser MODIS data, as the NDSI of pixels covering a larger area will be influenced by reflectance values from both higher reflectance tree and lower reflectance non-tree areas. Landsat pixel NDSI values would similarly be affected, but the total area of affected pixels would be reduced since smaller Landsat pixels are less likely to cover both tree and non-tree areas. Incorporating the Normalized

Difference Vegetation Index (NDVI), a spectral band ratio used for identifying vegetation health and sometimes incorporated along with NDSI to identify snow cover in forest areas (Klein et al., 1998), may improve STARFM results in the fall season.

Unlike accuracy, precision, recall, and F-score moving averages were at their lowest in the summer. Because precision and recall metrics are specific to the presence of our positive class of "snow", they can experience large changes in value when the number of positive class pixels are low. In this case, a misclassification of even a small amount of positive class pixels will drastically increase the value of the denominator in Equations 2.4 and 2.5, which in turn decreases the *p* and *r* values. As the harmonic mean of *p* and *r*, *F* values (Equation 2.6) are also low in the summer for the same reasons.

2.4.2.1 ASO Data

High accuracy, precision, recall, and F-score were also observed when comparing overlapping STARFM and ASO lidar data. Although the values of these metrics are slightly lower than those computed from comparisons with Landsat data, the accuracy value of 0.89 and precision, recall, and F-score values at 0.90 or above indicate that STARFM does provide information on snow cover that is comparable with snow cover information acquired from lidar acquisition. We recognize that this is only one data point, as only one ASO flight occurred during our model validation time frame. STARFM is clearly not a substitute for airborne lidar data when it comes to monitoring SCA as it is limited by cloud cover where airborne lidar data is not. Although airborne lidar data is widely recognized as one of the most accurate sources of information on snow depth, and therefore snow presence (Nolin, 2010; Painter et al., 2016), it has a high operational cost and is therefore temporally limited. It is encouraging then that STARFM optical data,

when available, can replicate airborne lidar SCA observations with approximately 90% accuracy.

Satellite optical remote sensing data are limited compared to airborne lidar snow data, however, when identifying snow cover in forested areas. While airborne lidar data can detect and identify snow under forest canopy, surface reflectance datasets from satellites cannot. This means that a forested area can be falsely identified as having no snow when snow is present but hidden from the overhead view. Selkowitz et al. (2017) addressed this problem by designing an algorithm that uses a local window search to further evaluate Landsat pixels in forested areas that are identified by the TMSCAG model as non-snow-covered but are acquired during times when snow cover is likely. These algorithms take ancillary data on land cover type, forest canopy cover, digital elevation models, and incoming solar radiation and identify thresholds that must be met to determine if the pixel in question should be adjusted to an *f*SCA value greater than zero.

Conversely, it is possible that STARFM may be identifying forested areas as snow-covered due to high NDSI values resulting from intercepted snow on tree canopies, when in fact there may be no snow on the ground underneath them. This is likely not a noteworthy source of error in our analysis of STARFM and ASO data due to our date of comparison being at or near peak SWE. This may be a more significant source of error, however, during other times of the year when snow has not yet accumulated beneath forest canopy but is instead stored within the canopy. Snowfall that has been intercepted by forest canopy can often sublimate before it has the chance to reach the ground. Broxton et al. (2015) found that this behavior can happen with up to 25% of snowfall in

mid-latitude, mountainous environments. Misidentifying this snow in the forest canopy as snow on the ground should be considered as a possible source of error in overestimating water inputs to hydrologic and CZ models.

We did not attempt to correct our results based on forest canopy density, nor did we incorporate measurements of NDVI into our binary snow cover classification for tree-covered areas (Klein et al., 1998) for the sake of the simplicity of using consistent model inputs over our entire domain. Thus, we suspect limitations in the ability of our satellite data to detect snow on the ground due to canopy cover are the primary cause of the deviation of SCA derived from STARFM and SCA derived from ASO. In addition, Lundquist et al. (2013) found that forests in climates with colder, drier winters, such as the East River watershed, retain snow longer under tree canopies than in open areas during the spring. This is largely due to the greater influence of shortwave radiation in inducing melting in areas with colder winter temperatures rather than emitted longwave radiation from trees and other vegetative material, which is more dominant in melt regimes in climates with warmer winter temperatures. Addressing this optical data blind spot in future work would greatly improve the predictive ability of STARFM in forest areas, which is the majority land cover type of the East River watershed, especially during the spring melt season.

2.4.2.2 Random Raster Test

The random raster tests we conducted were intended to determine whether or not our model was better at correctly identifying snow in the landscape than an algorithm that produced a random spatial distribution of snow, with the underlying assumption that the location of snow cover in a landscape is non-random. When we compared the accuracy,

precision, recall, and F-score from our temporal analysis of STARFM to the same metrics from our random raster tests, we saw a clear gap in performance between the two, with the model outperforming the random raster tests in its placement of snow throughout the landscape in all but a handful of instances. This indicates that, for the majority of the time, our model was able to capture the spatial variability of snow in the landscape related to non-random snow distribution processes.

On two occasions, random pixel placement recall values were slightly above those achieved by STARFM, indicating that in these few cases, the random algorithm provided better information regarding snow cover than STARFM. We believe both of those instances can be attributed to skewing from a lack of data due to cloud cover and due to the February-March 2016 MODIS outage.

The gap between modelled and randomly assigned precision, recall, and F-score values was largest in spring and fall, the time when both algorithms were expected to have the most difficulty in identifying snow placement on the landscape due to increased heterogeneity. While performance dropped for both algorithms during these times, the fact that STARFM significantly outperformed the random raster tests, even during the fall when STARFM was underpredicting snow cover, is again an indicator that the model was able to provide more information regarding snow placement in a heterogeneous landscape than the algorithm placing snow at random.

We saw a similar relationship when considering accuracy values for the two methods. Again, the largest gaps in accuracy values between STARFM and the random raster test occurred in spring and fall. One major difference, however, was that during the middle of winter accuracy was consistently much higher for STARFM, a relationship

which was not as strong for our other metrics. We once again attribute this improved performance to the way the metrics were calculated. By only considering the positive class – "snow" - precision, recall, and F-score are much less sensitive to small changes when there are a very large number of positive class pixels. Accuracy, by contrast, considers the correct instances of identification of both classes equally, and so can be a better metric to consider when the number of positive class instances is high. The opposite is true, meaning accuracy is less useful, when the positive class – "snow" - is nearly non-existent in the summer months. During this time, accuracy values for both methods were equally high at 100%. This suggests that when the number of positive pixels is extremely small, they are largely insignificant in the calculation of accuracy.

2.4.3 Spatial Analysis

While we expected to see clear relationships between landscape features and model performance in a heterogeneous landscape, we found surprisingly little relationship between accuracy and differences in elevation, slope, aspect, and land cover type. The highest mean accuracy value achieved by STARFM was for clear/unvegetated areas at 0.99, and the lowest was for mid-elevation areas at 0.96, a difference of only 0.03. We expected to see lower STARFM performance for areas of the landscape with greater heterogeneity, such as mid-elevations where the rain-snow transition zone fluctuates throughout the year (Klos et al., 2014) influencing the snowline, or on north-facing slopes with longer spring snowpack retention (Anderson et al., 2014). It appears, however, that STARFM does not have difficulties identifying these heterogeneities in the landscape at the 30-m spatial scale and can do so just as well as the Landsat data used for ground-truthing.

While mean accuracies are consistent, some landscape characteristics display slightly more variation, as indicated by larger standard deviations. The largest standard deviation value was calculated at 0.08 for tree areas and the lowest at 0.05 for grass areas, resulting in a difference of 0.03. Landscape characteristics that had higher standard deviation values included mid- and high elevation areas, north aspects, and land cover areas that fall into the tree and "other" land cover categories. While the standard deviation values were still quite small, some of the variability can likely be explained by low data availability resulting from cloud-cover, tree canopy obstruction, and cloud misidentification at high elevations. It is reasonable to expect larger variability in smaller datasets, as outliers will have more of an effect on their overall statistics. We suspect this is what influenced the standard deviations of all categories except for the tree areas. The tree area dataset was the largest dataset, covering 60% of the watershed. These data, however, are also fairly evenly distributed across all elevations, slopes, and aspects, which may explain why these data also have larger variabilities when considered on their own. Grass areas, the dataset with the smallest variability, are similarly distributed throughout the watershed, however, they cover much less area at only 15%. Differences in standard deviations between grass and tree areas may be due to inconsistencies in the model identifying snow among tree canopies, as explained in detail in Section 2.4.2.1.

Similar to the patterns in our data revealed by the differences in standard deviations, we saw similar patterns attributed to systematic relationships of data availability by obstructions from cloud-cover, tree canopies, and cloud misclassifications. The median values of the number of per-pixel predictions for mid- and high elevations and north aspects, the datasets with the smallest number of predictions due to the above

factors, are well below the midpoints for the ranges of these data. Larger variations in data availability for tree and "other" land cover areas can likely be explained by considering the amount of the watershed these land cover types occupy. Tree areas cover the largest amount of area at 60% of the watershed, and areas classified as "Other" cover only 1% of the watershed. The distribution of data availability of tree areas is likely due to their distribution on the landscape as previously mentioned. Conversely, "other" areas of the watershed have a high median number of predictions per pixel, but the number of data points is extremely small and contains vastly different types of land cover (e.g. developed areas and those covered permanently by water), therefore large variations in data availability in these locations is also not surprising.

Given our high mean accuracies and the fact that even our highest standard deviation values are still quite small, the slight discrepancies in model performance by landscape feature can be largely explained by data availability as opposed to errors from the model itself. This suggests that STARFM is viable at successfully replicating snow cover in complex, heterogeneous mountain landscapes. Future work with STARFM employing more detailed data pre-processing of model inputs may be able to test these conclusions through verification at higher resolutions.

2.5 Conclusion

We describe here a comprehensive method for validating STARFM for use in complex, heterogeneous mountain terrain to track seasonal snow cover. We identifiy the optimal model parameters and model domain for our study area. STARFM performance is validated using a data denial method to exclude individual Landsat images from model input files for comparison of model output to the real data acquired at that date. Binary

snow maps were created for the Landsat and STARFM data on the dates in consideration using an NDSI threshold. A confusion matrix derived through a classification analysis was employed to find accuracy, precision, recall, and F-score over the course of the year of analysis. Additional validation was provided by evaluating model results with ASO data and SCA maps created by random pixel placement. In addition to a temporal performance validation, STARFM was also validated spatially by considering its performance in relation to landscape features including elevation, slope, aspect, and land cover type.

Our results indicate that STARFM was able to successfully replicate Landsat resolution SCA data over a heterogeneous mountain area for an entire year, and that it performed equally well when given data including a variety of landscape features and land cover types. The primary limitation of STARFM in this study was data availability. When STARFM was given data, it performed well, but characteristics that make all remote sensing analyses challenging in mountain watersheds, including clouds, shadows, and data viewing angles, can limit its effectiveness. We recommend future studies using STARFM in mountain areas be aware of these limitations and attempt to mitigate them by utilizing more selective cloud masks or input datasets that contain their own cloud identification algorithms validated for mountain areas.

References

Anderson, B. T., McNamara, J. P., Marshall, H., & Flores, A. N. (2014). Insights into the physical processes controlling correlations between snow distribution and terrain properties. *Water Resources Research, 50*, 4545-4563. doi:10.1002/2013WR013714

Anderton, S. P., White, S. M., & Alvera, B. (2004). Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes, 18*, 435-453. doi:10.1002/hyp.1319

Andreadis, K. M., & Lettenmaier, D. P. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources, 29*, 872-886. doi:10.1016/j.advwatres.2005.08.004

Berman, E. E., Bolton, D. K., Coops, N. C., Mityok, Z. K., Stenhouse, G. B., & Moore, R. D. (2018). Daily estimates of Landsat fractional snow cover driven by MODIS and dynamic time-warping. *Remote Sensing of Environment, 216*, 635–646. doi:10.1016/j.rse.2018.07.029

Boz, M. B., Calvert, K., Brownson, J. R. S. (2015). An automated model for rooftop PV systems assessment in ArcGIS using LIDAR. *AIMS Energy 3*(3), 401-420. doi:10.3934/energy.2015.3.401

Broxton, P. D., van Leeuwen, W. J. D., Biederman, J. A. (2019). Improving Snow Water Equivalent Maps with Machine Learning of Snow Survey and Lidar Measurements. *American Geophysical Union,* 27 pp. doi:10.1029/2018WR024146

Buhler, Y., Adams, M. S., Bosch, R., & Stoffel, A. (2016). Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations. *The Cryosphere, 10,* 1075-1088. doi:10.5194/tc-10-1075-2016

Cammalleri, C., Anderson, M.C., Gao, F., Hain, C.R., & Kustas, WP. (2014). Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agricultural and Forest Meteorology, 186*, 1-11. doi:10.1016/j.agrformet.2013.11.001

Carroll, R. W. H., Bearup, L. A., Brown, W., Dong, W., Bill, M., & Williams, K. H. (2018). "Factors controlling seasonal groundwater and solute flux from snow-dominated basins". *Hydrological Processes, 32*, 2187-2202. https://doi.org/10.1002/hyp.13151

Che, T., Li, X., Jin, R., Armstrong, R. & Zhang, F. (2008). Snow depth derived from passive microwave remote-sensing data in China. *Annals of Glaciology 29,* 145-154. doi:10.3189/172756408787814690

Daly, S. F., Vuyovich, C. M., Deeb, E. J., Newman, S. D., Baldwin, T. B., & Gagnon, J. J. (2012). Assessment of the snow conditions in the major watersheds of Afghanistan using multispectral and passive microwave remote sensing. *Hydrological Presses, 26,* 2631-2642. doi:10.1002/hyp.9367

Elder, K., Rosenthal, W., & Davis, R. E. (1998). Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes, 12,* 1793-1808.

Emelyanova, I. V., McVicar, T. R., Van Niel, T. G., Li, L. T., & van Dijk, A. I. J. M. (2012). On blending Landsat-MODIS surface reflectances in two landscapes with contrasting spectral, spatial and temporal dynamics. WIRADA Project 3.4: Technical Report. CSIRO Water for a Healthy Country Flagship, Australia. 72 pp.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics, 45,* RG2004, 1-33. doi:10.1029/2005RG000183

Franz, K. J., & Karsten, L. R. (2013). Calibration of a distributed snow model using MODIS snow covered area data. *Journal of Hydrology, 494,* 160-175. dx.doi:/10.1016/j.jhydrol.2013.04.026

Gallagher, M. (2018). "Utilizing satellite fusion methods to assess vegetation phenology in a semi-arid ecosystem." [Boise, Idaho] : Boise State University, 2018. BSU Albertsons Library ebook -- Online LD458.8.G5 G355 2018eb http://catalog.boisestate.edu/cgi-bin/Pwebrecon.cgi?LANGUAGE=English&DB=local&BBID=1295025

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(8), 2207–2218. doi:10.1109/TGRS.2006.872081

Gaskill, D. L., Mutschler, F. E., Kramer, J. H., Thomas, J. A., & Zahony, S. G. (1991). Geologic map of the Gothic Quadrangle. Colorado: Gunnison County.

Gevaert, C. M. & Garcia-Haro, F. J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment, 156,* 34-44. dx.doi:10.1016/j.rse.2014.09.012

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202,* 18-27. dx.doi:10.1016.j.rse.2017.06.031

Hall, D. K., Riggs, G. A., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment, 54,* 127–140. doi:10.1016/0034-4257(95)00137-P

Hiemstra, C. A., Liston, G. E., & Reiners, W. A. (2006). Observing, modelling, and validating snow redistribution by wind in a Wyoming upper treeline landscape. *Ecological Modelling, 197,* 35–51. doi:10.1016/j.ecolmodel.2006.03.005

Hilker, T., Wulder, M. A., Coops, N. C., Seitz, N., White, J. C., Gao, F., & Stenhouse, G. (2009). Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing of Environment, 113,* 1988–1999. doi:10.1016/j.rse.2009.05.011

Homan, J. W., Luce, C. H., McNamara, J. P., & Glenn, N. F. (2011). Improvement of distributed snowmelt energy balance modeling with MODIS-based NDSI-derived fractional snow-covered area data. *Hydrological Processes, 25*, 650-660. doi: 10.1002/hyp.7857

Horn, B. K. P. (1981). Hill Shading and Reflectance Map. *Proceedings of the IEEE, 69*(1), 14-47.

Hou, J., Huang, C., Zhang, Y., Guo, J., & Gu, J. (2019). Gap-Filling of MODIS Fractional Snow Cover Products via Non-Local Spatio-Temporal Filtering Based on Machine Learning Techniques. *Remote Sensing*, *11*(90), 1-24. doi:10.3390/rs11010090

Kerr, T., Clark, M., Hendrikx, J., & Anderson, B. (2013). Snow distribution in a steep mid-latitude alpine catchment. *Advances in Water Resources, 55*, 17-24. dx.doi:10.1016/j.advwatres.2012.12.010

Klein, A. G., Hall, D. K., & Riggs, G. A. (1998). Improving snow cover mapping in forests through the use of a canopy reflectance model. *Hydrological Processes, 12*, 1723-1744. doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1723::AID-HYP691>3.0. CO;2-2

Klos, P. Z., Link, T. E., & Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate. *Geophysical Research Letters*, *41*, 4560–4568. doi:10.1002/2014GL060500

Landfire. (2016), Existing vegetation type layer, LANDFIRE 2.0.0, U.S. Department of the Interior, Geological Survey. Accessed 30 September 2020

Lundquist, J. D., Dickerson-Lange, S. E., Lutz, J. A., & Cristea, N. C. (2013). Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling. *Water Resources Research, 49,* 6356-6370. doi:10.1002/wrcr.20504

Margulis, S. A., Liu, Y., & Baldo, E. (2019). A Joint Landsat- and MODIS-Based Reanalysis Approach for Midlatitude Montane Seasonal Snow Characterization. *Frontiers in Earth Science, 7*(272). doi:10.3389/feart.2019.00272

Maurer, E. P., Rhoads, J. D., Dubayah, R. O., & Lettenmair, D. P. (2003). Evaluation of the snow-covered area product from MODIS. *Hydrological Processes, 17,* 59-71. doi: 10.1002/hyp.1193

McClung, D., & Schaerer, P. (2006). The Avalanche Handbook. 3rd ed. Seattle: The Mountaineers.

Nolin, A. W. (2010). Recent advances in remote sensing of seasonal snow. *Journal of Glaciology, 56*(200), 1141-1150.

Olsoy, P. J., Mitchell, J., Glenn, N. F., & Flores, A. N. (2017). Assessing a multi-platform data fusion technique in capturing spatiotemporal dynamics of heterogeneous dryland ecosystems in topographically complex terrain. *Remote Sensing, 9* (981), 19. dx.doi:10.3390/rs9100981

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment, 113,* 868-879. doi:10.1016/j.rse.2009.01.001

Painter, T.H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C., & Winstral, A. (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment, 184*, 139-152. dx.doi:10.1016/j/rse.2016.06.018

Painter, T. (2018). ASO L4 Lidar Snow Depth 50m UTM Grid, Version 1. [2016-04-04]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Archive Center. doi:10.5067/STOT5I0U1WVI Accessed July 16, 2020.

Peters, F. (2016). "Analyzing the phenologic dynamics of kudzu (Pueraria montana) infestations using remote sensing and the normalized difference vegetation index." (2016). Electronic Theses and Dissertations. Paper 2437. doi:10.18297/etd/2437

Rittger, K., Painter, T. H., & Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources, 51,* 367–380. doi:10.1016/j.advwatres.2012.03.002

Roy, A., Royer, A., & Turcotte, R. (2010). Improvement of springtime streamflow simulations in a boreal environment by incorporating snow-covered area derived from remote sensing data. *Journal of Hydrology, 390,* 35-44. doi:10.1016/j.jhydrol.2010.06.027

Selkowitz, D. J., Painter, T. H., Rittger, K., Schmidt, G., & Forster, R. (2017). The USGS Landsat Snow Covered Area Products: Methods and Preliminary Validation CHAPTER 5. In Automated Approaches for Snow and Ice Cover Monitoring Using Optical Remote Sensing (pp. 76). The University of Utah.

Singh, D. (2011). Generation and evaluation of gross primary productivity using Landsat data through blending with MODIS data. *International Journal of Applied Earth Observation and Geoinformation, 13,* 59-69. doi:10.1016/j.jag.2010.06.007

Skiles, S. M., & Painter, T. (2017). Daily evolution in dust and black carbon content, snow grain size, and snow albedo during snowmelt, Rocky Mountains, Colorado. *Journal of Glaciology, 63*(237), 118-132. doi:10.1017/jog.2016.125

Tan, Z., Yue, P., Di, L., & Tang, J. (2018). Deriving High Spatiotemporal Remote Sensing Images Using Deep Convolutional Network. *Remote Sensing, 10*(1066), 1-16. doi:10.3390/rs10071066

Tennant, C. J., Crosby, B. T., & Godsey, S. E. (2015). Elevation-dependent responses of streamflow to climate warming. *Hydrological Processes, 29,* 991-1001. doi:10.1002/hyp.10203

Tong, R., Parajka, J., Komma, J., & Blöschl, G. (2020). Mapping snow cover from daily Collection 6 MODIS products over Austria. *Journal of Hydrology, 590*(125548), 1-10. doi:10.1016/j.jhydrol.2020.125548

Veatch, W., Brooks, P. D., Gustafson, J. R., & Molotch, N. P. (2009). Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site. *Ecohydrology, 2,* 115-128. https://doi.org/10.1002/eco.45

Vermote, E. F., Roger, J. C., & Ray, J. P. (2015). MODIS Surface Reflectance User's Guide Collection 6 Version 1.4

USGS. (2020). Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) Product Guide Version 3.0 LSDS-1368

Walker, J. J., de Beurs, K. M., Wynne, R. H., & Gao, F. (2012). Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment, 117,* 381-393. doi:10.1016/j.rse.2011.10.014

Walters, R. D., Watson, K. A., Marshall, H. P., McNamara, J. P., & Flores, A. N. (2014). A physiographic approach to downscaling fractional snow cover data in mountainous regions. *Remote Sensing of Environment, 152,* 413–425. doi:10.1016/j.rse.2014.07.001

Wang, P., Gao, F., & Masek, J. G. (2014). Operational Data Fusion Framework for Building Frequent Landsat-Like Imagery. *IEEE Transactions on Geoscience and Remote Sensing, 52*(11), 7353-7365. doi: 10.1109/TGRS.2014.2311445

Wang, Z., Schaaf, C. B., Sun, Q., Kim, J., Erb, A. M., Gao, F., Roman, M. O., Yang, Y., Petroy, S., Taylor, J. R., Masek, J. G., Morisette, J. T., Zhang, X, & Papuga, S. A. (2017). Monitoring land surface albedo and vegetation dynamics using high spatial and temporal resolution synthetic time series from Landsat and the MODIS BRDF/NBAR/albedo product. *International Journal of Applied Earth Observation and Geoinformation*, *59*, 104–117. doi:10.1016/j.jag.2017.03.008

Zhang, H., Zhang, F., Che, T., & Wang, S. (2020). Comparative evaluation of VIIRS daily snow cover product with MODIS for snow detection in China based on ground observations. *Science of the Total Environment, 724*(138156), 1- 11. doi:10.1016/j.scitotenv.2020.138156

Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment, 114,* 2610-2623. doi:10.1016/j.rse.2010.05.032

CHAPTER THREE: SPATIAL ANALYSIS OF SNOW-COVERED AREA PATTERNS BY WATER YEAR FROM STARFM RESULTS

3.1 Introduction

In highly heterogeneous mountain areas, the distribution of snow cover typically follows predictable patterns driven by elevation, slope, aspect, and vegetation. In general, higher elevations are more snow precipitation dominated as opposed to rain precipitation dominated (Tennant et al., 2015), and often have greater snow depths (Anderson et al., 2014) that persist longer in the landscape once melting begins (Anderton et al., 2004). Flatter slopes tend to accumulate more snow than steeper slopes due to gravitational forces (McClung & Schaerer, 2006), with very steep slopes often staying almost permanently snow-free (Farinotti et al., 2010). In the northern hemisphere, differences in solar radiation between slopes with north and south aspects promotes greater accumulation and later melt on predominantly north-facing slopes (Elder et al., 1998, Buhler et al., 2016). Vegetation, particularly forest canopy, can both intercept falling snow, creating a shallower snowpack directly beneath them, and help with snowpack retention by shading snow from incoming solar radiation (Veatch et al., 2009).

These relationships, however, are not always this straightforward. Localized differences in topographic features can create variability in snow cover behaviors. For example, while measuring snow depth with an unmanned aerial system, Buhler et al. (2016) found much higher variability in snow depth on their mountain top study site than their valley bottom site, which had much smaller depth gradients. At high elevation they

observed instances where bowls with deep snowpack were located near areas where all snow had been blown away. Slopes with varying steepness and aspects can also exhibit patterns of snow accumulation contrary to the general rules of gravity or solar radiation. Wind scour can be very effective at removing snow on windward-facing slopes, while more sheltered environments on leeward-facing slopes promote more snow accumulation, regardless of aspect or steepness (Hiemstra et al., 2006). In addition to influencing snow accumulation and melt by physically blocking falling snow and solar radiation, trees can have a significant influence on snow melt patterns through their release of longwave radiation. Lundquist et al. (2013) found that snow in warmer climates tended to persist longer in the openings between trees due to emitted longwave radiation exceeding incoming shortwave radiation, while snow persisted longer beneath tree canopies in colder climates.

Capturing these fine scale variabilities in snow cover in heterogeneous landscapes has been challenging with traditional remote sensing and in situ data collection methods. In their review of previous field studies considering the spatial variability of snow processes, Clark et al. (2011) noted that spatial variability of snow generally increased with spatial scale, so studies with limited spatial extents risked underestimating this natural variability. Conversely, field studies that attempted to capture a larger area by increasing space between point data capture risked missing these small-scale spatial correlations in the landscape. Remote sensing can make up for limitations of spatial extent due to the ability to capture data over large areas, but has limitations due to data resolution, both spatial and temporal. Remote sensing instruments have historically been able to provide data captured at high spatial resolutions, which can be used to identify

snow cover patterns in highly heterogeneous landscapes, or at high temporal resolutions, which can be used to identify quickly changing snow patterns, but not both.

While emerging remote sensing technology has increased both the temporal and spatial resolution of snow cover products (e.g. the Sentinel-2 instrument provides optical data at higher spatial resolutions of 10-20 m every 5 days), model and data downscaling are still required to produce daily high spatial resolution datasets. Such fine-scale resolutions, however, are necessary to fully capture the variability in snow cover due to catchment specific landscape characteristics and understand how snow patterns evolve overtime.

Here we describe the creation of a 20-year, daily 30-m dataset developed with STARFM for the purpose of monitoring SCA in the East River watershed. This effort builds on our previous work validating STARFM as an effective method to represent variability in SCA in a mountain landscape over the course of an entire year. We also perform preliminary analyses of this dataset to both address high-level scientific questions related to annual patterns of SCA in our watershed and demonstrate the type of analyses that this dataset, and others like it, can enable for answering further questions related to snow cover processes in the future.

3.2 Methods

3.2.1 Data

The study area used for the creation of the 20-year, daily dataset with STARFM is the same model domain area described in section 2.2.1. The time period used in the creation of the dataset spans WYs 2001-2020. The MODIS instrument came online in early 2000, so the first complete WY for which MODIS data is available is WY 2001.

We again used the MODIS product MOD09GA.006 (Terra Surface Reflectance Daily L2G Global 1 km and 500-m) daily surface reflectance values pre-processed in the same manner described in Section 2.2.2.

We used several different Landsat surface reflectance products to create the 20-year dataset, as several missions occurred throughout the study period and were, at times, discontinuous (Table 3.1). For the period spanning WY 2001 through May 2003, we used the USGS Landsat 7 Surface Reflectance Tier 1 (LE07/C01/T1_SR) product. These data are processed from the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor. The last ETM+ data acquisition used was collected on 25 May 2003, just prior to when the Scan Line Corrector (SLC) failed on the instrument on 31 May 2003 (Markham et al., 2004). For the period spanning 24 May 2003 (Landsat 5 and 7 data acquisitions were one day apart) through 15 November 2003, we used the USGS Landsat 5 Surface Reflectance Tier 1 (LT05/C01/T1_SR) product. These data are processed from the Landsat 5 Thematic Mapper (TM) sensor. TM data became unavailable on GEE after 15 November 2011. For the period spanning 15 November 2011 through 4 May 2013, we again used Landsat 7 ETM+ data, despite known SLC errors for this period. We chose not to apply any SLC error correction methods to the ETM+ data or the STARFM results produced from it, instead leaving the decision to future users of these data to decide if and how they wish to apply SLC corrections. Sufficiently cloud-free Landsat 8 OLI data were not available on GEE until 3 May 2013. For the final period spanning 4 May 2013 through 6 October 2020, we used USGS Landsat 8 Surface Reflectance Tier 1 (LC08/C01/T1_SR). All Landsat surface reflectance data described herein have been processed to radiometric and geodetic accuracy with Level-1 Precision and Terrain Correction standards (L1TP).

ETM and ETM+ data have been atmospherically corrected using LEDAPS, they include cloud, shadow, water, and snow masks produced with CFMASK, a per-pixel saturation mask, and are projected onto a standardized reference grid (USGS, 2020). As described in Section 2.2.2, OLI surface reflectance data are processed similarly, with the exception of being atmospherically corrected using the updated method LaSRC (USGS, 2020), but otherwise meet the same standards as ETM and ETM+ data.

Table 3.1: Landsat datasets and the dates each was used for the generation of the 20-year, daily STARFM NDSI dataset.

Landsat Dataset	Beginning Date	End Date
Landsat 7 ETM+	28 September 2000	25 May 2003
Landsat 5 ETM	24 May 2003	15 November 2011
Landsat 7 ETM+	14 November 2011	4 May 2013
Landsat 8 OLI	3 May 2013	6 October 2020

3.2.2 Data Pre-Processing and STARFM

Landsat and MODIS data pre-processing were completed using GEE as described in Section 2.2.3, with the only notable difference being the band numbers used for the calculation of NDSI. Landsat 5 ETM and Landsat 7 ETM+ band orders are slightly different than Landsat 8 OLI band order, requiring the use of ETM and ETM+ Band 2 (green) and Band 5 (SWIR 1) to calculate NDSI instead of OLI Band 3 (green) and Band 6 (SWIR 1).

Pre-processed MODIS and Landsat data were used as inputs to the STARFM model, described at length in Section 2.2.4. The model parameters used during validation analysis remained unchanged, however, this time STARFM was run on a Boise State's R2 high-performance computing cluster due to the size of data files and R2's storage capabilities.

3.2.3 Post-Processing of STARFM Data

In addition to generating daily NDSI results from STARFM for WYs 2001-2020, we also extracted daily cloud cover information for MODIS, Landsat, and STARFM files. To find this, we calculated the amount of data present for each layer in the MODIS, Landsat, and STARFM raster stacks for each date of analysis. We assigned a value of 1 to all pixels that contained any non-NA value, summed these values by layer, then normalized them by the total number of pixels in each layer. This produced values between 0-1 that represent the fraction of pixels in each layer that containing data. For the MODIS and Landsat input files, this value between 0-1 was the fraction of data that remained after the cloud mask was applied (described in detail in Section 2.2.3), also known as the inverse cloud fraction. For STARFM results, a value of less than 1 for a

raster layer implies that there were pixels in that layer for which STARFM could not predict a value. There are a few reasons this may happen, the first being that the corresponding Landsat and MODIS pixels for that date were cloud covered, and therefore STARFM had no information by which to use to make a prediction. A no-value STARFM pixel can also occur when just one of the input files is missing a value for the corresponding pixel.

According to the theoretical underpinnings of STARFM, we can assume that when a prediction is made on a date when an input pair (MODIS and Landsat acquisitions) are present, the observed Landsat surface reflectance data and the synthetic data will be exactly the same (Gao et al., 2006). However, our results show that this is not always the case. STARFM generates images identical to Landsat surface reflectance only when 100% of MODIS data are present in the prediction pair. If there is any missing MODIS data, the fraction of data available in the synthetic STARFM raster layer is always less than the fraction of data available in the Landsat image from that same date. Our best explanation for why this occurs has to do with the STARFM's sample filtering capabilities. According to Gao et al. (2006), fine-resolution neighbor pixels from the Landsat image are selected that are spectrally similar to the fine-resolution central pixel of the moving window. Before these pixels are compared to their respective resampled MODIS pixels, however, STARFM filters out any neighboring pixels it considers "poor quality", which means they are not able to provide better spectral and spatial information than that of the moving window's central pixel. This sample filtering method is likely the reason why Landsat pixels are being excluded from STARFM results, even though both images are from the same date. Because of this issue with the output, the fact that the

sample filtering algorithm is not a model parameter that can adjusted by the user, and that we are using Landsat data as our "ground-truth" data in this study, we made the decision to replace any STARFM layer predicted on the date of a Landsat acquisition with the Landsat data before performing any analysis on our results. The only exceptions to this rule were for dates when 100% of MODIS data were present, and therefore replacing STARFM data with Landsat data provided no further information for the results.

3.2.4 Selection of Individual Water Years for In-Depth Analysis

After the STARFM dataset had been created, we chose three separate water years for in-depth analysis of SCA patterns from our STARFM results with the goal of understanding how SCA varies within the East River watershed during an average, wet, and dry water year. We selected the years meeting these criteria by reviewing Butte SNOTEL station (wcc.nrcs.usda.gov/snow/, site number 380) maximum SWE data in the same manner as we did when choosing the water year for the STARFM validation analysis in Section 2.2.1. The mean maximum SWE for WYs 2000-2018 was 14.39 in, so we chose WY 2010 with a maximum SWE of 13.3 in as our average water year (which happened to have the same maximum SWE value as our validation year of WY 2016). We chose WY 2008 with the highest maximum SWE value of 23.7 in as our wet water year, and WY 2012 with the lowest maximum SWE value of 8.2 in as our dry water year (Figure 3.1).

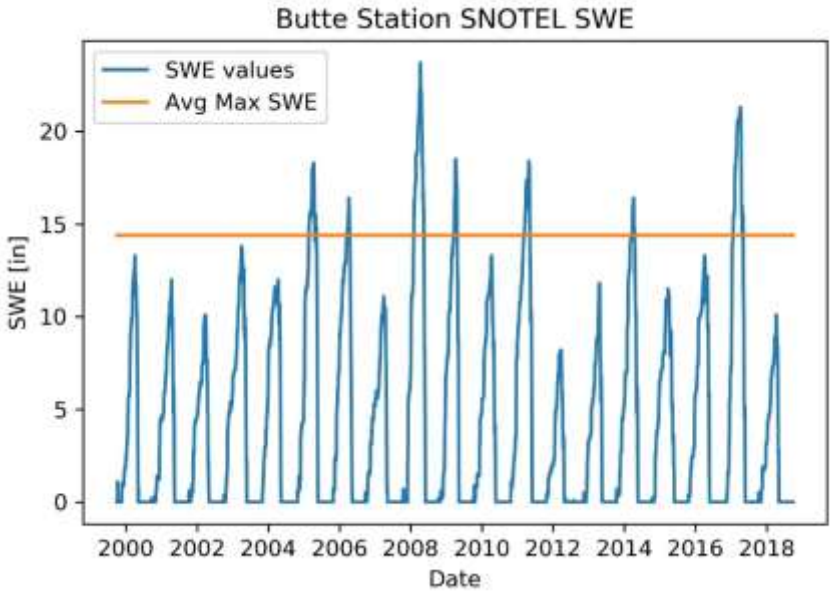


Figure 3.1: SWE values by WY from the Butte SNOTEL station located within the East River Watershed. The orange bar represents the mean maximum SWE value from the 2000-2018 WY period.

3.2.5 Spatial Snow-Cover Analysis of STARFM Results

We created binary "snow" and "no-snow" maps for each day of our STARFM results for our chosen water years using the NDSI threshold of 0.4 in the same manner as described in Section 2.2.5. We then determined the number of snow-covered days by pixel for each water year by summing all raster layers (one for each day) for that year together. Because each snow-covered pixel had already been assigned a value of 1, the resulting sum for each pixel was equal to the number of days for that year when that pixel was classified as "snow" (Figure 3.2).

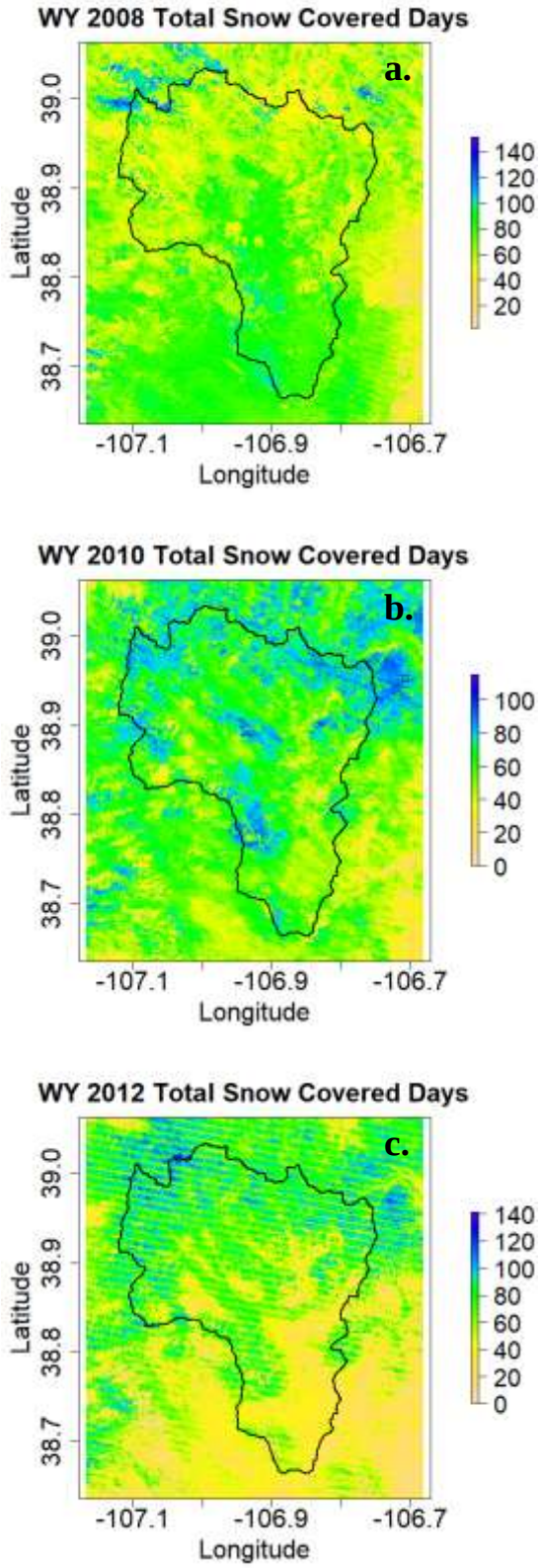


Figure 3.2: The total number of snow-covered days by pixel for each water year of analysis, a) WY 2008, b) WY 2010, and c) WY 2012.

A simple per-pixel sum of snow-covered instances, however, may not be representative of actual SCA trends due to missing data filtered out by our cloud and cloud-shadow masks (see Section 2.2.3 for further detail). For example, this analysis reports more snow-covered days in the lower elevations of the watershed (Figure 3.2a), which is not what we would expect to see for a water year with an extremely high max SWE. It is possible that persistent cloud-cover in the valleys was consistently being misidentified as snow, but as we saw in our data availability analysis (Section 2.3.3.1), our cloud filters are much more likely to misidentify snow as clouds at higher elevations. Therefore, we chose to normalize the data by dividing the number of snow-covered days for each pixel by the total number of days valid data were available, regardless of snow cover status, for that pixel. Sproles et al. (2018) used a similar method to calculate monthly snow cover frequencies from MODIS data for input into their streamflow forecasting model. This produced a value representing the percent of snow-covered days for which a pixel had data (Figure 3.3) and the numerical distribution of those percent values for all pixels (Figure 3.4), repeated for all three water years. The numerical distributions of percent snow-covered days provide a perspective for which to evaluate differences in snow cover duration for the entire watershed over multiple water years. For example, Figure 3.4a shows that a clear majority of pixels for WY 2008 were snow-covered less than 50% of the time. However, the presence of the right tail of the distribution for this year shows that the pixels that did retained snow did so for much longer than for other water years (Figures 3.4b and 3.4c).

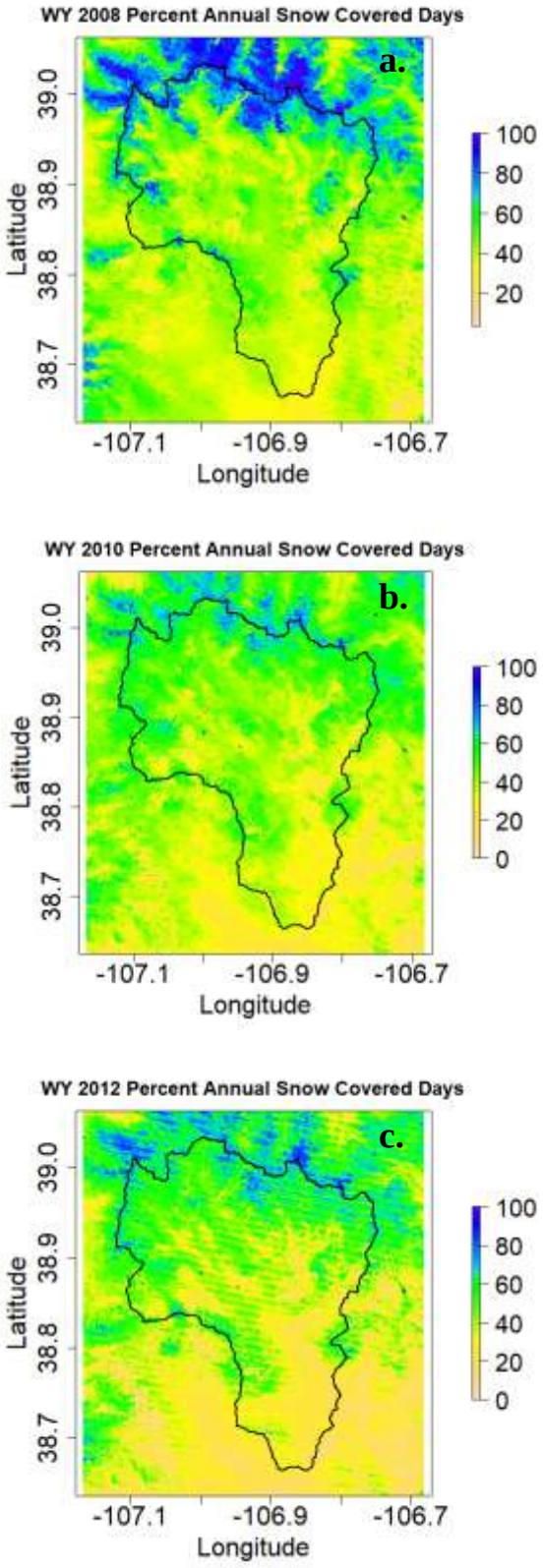


Figure 3.3: Percent annual snow-covered days for a) WY 2008, b) WY 2010, and c) WY 2012.

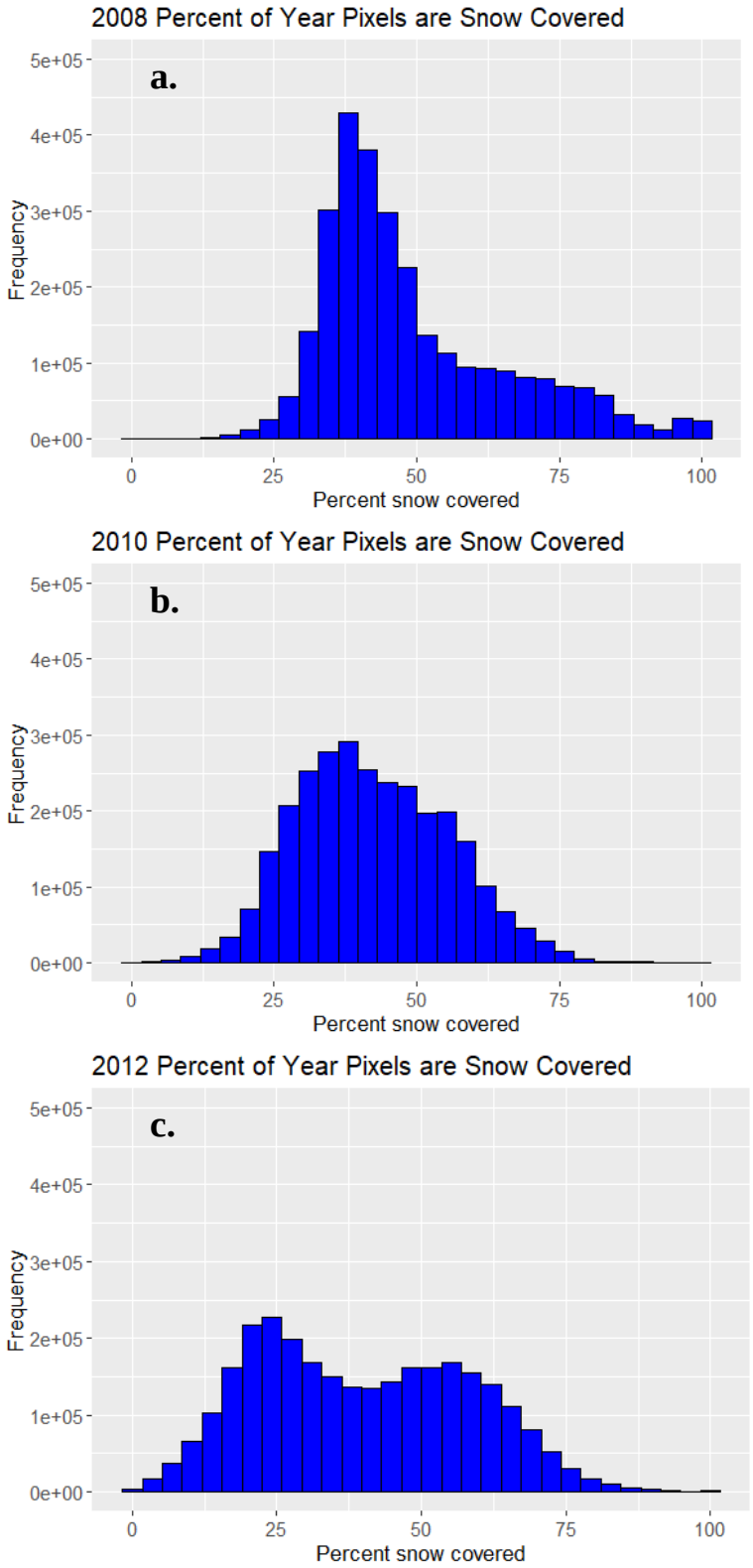


Figure 3.4: The distribution of percent annual snow-cover for a) WY 2008, b) WY 2010, and c) WY 2012.

3.2.5.1 Elevation

With the annual percent snow cover determined for the entire study area for each water year, we delineated subset areas based on landscape features to examine their relationships with percent snow cover. When considering SCA and elevation, we expected to see a positive correlation of increasing SCA that persisted for more days in higher elevation areas (Anderson et al., 2014). This relationship is not always straightforward, however, as high elevations tend to have sparse or no vegetation, so wind scour can leave some areas devoid of snow, while depositing that snow in nearby locations (Buhler et al., 2016).

We used the Low (2399-2963 m), Medium (>2963-3314 m), and High (>3144-4332 m) elevation bands as calculated from the SRTM 30-m DEM (described in Section 2.2.6.1). We then calculated the mean and standard deviation of the percent annual snow cover values for the three representative water years by elevation band (Figures 3.5, 3.6, and 3.7). We also counted and displayed the numerical distribution of the annual percent snow cover across elevation bands for each water year to explore how the percent snow cover was distributed across elevation bands.

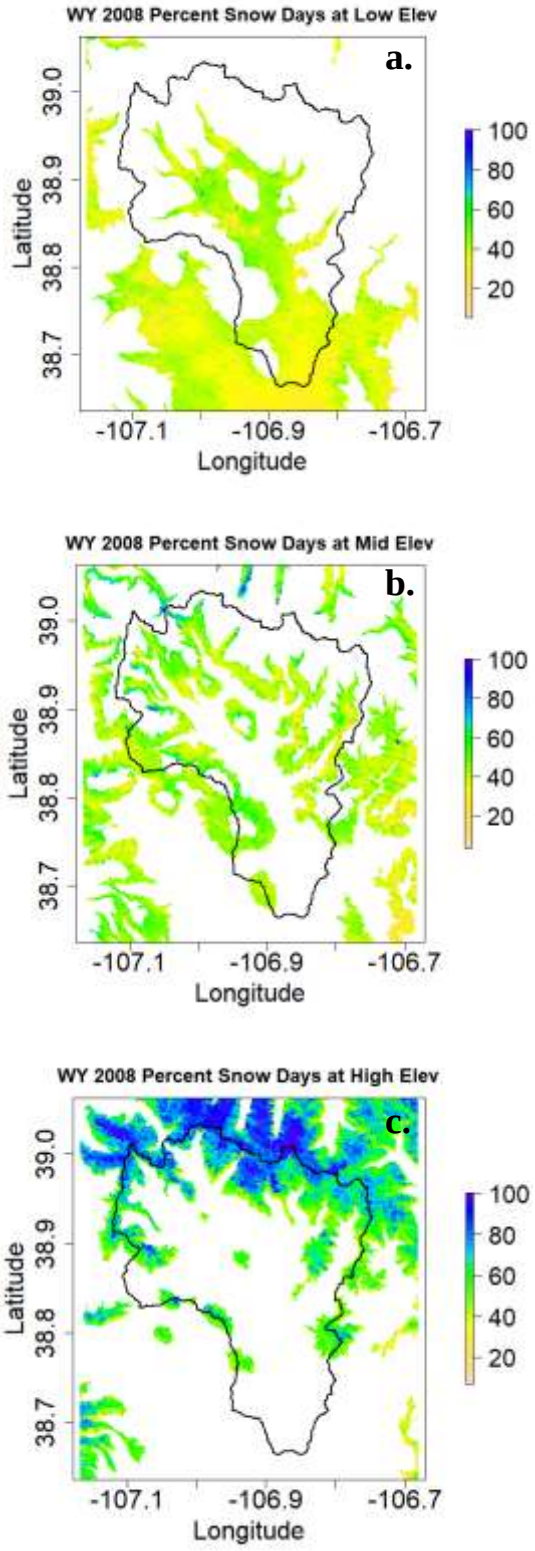


Figure 3.5: Percent annual snow-covered days by pixel for WY 2008 at the a) low elevation band, b) mid elevation band, and c) high elevation band.

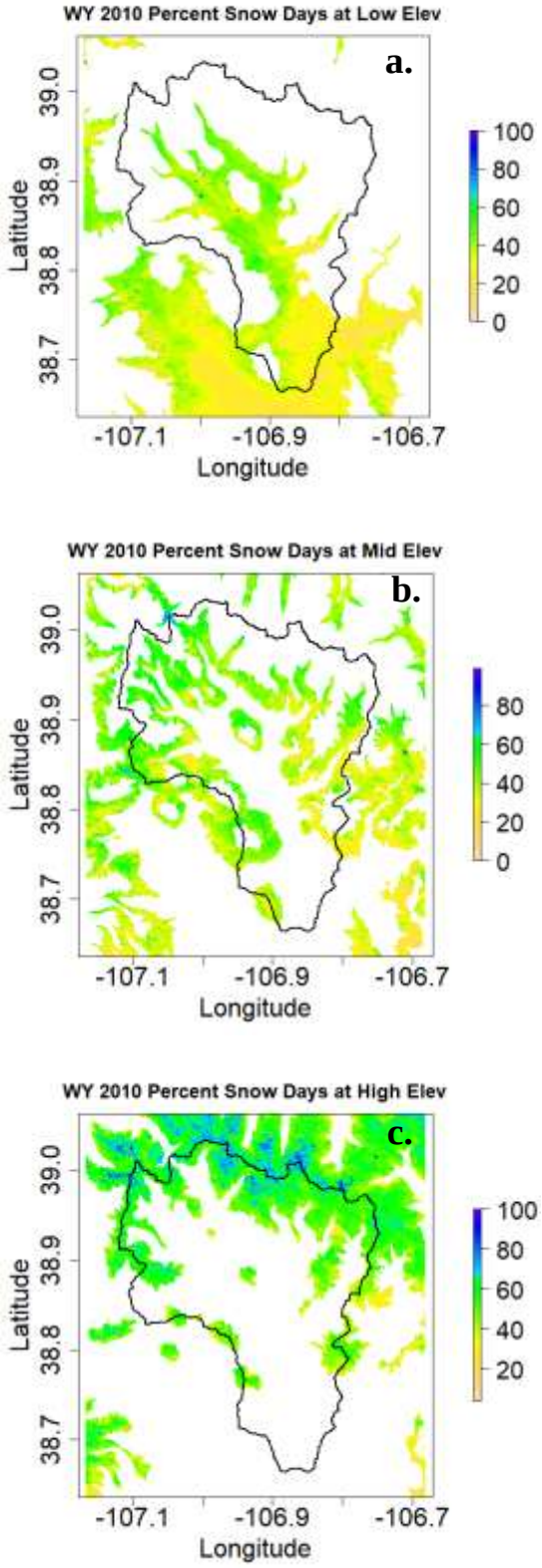


Figure 3.6: Percent annual snow-covered days by pixel for WY 2010 at the a) low elevation band, b) mid elevation band, and c) high elevation band.

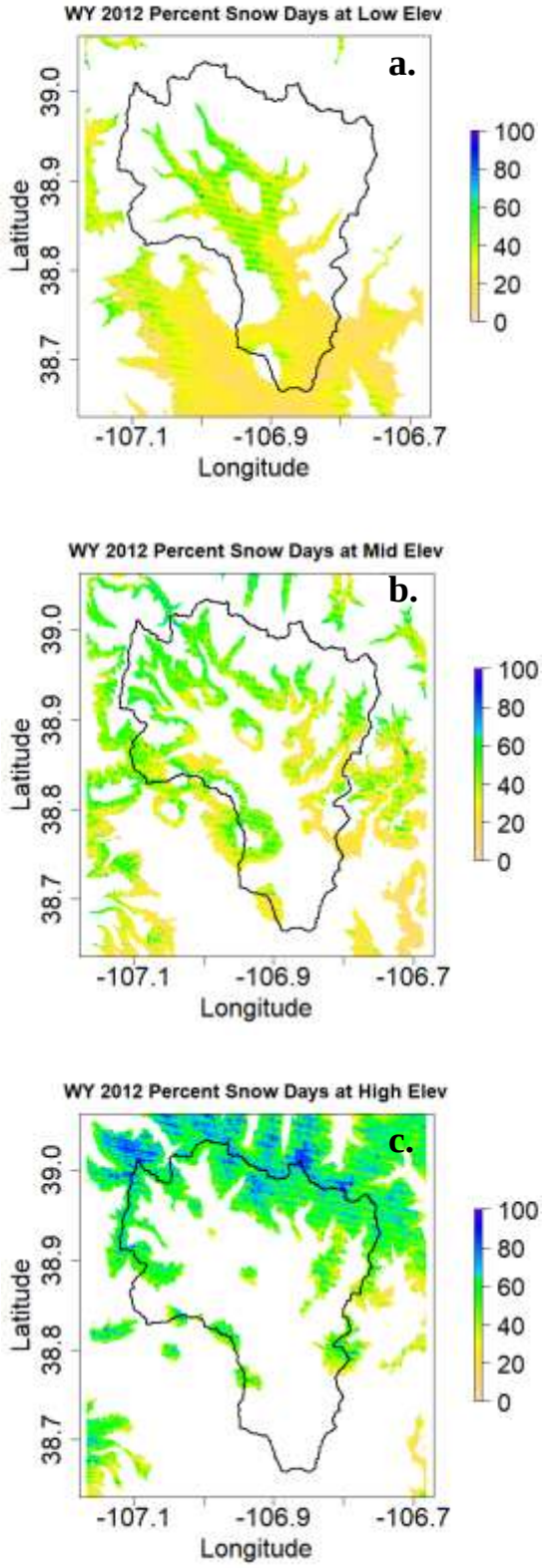


Figure 3.7: Percent annual snow-covered days by pixel for WY 2012 at the a) low elevation band, b) mid elevation band, and c) high elevation band.

3.2.5.2 Slope

Flatter areas of a watershed tend to accumulate more snow due to sluffing and avalanching occurring on adjacent slopes with steeper inclines (McClung & Schaerer, 2006). As a result of this accumulation, snow in these lower sloped areas can be deeper and persist longer in the landscape due to slower melting (Elder et al., 1998; Kerr et al., 2013). However, areas with flatter slopes often occur in lower elevation areas of a watershed, as is the case with our study area (Figures 2.14b and 2.15b), indicating that relationships with snow cover duration and elevation may be the dominating effect in these areas (Daly et al., 2012; Che et al., 2016).

We use our Low (0-11°), Medium (>11-20°), and High (>20-74°) slope bands as calculated from the SRTM 30-m DEM (described in Section 2.2.6.2). As with elevation, we calculated the mean and standard deviation of the percent annual snow cover values for our three water years by slope band (Figures 3.9, 3.10, and 3.11).We also counted and displayed the numerical distribution of annual percent snow cover across slope bands for each water year to explore how the percent snow cover was distributed across slope bands.

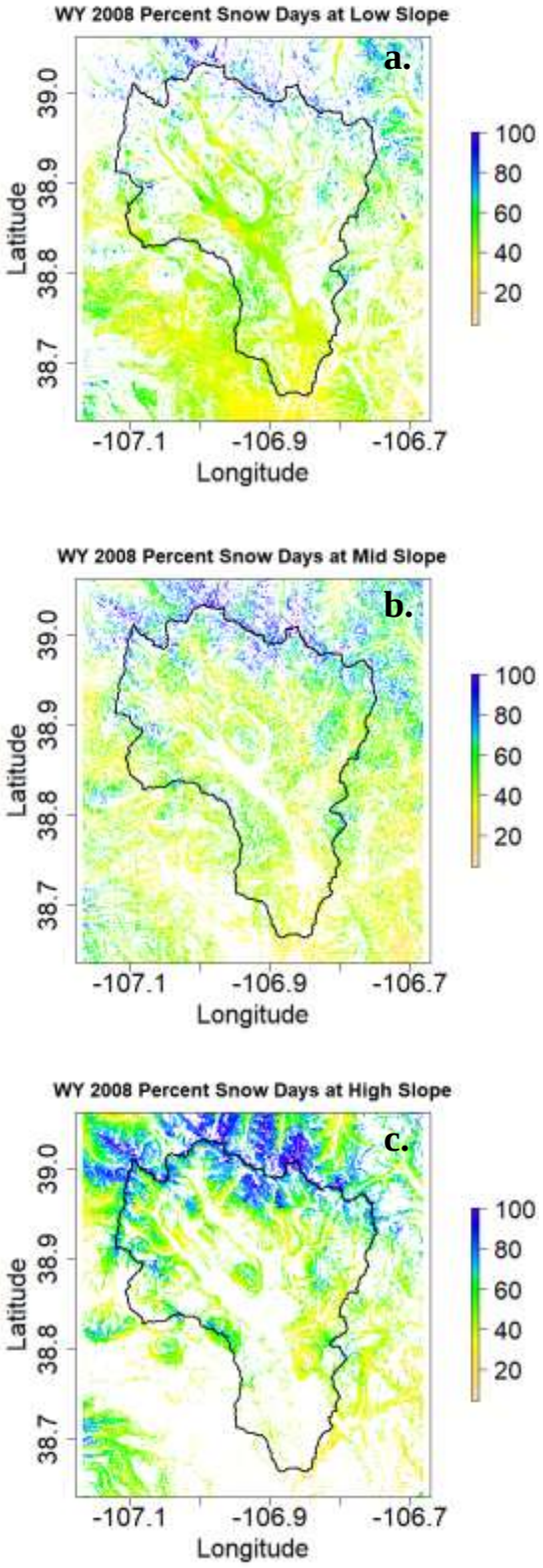


Figure 3.8: Percent annual snow-covered days by pixel for WY 2008 at the a) low slope band, b) mid slope band, and c) high slope band.

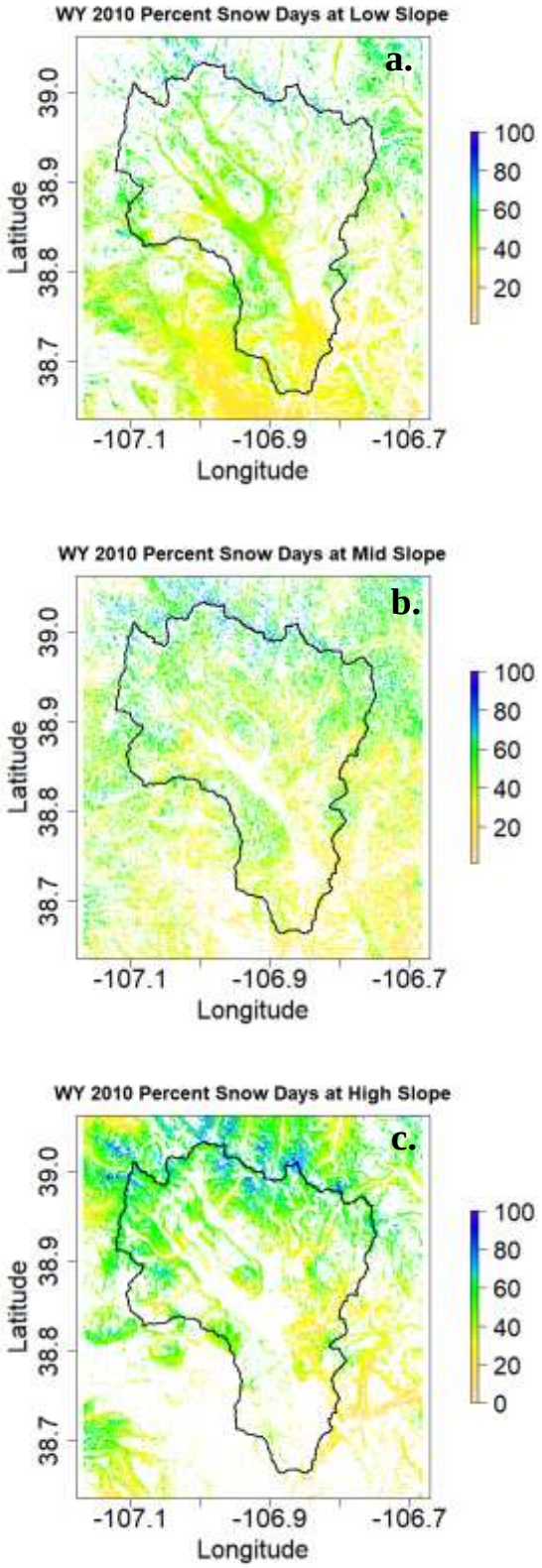


Figure 3.9: Percent annual snow-covered days by pixel for WY 2010 at the a) low slope band, b) mid slope band, and c) high slope band.

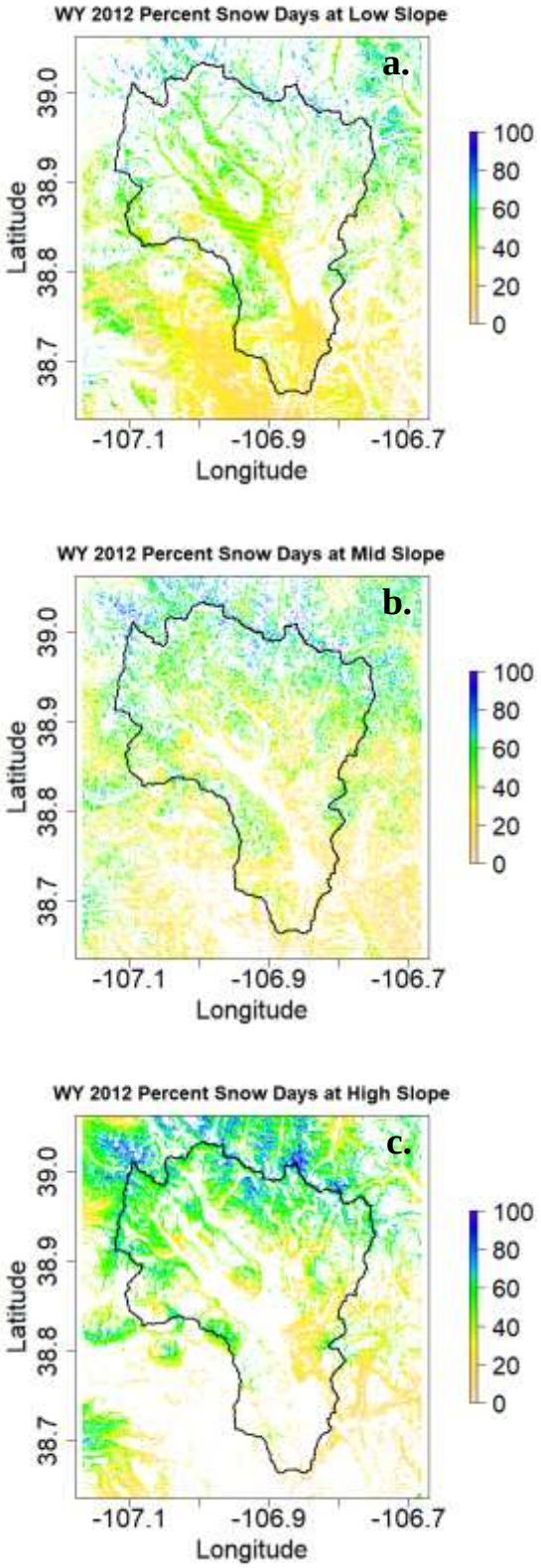


Figure 3.10: Percent annual snow-covered days by pixel for WY 2012 at the a) low slope band, b) mid slope band, and c) high slope band.

3.2.5.3 Aspect

We expected to see greater SCA on north/east aspects in our watershed due to unequal solar loading on north/east and south/west aspects in the Northern Hemisphere. We also recognized that variations in SCA due to aspect may be observed due to differences in accumulation on western, windward-facing slopes, versus eastern, leeward-facing slopes (Hiemstra et al., 2006).

We evaluated these potential SCA trends by using our aspect directions as calculated from our STRM 30-m DEM data (described in Section 2.2.6.3). We calculated mean and standard deviations from the percent annual snow cover values for our three water years by aspect direction. We counted and displayed the numerical distribution of annual percent snow cover across all aspect directions for each water year to explore how the percent snow cover was distributed across aspect direction.

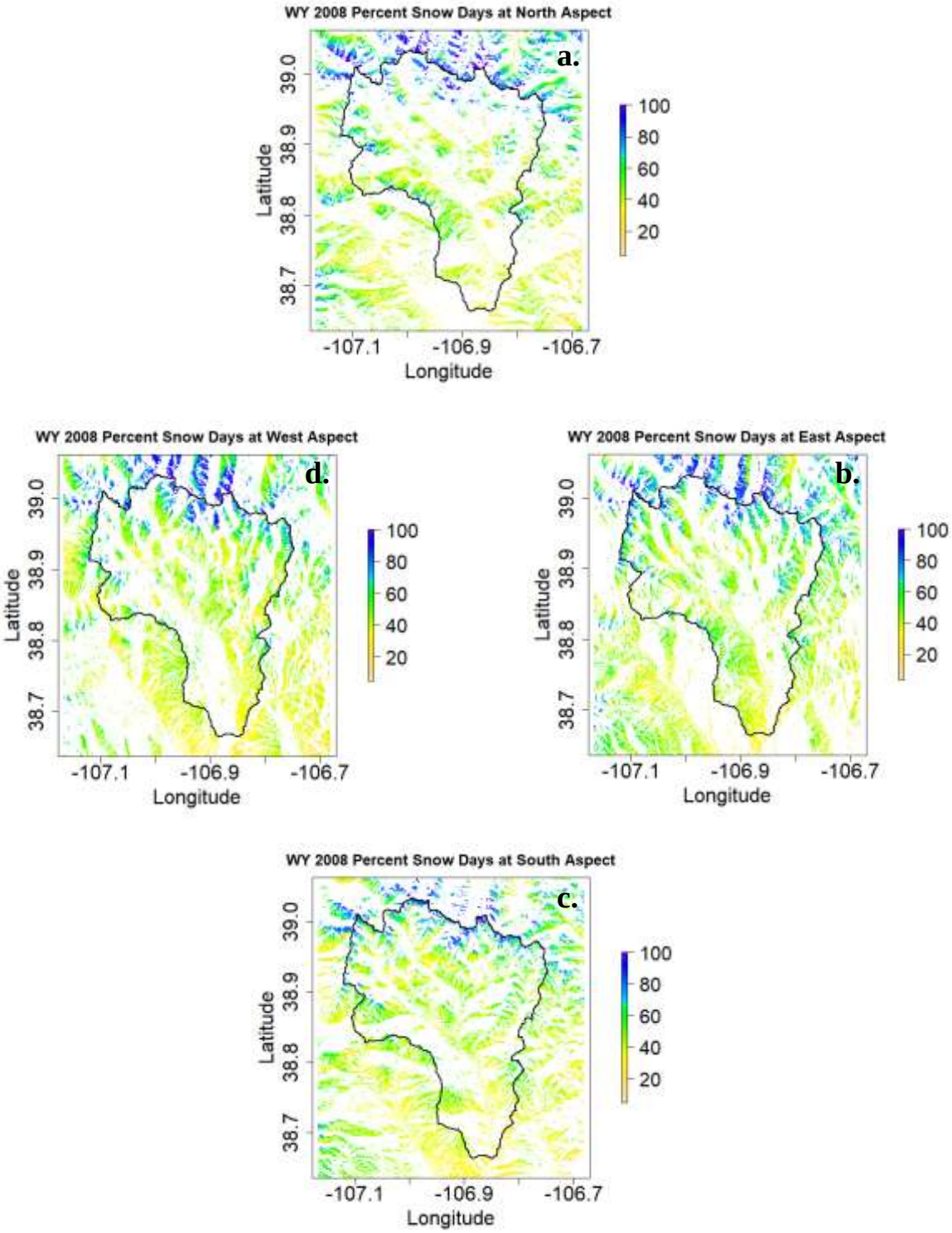


Figure 3.11: Percent annual snow-covered days by pixel for WY 2008 at the a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

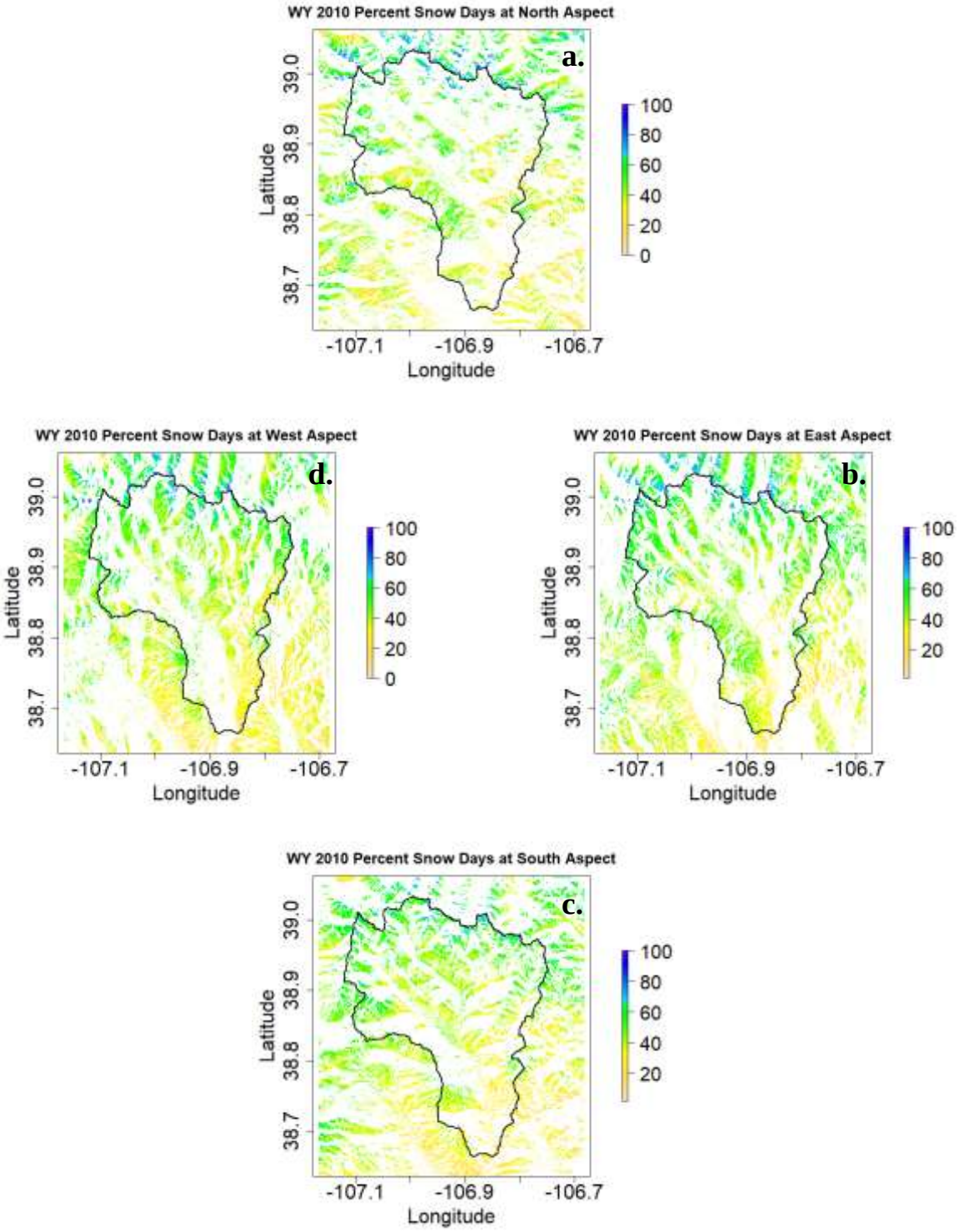


Figure 3.12: Percent annual snow-covered days by pixel for WY 2010 at the a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

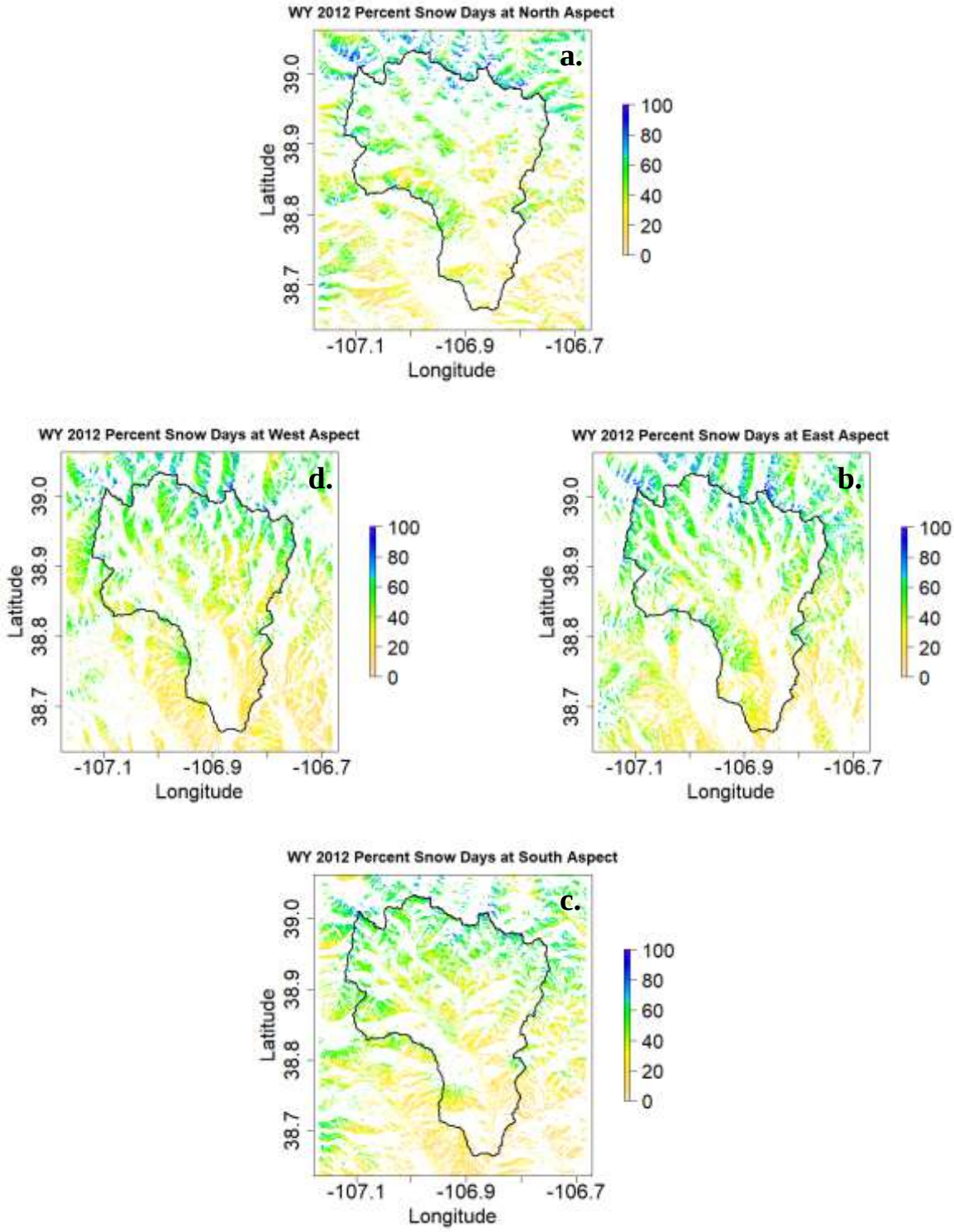


Figure 3.13: Percent annual snow-covered days by pixel for WY 2012 at the a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

3.2.5.4 Land Cover Type

Land cover and vegetation type and their distribution can influence SCA patterns in a variety of ways. Trees often vertically intercept falling snow (Anderson et al., 2014), while bushes influence small-scale heterogeneity in snow depth by laterally trapping snow (Clark et al., 2011). Forest canopy cover shades snow on the ground from solar radiation and induces slower melting than open areas (Veatch et al., 2009). In other instances, however, trees can enhance melting by emitting longwave radiation that exceeds incoming solar shortwave radiation (Lundquist et al., 2013).

Additionally, factors such as elevation, slope, and aspect that influence persistence in snow cover can also influence vegetation distribution. For example, Anderson et al. (2014) found that forested areas in the Dry Creek Experimental Watershed in southwest Idaho tended to be located on high elevation, north-facing slopes, factors which also contributed to greater snowfall amounts at those locations.

To evaluate relationships between land cover and SCA, we used our land cover classifications determined from USGS Landfire 30-m data (Landfire, 2016) (described in Section 2.2.6.4). We omit the “other” land cover classification as this area is small (1% of the watershed) and consists of very different land cover types (human development, water, and permanent snow/ice, which itself is a negligible 0.04% of the watershed area), and is thus likely to not generate meaningful relationships. For the rest of our land cover types, we calculated the mean and standard deviation of the annual percent snow cover values for our three water years by land cover classification area. We also counted and displayed the numerical distribution of annual percent snow cover across all land cover

types for each water year to better understand the spatial distribution of snow cover as determined by land cover type.

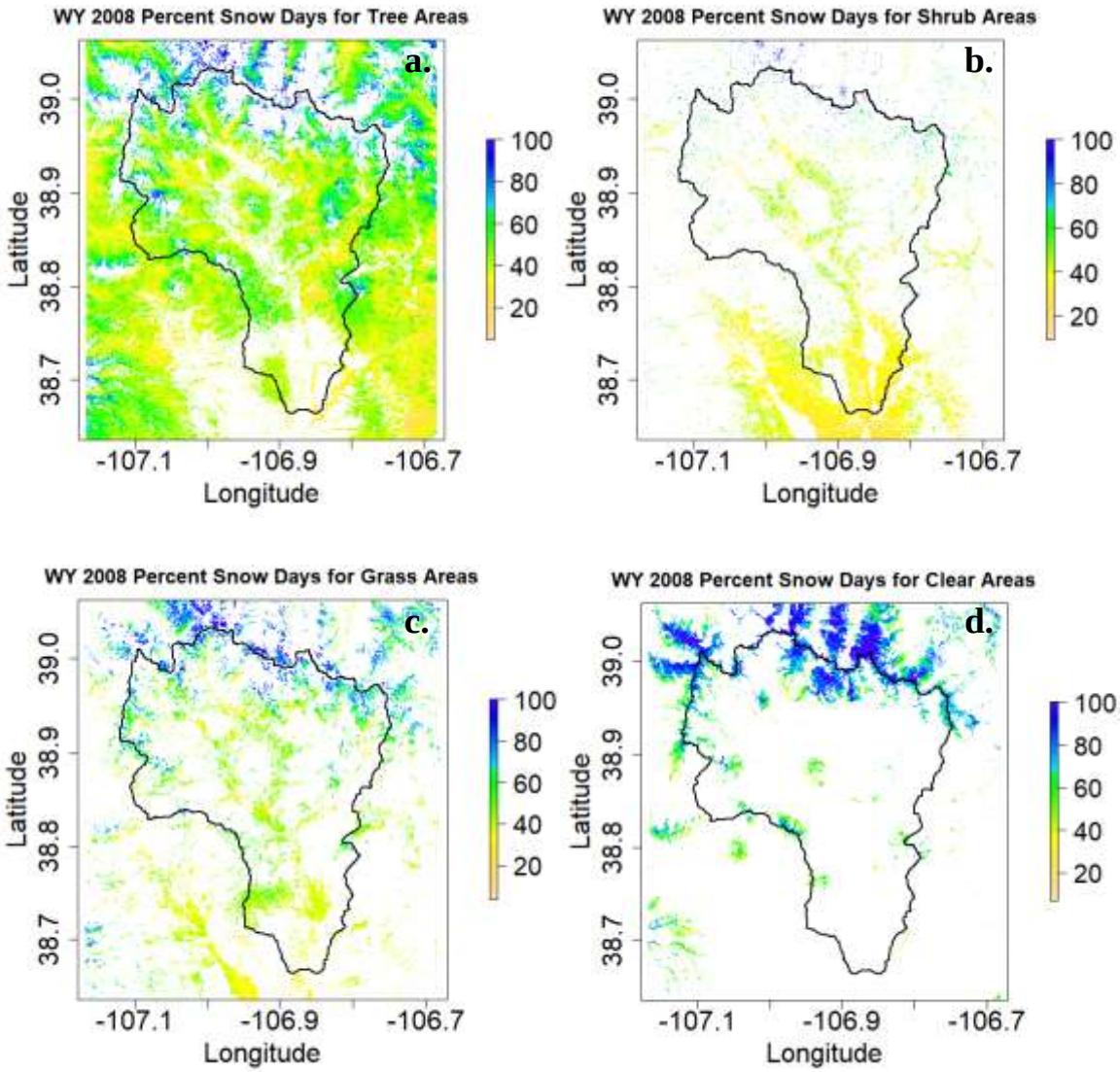


Figure 3.14: Percent annual snow-covered days by pixel for WY 2008 for a) tree areas, b) shrub areas, c) grass areas, and d) clear areas.

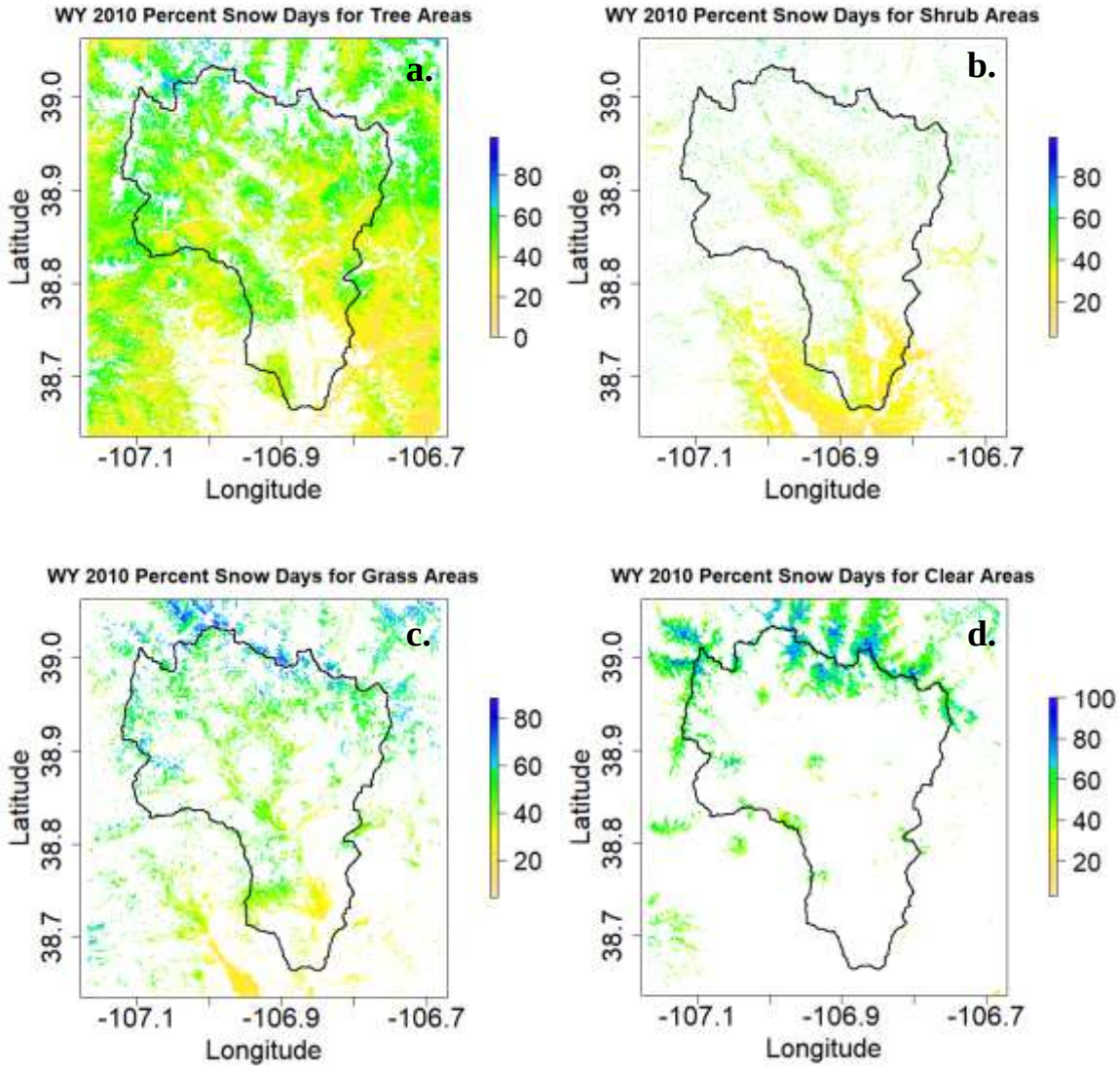


Figure 3.15: Percent annual snow-covered days by pixel for WY 2010 for a) tree areas, b) shrub areas, c) grass areas, and d) clear areas.

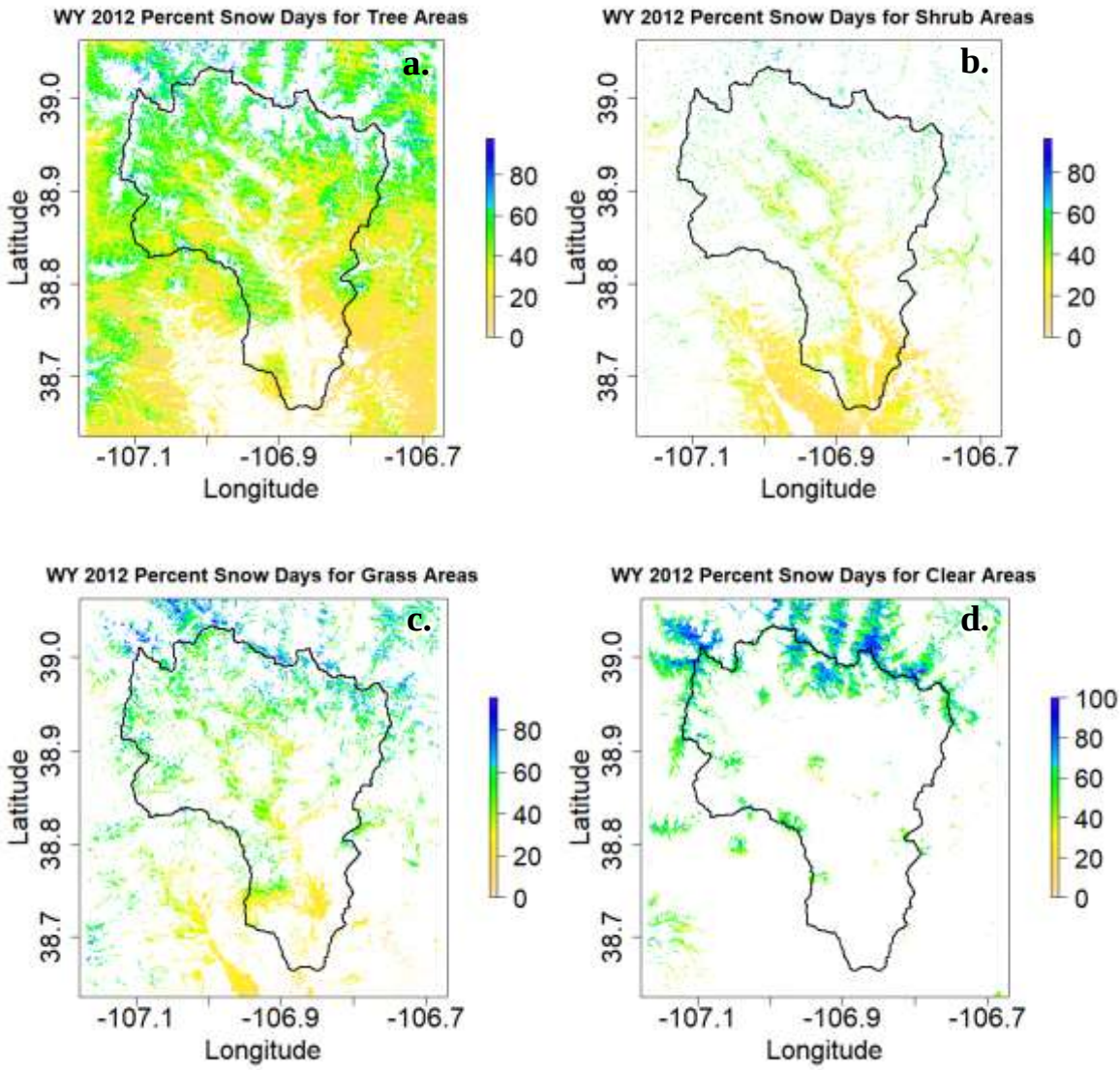


Figure 3.16: Percent annual snow-covered days by pixel for WY 2012 for a) tree areas, b) shrub areas, c) grass areas, and d) clear areas.

3.2.6 Comparison of SCA with Stream Discharge Data

To provide an example of how information regarding changes in SCA over a water year can be related to CZ processes in the local watershed, we examined the East River's discharge at the watershed outlet for each of the three analyzed water years. The volume, timing, and shape of an annual stream discharge curve for a snow-dominated watershed can be heavily influenced by snow accumulation and melt patterns (Knowles

et al., 2015), surface and subsurface conditions (Barnhart et al., 2016), and routes the water takes before it becomes streamflow (Tokunaga et al., 2019). Thus, we expected to see signals from these characteristics reflected in our data.

We retrieved discharge data for all three water years from the USGS gauging station 09112500 East River at Almont, CO (waterdata.usgs.gov). This gauging station is located immediately upstream from the East River's confluence with the Taylor River at the south end of the watershed (Figure 3.17). Figure 3.18a shows the variation in the size and shape of the hydrograph peaks for the three analyzed water years, and Figure 3.18b displays the differences in the cumulative sums of the discharge for the entire water year.

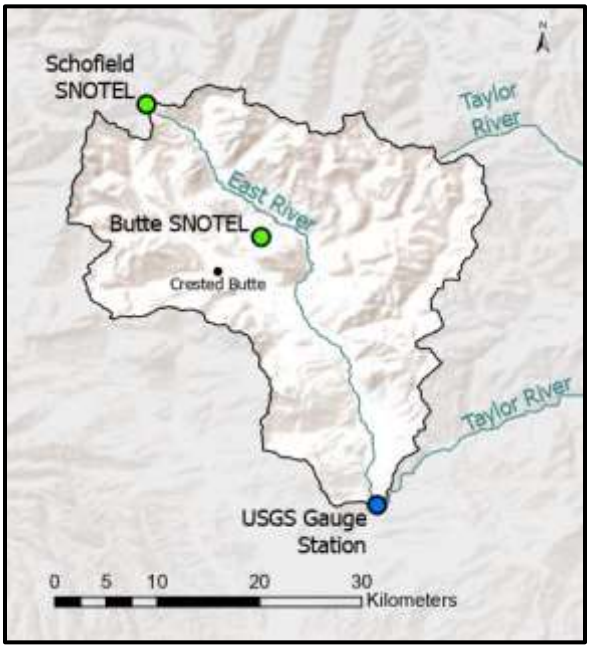


Figure 3.17: Location of the USGS gauge station 09112500 East River at Almont (blue)

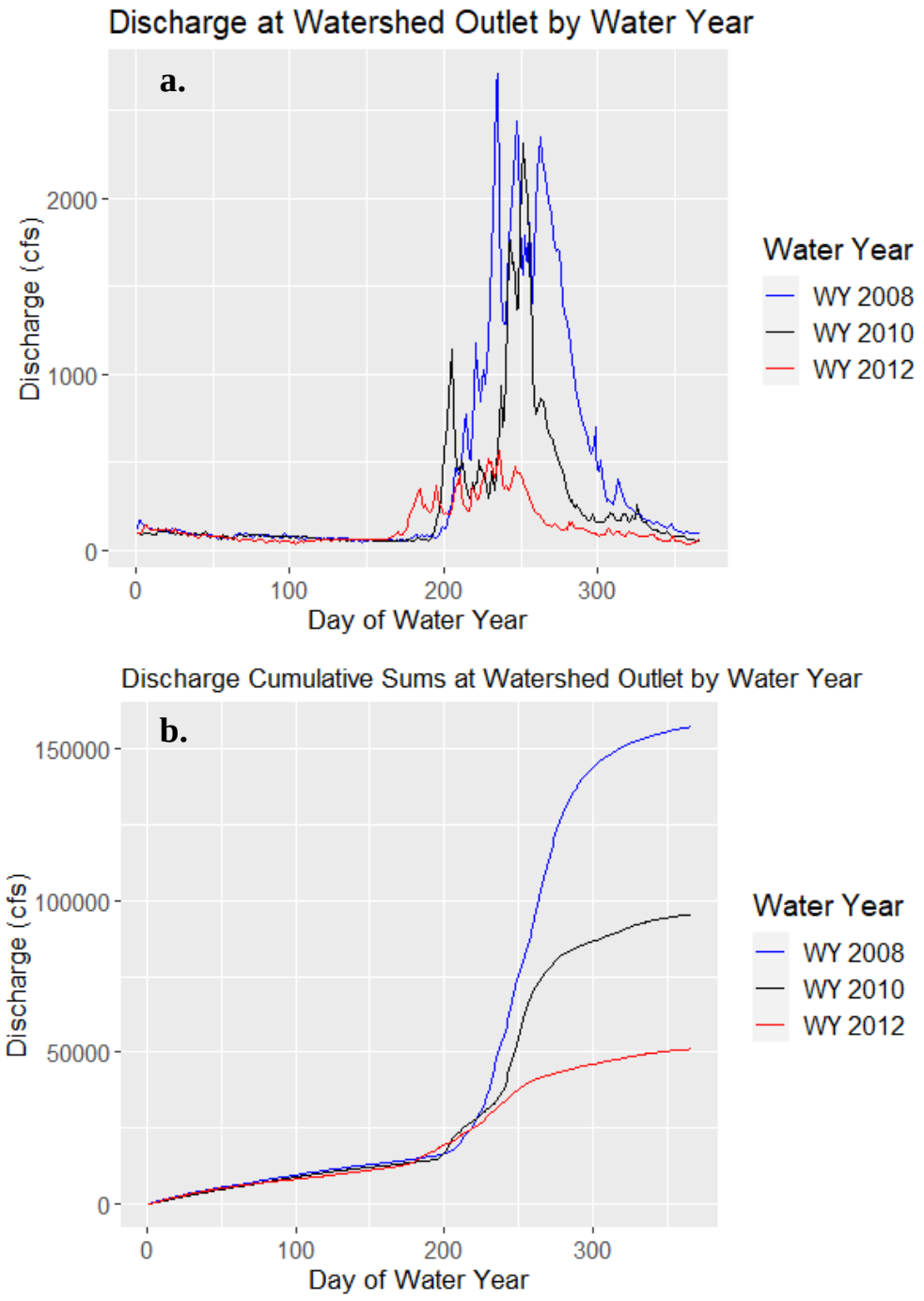


Figure 3.18: a) Discharge at the watershed outlet for the three water years of analysis, and b) the cumulative summations of the discharge values from a.

To compare SCA and stream discharge, we found the percent SCA of the entire watershed for each day of the water year (Figure 3.19). We then smoothed these SCA data by calculating a 10-day, centered moving average to correct for missing data due to cloud cover. To compare SCA with stream discharge, we found the latest day of the water year when our watershed SCA as calculated from our moving average was 50% or above. We identified the time lag between that point and the time at the center of mass of the hydrograph, or when 50% of the stream discharge for that water year occurred. The amount of time between when the majority of water as snow cover disappears from the watershed to when it is detected as streamflow at the outlet should provide a first order understanding of the CZ pathways by which water transitions from melt to streamflow. We repeated this process for all three water years of interest and compared the differences in the time lags for each.

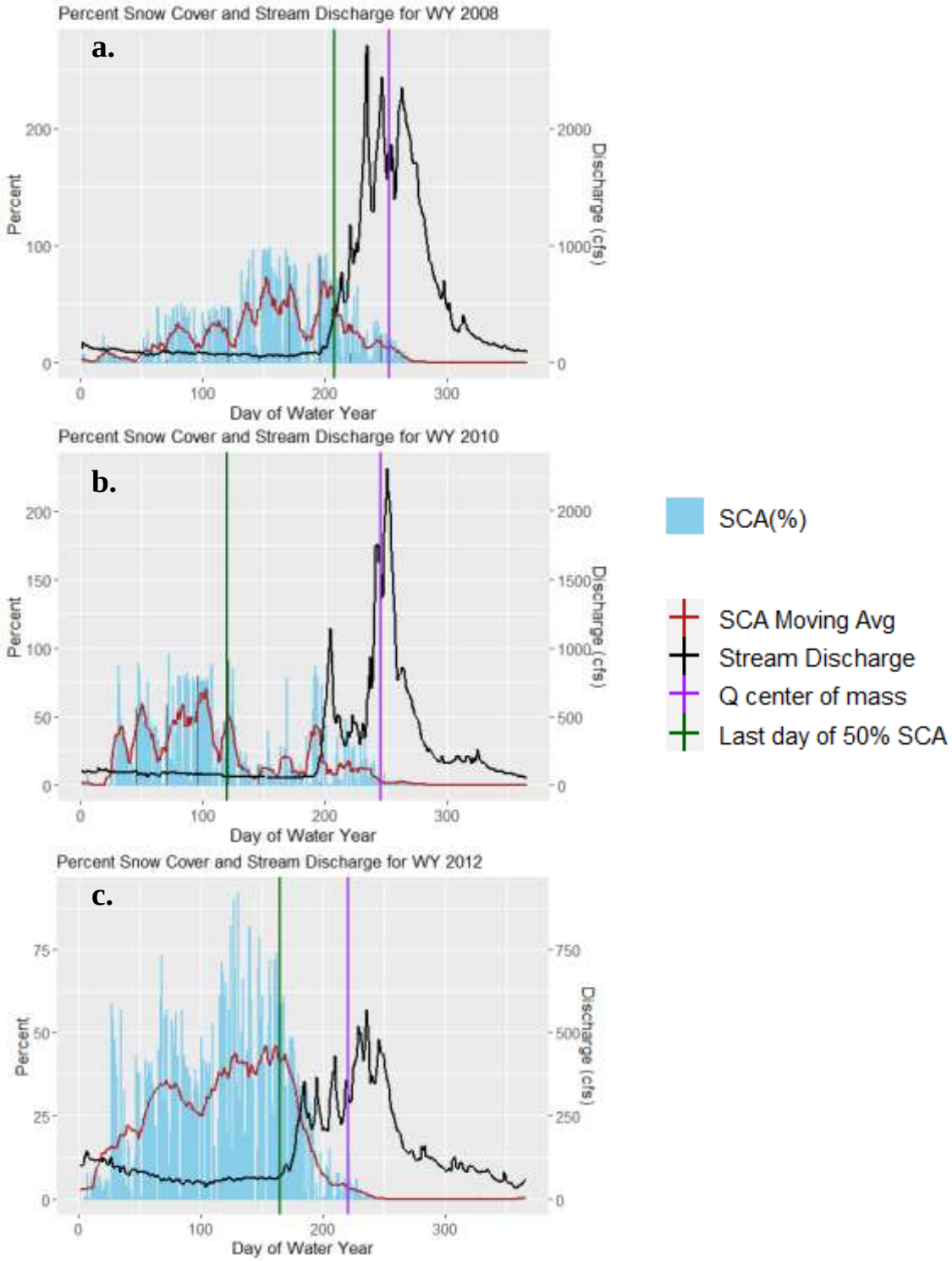


Figure 3.19: SCA and stream discharge (Q) by day of water year for WYs a) 2008, b) 2010, and c) 2012. In each graph: SCA as percent of the watershed by day (blue bars), SCA 10-day centered moving average (red line), daily discharge at the watershed outlet (black line), the last day the watershed had 50% or more SCA observed (green vertical line), and the center of mass of the hydrograph (purple vertical line).

3.3 Results

3.3.1 STARFM 20-Year Dataset

We created a 20-year, daily dataset of 30-m resolution NDSI values for WYs 2001-2020. The dataset contains NDSI values derived from Landsat data on all days of Landsat acquisition, and synthetic NDSI values predicted by STARFM for all other days in between. Because STARFM works by downscaling coarse spatial resolution data acquired on the date of interest, our dataset was not able to include information for dates when MODIS data was unavailable from GEE. These dates include 16 June-2 July 2001, 20-27 March 2002, 17-23 December 2003, 15 February 2016, 19-27 February 2016, 1 January 2019, 9 February 2019, and 28 February 2019.

3.3.2 Spatial Snow-Cover Analysis of STARFM Results

To demonstrate the utility of STARFM synthetic data, we closely analyzed daily SCA patterns for 3 individual water years from the 20-year dataset representing below average (dry, 2012), average (2010), and above average (wet, 2008) annual peak SWE. We accomplished this by calculating the annual percent of snow-covered days for each water year and analyzing how percent snow cover varied for the landscape features of elevation, slope, aspect, and land cover type.

3.3.2.1 Elevation

This analysis indicates that mean annual SCA increases with elevation, regardless of annual peak SWE, though variability in these values increases in drier years (Figure 3.20). Mean annual percent snow cover was the highest within each elevation band for WY 2008, the wettest water year, at 38% for low elevations, 43% for mid-elevations, 65% for high elevations, with standard deviations of 6%, 10%, and 16%, respectively.

For WY 2010, the average water year, mean annual percent snow cover was 34% for low elevations, 39% for mid-elevations, 54% for high elevations, with standard deviations of 8%, 11%, and 12%, respectively. Mean annual percent snow cover was lowest for all but the highest elevation band for the driest water year, WY 2012, with 27% for low elevations, 36% for mid-elevations, 55% for high elevations, with standard deviations of 11%, 15%, and 15%, respectively.

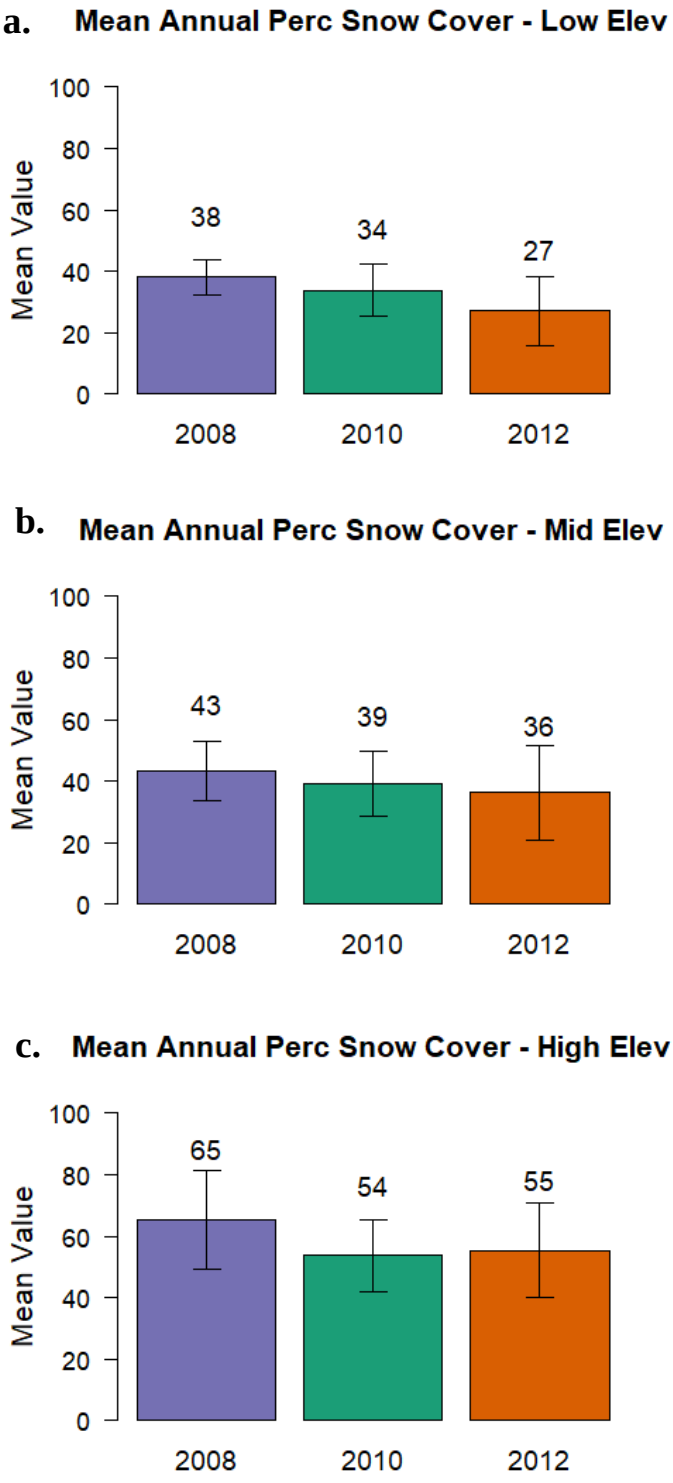
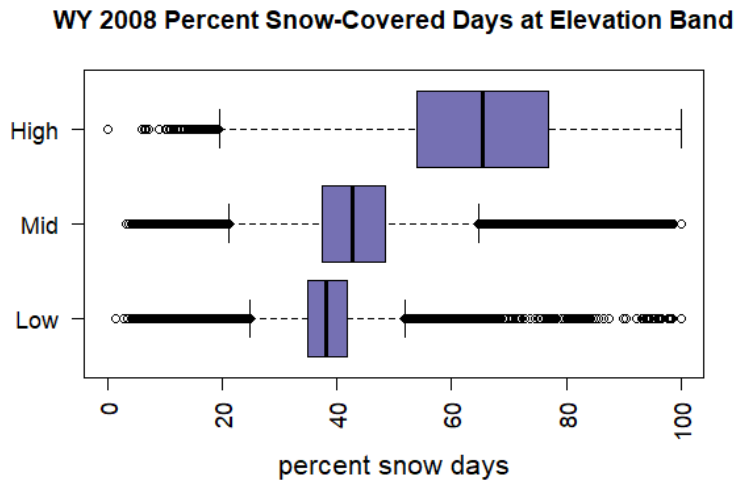


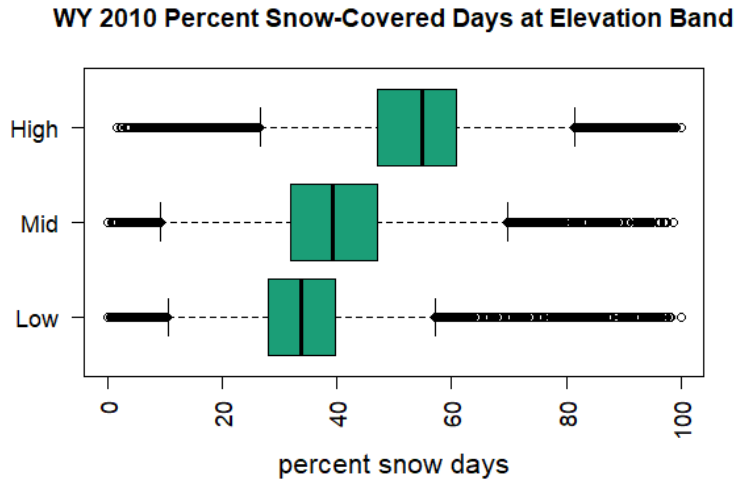
Figure 3.20: Mean annual snow-covered days and their standard deviations (error bars) by water year for a) low elevation band, b) mid elevation band, and c) high elevation band.

Analysis of the numerical distribution of percent annual snow cover by pixel also displayed a clear trend of increasing snow cover percent with elevation (Figure 3.21). This relationship was present for all three analyzed water years. Among these years, the wettest year, WY 2008, had considerably higher variation in percent annual snow cover for high elevations with a difference of 23% between the lower (Q1) and upper (Q3) quartiles (Table 3.2). The driest year, WY 2012, had considerably higher variation in percent annual snow cover for mid-elevations, with a difference of 26% between Q1 and Q3. The difference in percent snow cover is larger between the mid- and high elevations than between the low and mid-elevations. Median percent snow values are 5%-11% higher for mid-elevations than low elevations, whereas median percent snow values are 12%-23% higher for mid-elevations than high elevations.

a.



b.



c.

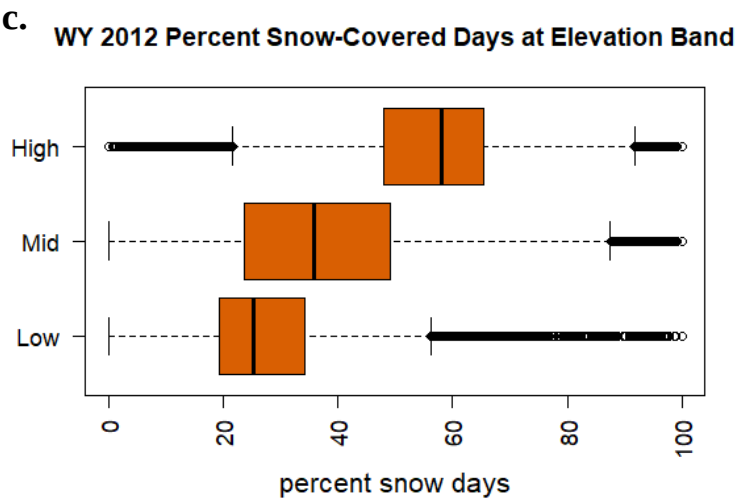


Figure 3.21: Numerical distributions of percent annual snow-covered days by elevation bands for a) WY 2008, b) WY 2010, and c) WY 2012.

Table 3.2: Percent annual snow-covered area metrics by elevation band for all water years.

Elevation	Q1	Median	Q3	Mean	St. Dev.
Water Year 2008					
Low	35	38	42	38	6
Mid	38	43	48	43	10
High	54	66	77	65	16
Water Year 2010					
Low	28	34	40	34	8
Mid	32	39	47	39	11
High	47	55	62	54	12
Water Year 2012					
Low	19	25	34	27	11
Mid	24	36	49	36	15
High	28	48	60	55	15

3.3.2.2 Slope

Mean annual percent snow cover also displayed a positive relationship with increasing slope, although not as pronounced as that observed for elevation (Figure 3.22). The wettest year, WY 2008, had the highest mean percent snow cover value of the three at 44% for low slopes, 49% for mid-slopes, 54% for high slopes, with standard deviations of 13%, 16%, and 19%, respectively. Mean annual percent snow cover for the average water year, WY 2010, was 40% for low slopes, 42% for mid-slopes, 45% for high slopes, with the smallest variations in the data as shown by standard deviations of 12%, 13%, and 14%, respectively. The driest water year, WY 2012, had the lowest mean annual percent snow cover values of 35% for low slope areas, 39% for mid-slope areas, 45% for high slope areas, and again had the largest variations in the data with standard deviations of 16%, 18%, and 19%, respectively.

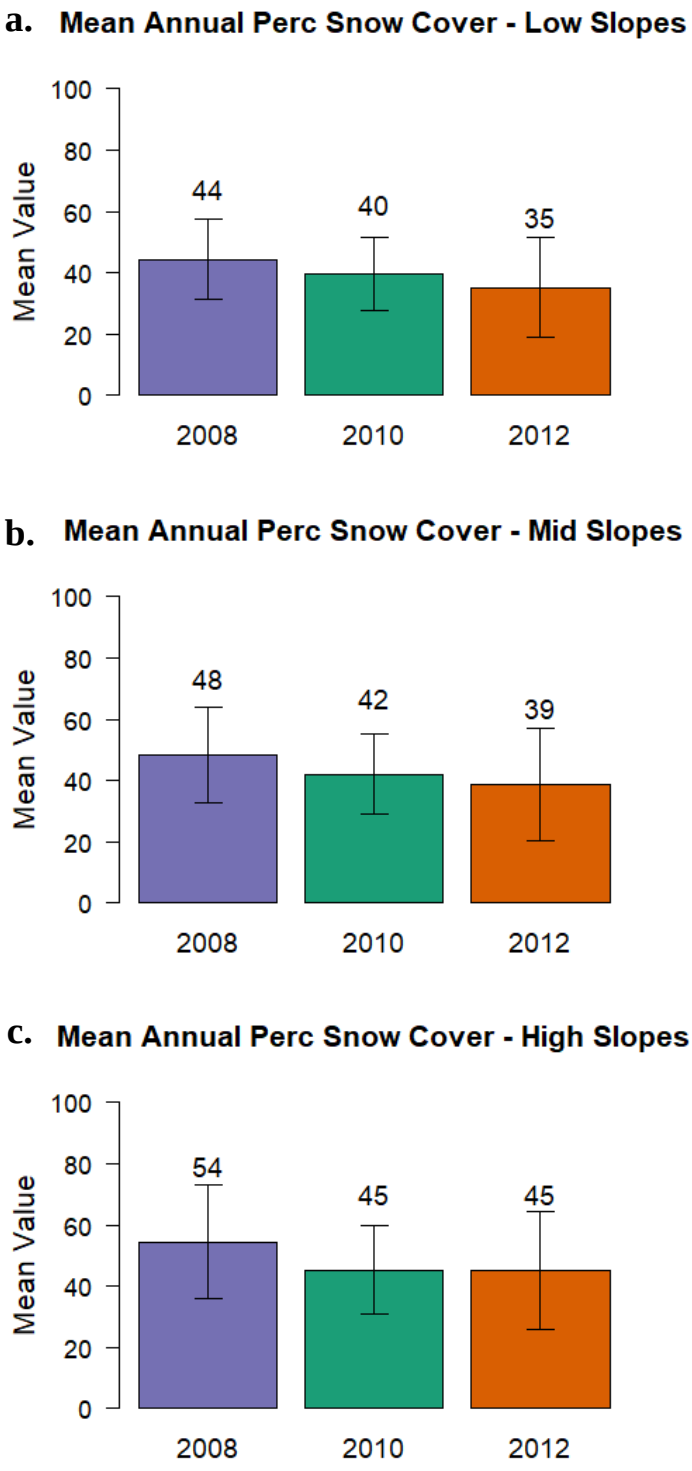


Figure 3.22 Mean annual snow-covered days and their standard deviations (error bars) by water year for a) low slope band, b) mid slope band, and c) high slope band.

Numerical distribution of annual percent snow cover by slope band also showed increasing percent snow cover with increasing slope (Figure 3.23). WY 2012, the dry year, had the most variation in the data, with differences of 25%, 30%, and 32% between Q1 and Q3 for low, mid-, and high slopes, respectively (Table 3.3). As with elevation, the differences in percent snow values between mid- and high slopes were larger than the differences between low and mid- slopes. Median percent snow values were 3-5% higher for mid-slopes than low slopes, and 6%-11% higher for high slopes than mid-slopes. In all water years the high slope bands had the greatest variation in the data compared to other slope bands for their same years. This is possibly because the high slope band contains the largest range of slopes, from 20° to the maximum slope value at 74°.

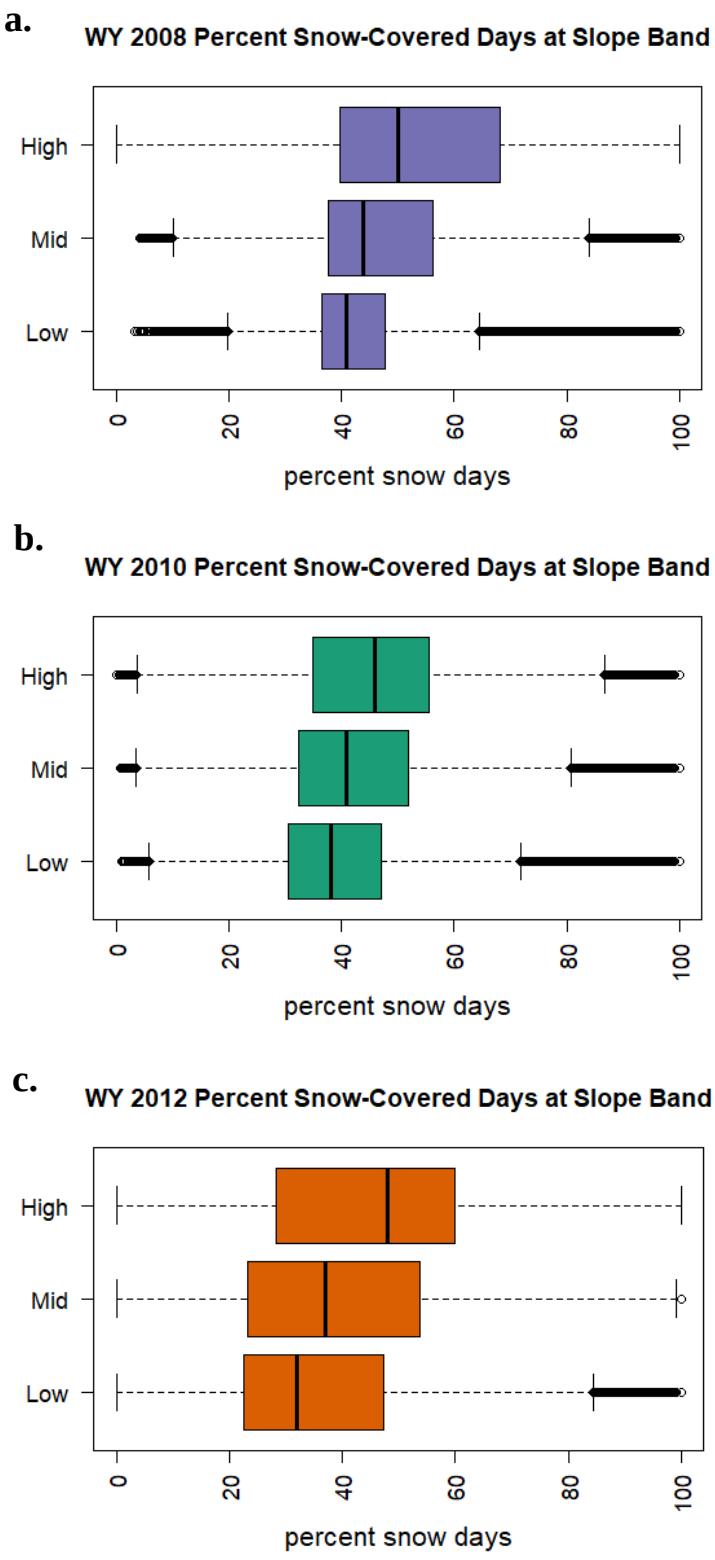


Figure 3.23: Numerical distributions of percent annual snow-covered days by slope bands for a) WY 2008, b) WY 2010, and c) WY 2012.

Table 3.3: Percent annual snow-covered area metrics by slope band for all water years.

Slope	Q1	Median	Q3	Mean	St. Dev.
Water Year 2008					
Low	37	41	48	44	13
Mid	38	44	56	48	16
High	40	50	68	54	19
Water Year 2010					
Low	31	38	47	40	12
Mid	32	41	52	42	13
High	35	46	56	45	14
Water Year 2012					
Low	23	32	47	35	16
Mid	23	37	54	39	18
High	28	48	60	45	19

3.3.2.3 Aspect

Mean annual percent snow cover values show clear, though still relatively weak, relationships between north/east and south/west aspects (Figure 3.24). WY 2008, the year with the highest mean annual percent snow cover, was the only year in which the highest value was found for north aspects (52%), followed by east aspects (51%), west aspects (48%), and its lowest annual value occurred for south aspects (46%). Standard deviations for the above were 18%, 16%, 16%, and 14%, respectively. For WY 2010, however, the mean annual percent snow cover was slightly lower for north aspects (43%) than it was for east aspects (45%), and higher for south aspects (41%) than for west aspects (41%). Standard deviations were 15%, 13%, 12%, and 13% respectively. Mean annual percent snow cover for WY 2012 was the lowest of the three years for north aspects (41%), east aspects (43%), south aspects (37%), and west aspects (38%). This water year also had the largest standard deviations of 20%, 18%, 18%, and 18%, respectively.

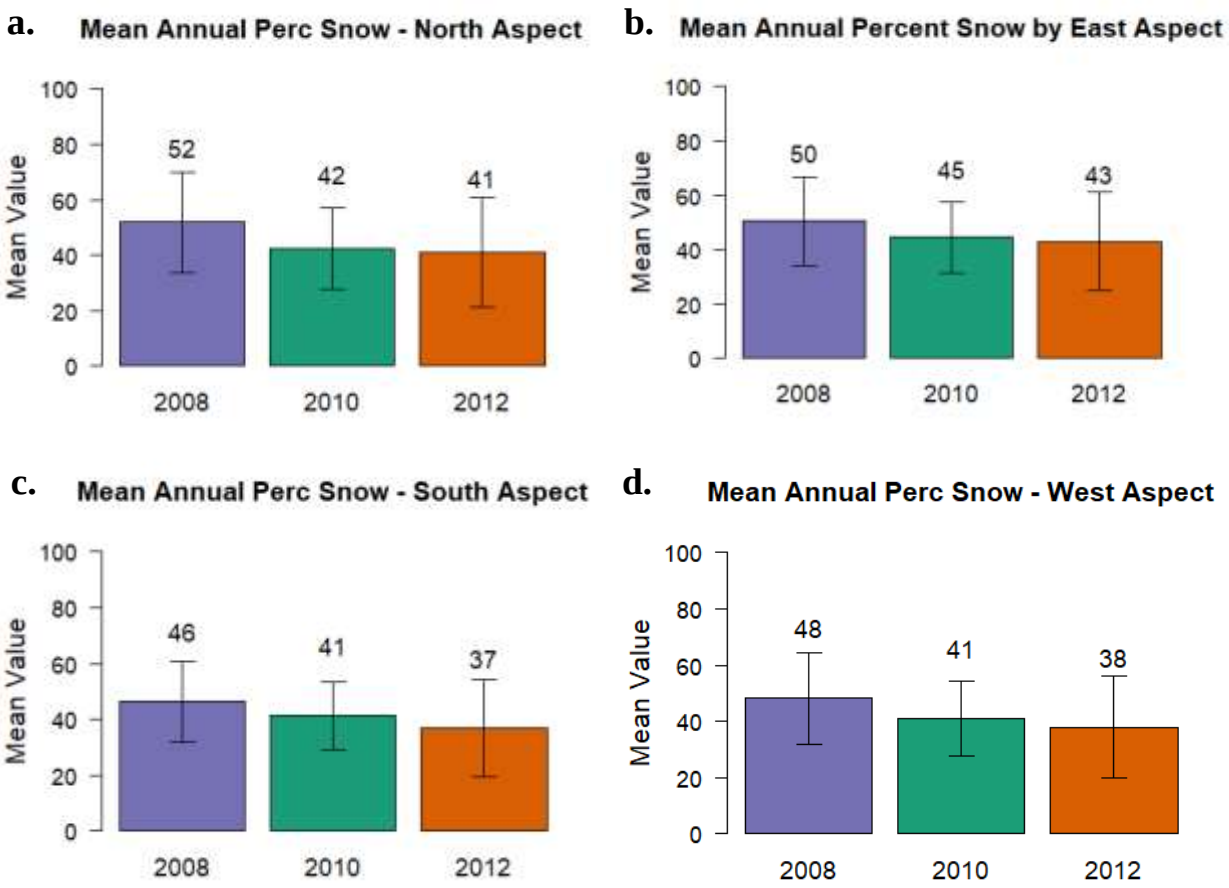


Figure 3.24: Mean annual snow-covered days and their standard deviations (error bars) by water year for a) north aspect, b) east aspect, c) south aspect, and d) west aspect.

Numerical distributions of annual percent snow cover by aspect also showed a weak distinction between north/east aspects and south/west aspects (Figure 3.25). Median values for north/east aspects were 3%-4% higher than south/west aspects for WY 2008, 1%-5% higher for WY 2010, and 4%-9% higher for WY 2012 (Table 3.4). Similar to the distributions of percent snow cover by slope, WY 2012 also had the largest variations in the data of all three water years.

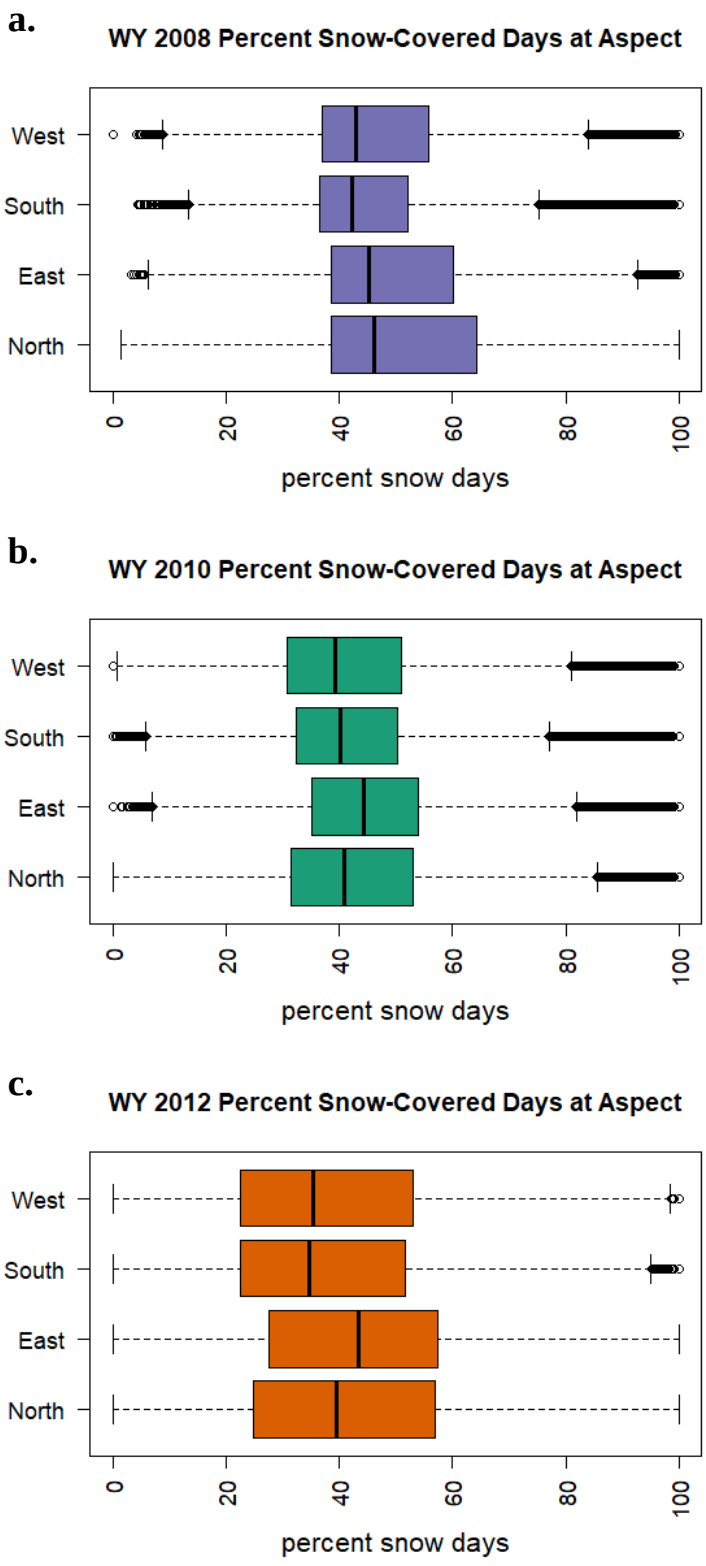


Figure 3.25: Numerical distributions of percent annual snow-covered days by aspect direction for a) WY 2008, b) WY 2010, and c) WY 2012.

Table 3.4: Percent annual snow-covered area metrics by aspect direction for all water years.

Aspect	Q1	Median	Q3	Mean	St. Dev.
Water Year 2008					
North	39	46	64	52	18
East	39	45	60	51	16
South	37	42	52	46	14
West	37	43	56	48	16
Water Year 2010					
North	31	41	53	42	15
East	35	44	54	45	13
South	32	40	50	41	12
West	31	39	51	41	13
Water Year 2012					
North	25	39	57	41	20
East	28	43	57	43	18
South	23	35	52	37	18
West	23	35	53	38	18

3.3.2.4 Land Cover Type

Mean annual percent snow cover varied widely by land cover type, though consistent patterns emerged across water years within each type, with unvegetated or clear areas typically holding the highest values and shrub areas holding the lowest (Figure 3.26). For WY 2008, mean annual percent snow cover was 46% for tree areas, 40% for shrub areas, 52% for grass areas, and 73% for unvegetated or clear areas. Standard deviations for the above were 13%, 10%, 17%, and 16%, respectively. For WY 2010, mean percent annual snow cover was 40% for tree areas, 35% for shrub areas, 47% for grass areas, and 58% for unvegetated or clear areas. Standard deviations for WY 2010 for these land cover types were 12%, 10%, 12%, and 12%, respectively. For the dry WY 2012, mean annual percent snow cover was 36% for tree areas, 29% for shrub areas, 46%

for grass areas, and 62% for unvegetated or clear areas. WY 2012 standard deviations for the above were 17%, 14%, 17%, and 13%, respectively.

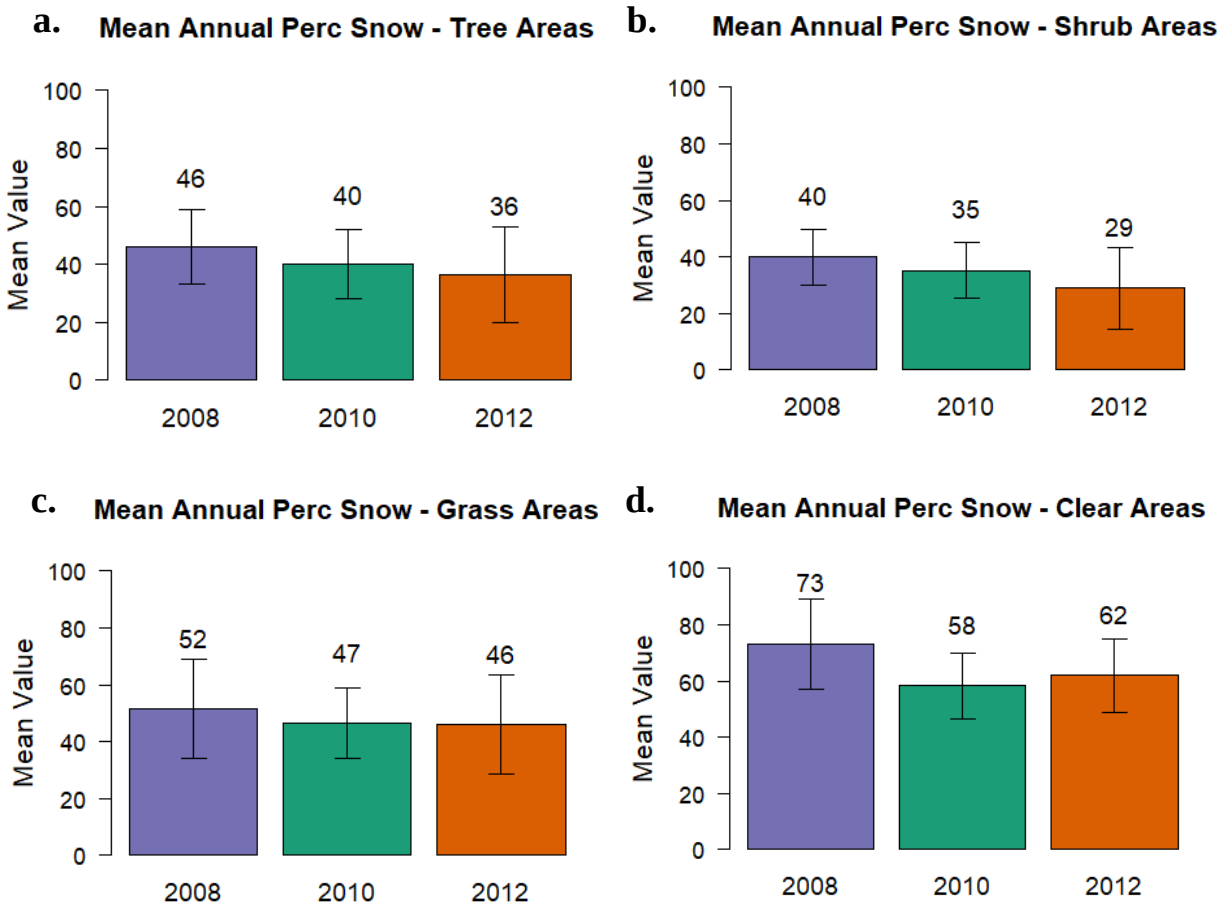


Figure 3.26: Mean annual snow-covered days and their standard deviations (error bars) by water year for a) tree areas, b) shrub areas, c) grass areas, and d) clear areas.

The numerical distributions of annual percent snow cover by land cover type varied widely but also displayed consistent patterns between land cover type for all water years (Figure 3.27). Percent snow cover for clear/unvegetated areas was consistently the highest of all land cover types, with median values of 75%, 59%, and 63% for WYs 2008, 2010, and 2012, respectively (Table 3.5). Conversely, percent snow cover for shrub

areas was consistently the lowest of all land cover types, with median values of 38%, 34%, and 25% for WYs 2008, 2010, and 2012, respectively. We believe the patterns between these two land cover types is largely due to where they are primarily located in the landscape. Most clear/unvegetated areas are at high elevations above tree line, whereas most shrub areas are at lower elevations. Grass and tree areas are dispersed evenly throughout the watershed at a variety of elevations, aspects, and slopes, which may help to explain the generally larger variations in the data for these land cover types.

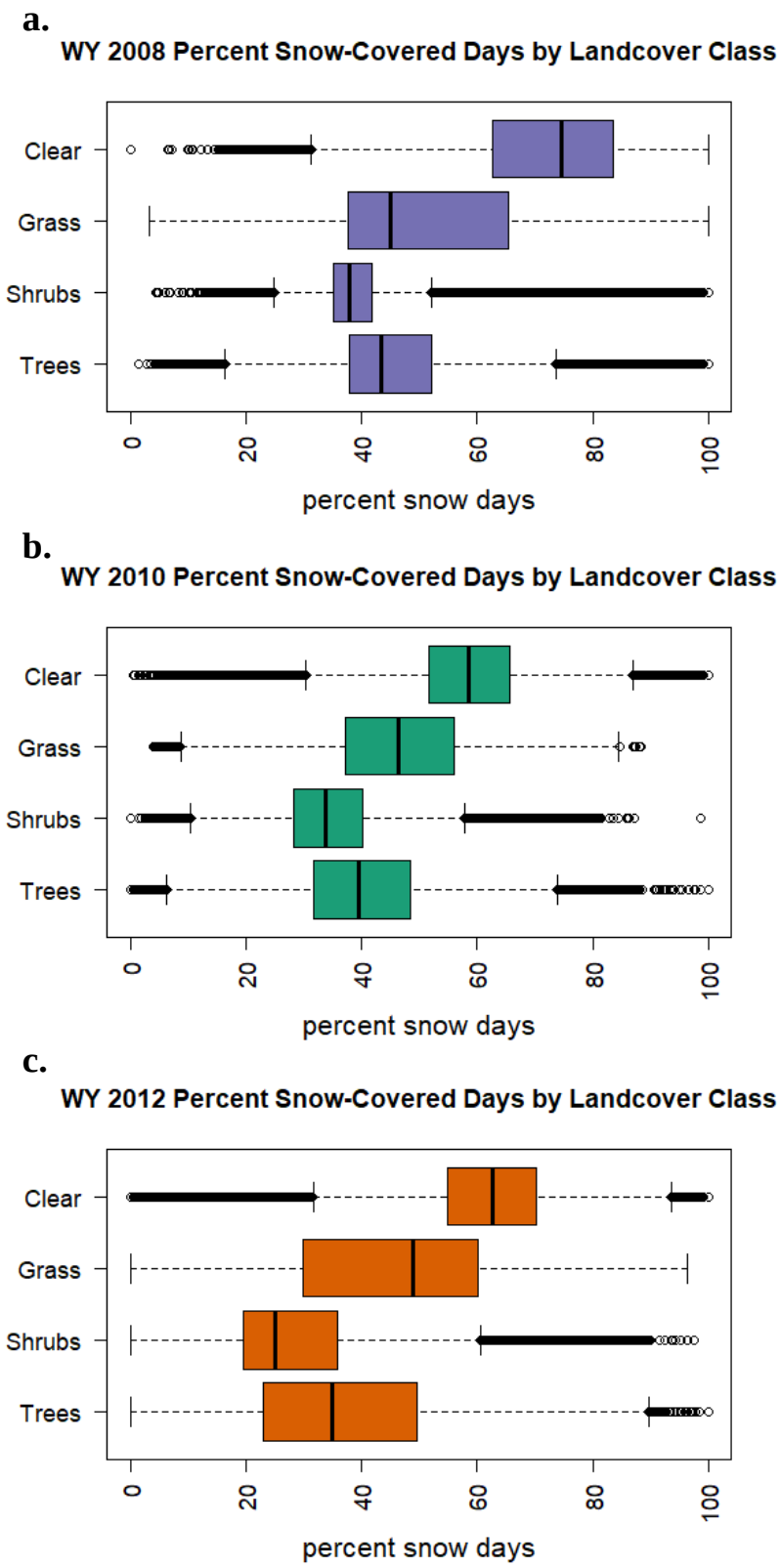


Figure 3.27: Numerical distributions of percent annual snow-covered days by land cover type for a) WY 2008, b) WY 2010, and c) WY 2012.

Table 3.5: Percent annual snow-covered area metrics by land cover type for all water years.

Land Cover Type	Q1	Median	Q3	Mean	Std. Dev.
Water Year 2008					
Trees	38	44	52	46	13
Shrubs	35	38	42	40	10
Grass	38	45	65	52	17
Clear	63	75	84	73	16
Water Year 2010					
Trees	32	40	49	40	12
Shrubs	28	34	40	35	10
Grass	37	47	56	47	12
Clear	52	59	66	58	12
Water Year 2012					
Trees	23	35	50	36	17
Shrubs	20	25	36	29	14
Grass	30	49	60	46	17
Clear	55	63	70	62	13

3.3.2.5 Spatial Snow Cover by Water Year

Snow cover patterns across landscape characteristics predictably trended within the relative amount of precipitation received (as represented by max SWE amounts) that year. Mean and median percent snow cover were highest for the wet WY 2008 and lowest for the dry WY 2012 in almost all instances. Exceptions included: mean annual percent snow cover at high elevations, where the dry WY 2012 was 1% higher than the average WY 2010; median annual snow cover at high slopes, where WY 2012 was 2% higher than WY 2010; mean annual snow cover for clear/unvegetated areas, where WY 2012 was 4% higher than WY 2010; and median annual snow cover clear/unvegetated areas, where WY 2012 was higher than WY 2010 by 4%. Median annual snow cover values for grass areas were highest for the dry WY 2012 at 49% and lowest for the wet WY 2008 at 45%.

3.3.2.6 In-Depth Elevation Analysis

This initial analysis presented data for percent snow cover distributions for slope that looked markedly similar to distributions for elevation (Figures 3.20 and 3.22), raising the question of potentially confounding variables. We therefore sought to explore whether SCA for these landscape characteristics were in fact primarily influenced by elevation, and whether a relationship with elevation existed for landscape characteristics whose data distributions had large variability. For example, if elevation had a greater influence on percent annual snow cover than did slope, we expected to see increasing annual percent snow cover by elevation within each slope band. If this was not the case, then annual percent snow cover values within each slope band should be relatively similar regardless of the elevation where they were located.

To do this, we isolated slope bands, hillslope aspects, and land cover types by elevation band and plotted the distribution of percent annual snow cover for each. The distribution of annual percent snow cover by low (Figure 3.28), mid- (Figure 3.29), and high (Figure 3.30) slopes was similar for all water years with increasing snow cover by elevation regardless of slope category. The distribution of annual percent snow cover by aspect was highly variable and showed very weak relationships between snow cover and hillslope aspect (north and south aspects shown in Figures 3.31, 3.32, and 3.33. east and west aspects shown in Appendix A). However, when we examined the distribution of percent annual snow cover for all aspects, we saw clear trends of increasing snow cover with elevation, regardless of water year.

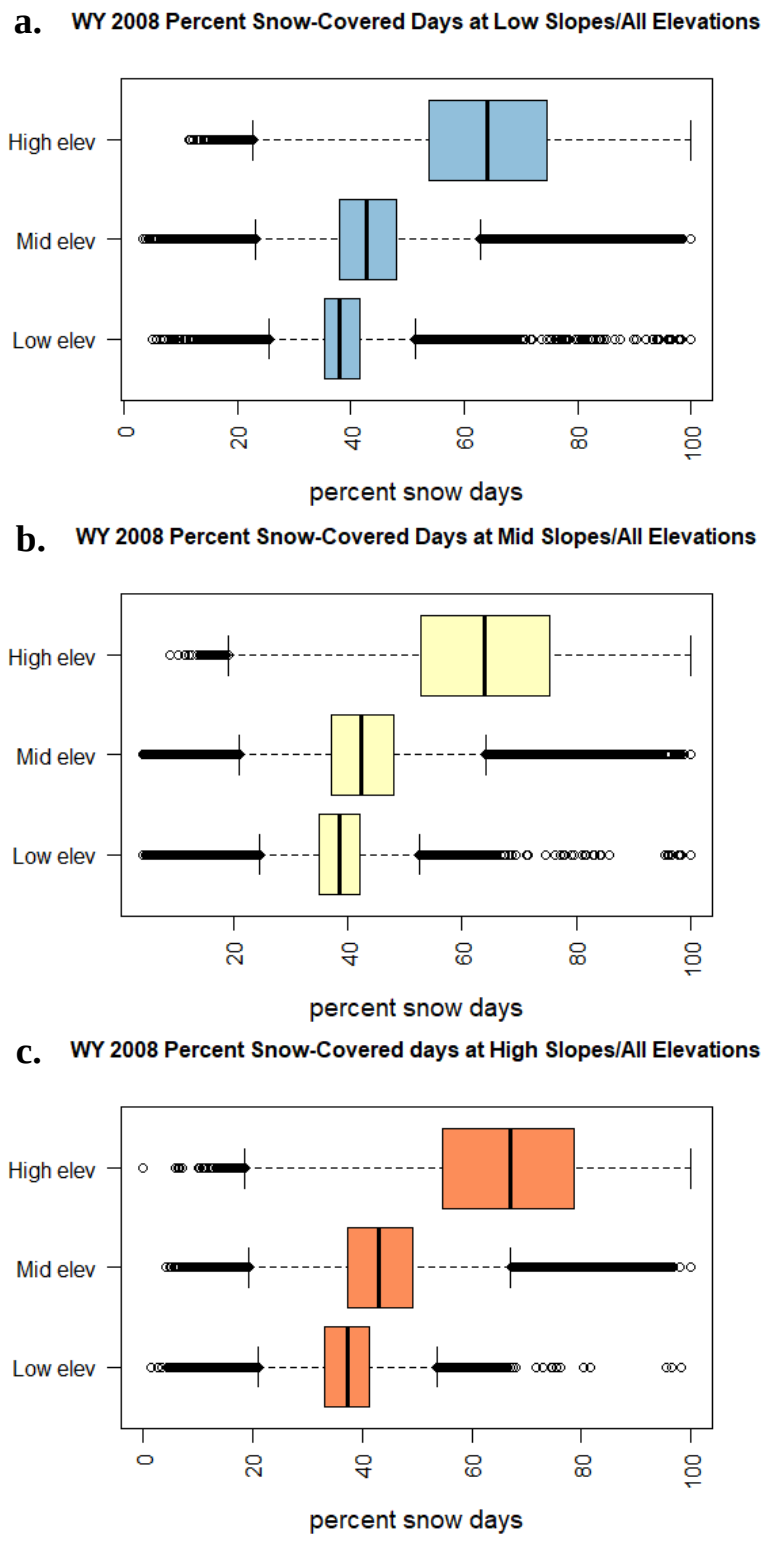


Figure 3.28: Numerical distributions of WY 2008 percent annual snow-covered days by elevation band for a) low slopes, b) mid-slopes, and c) high slopes.

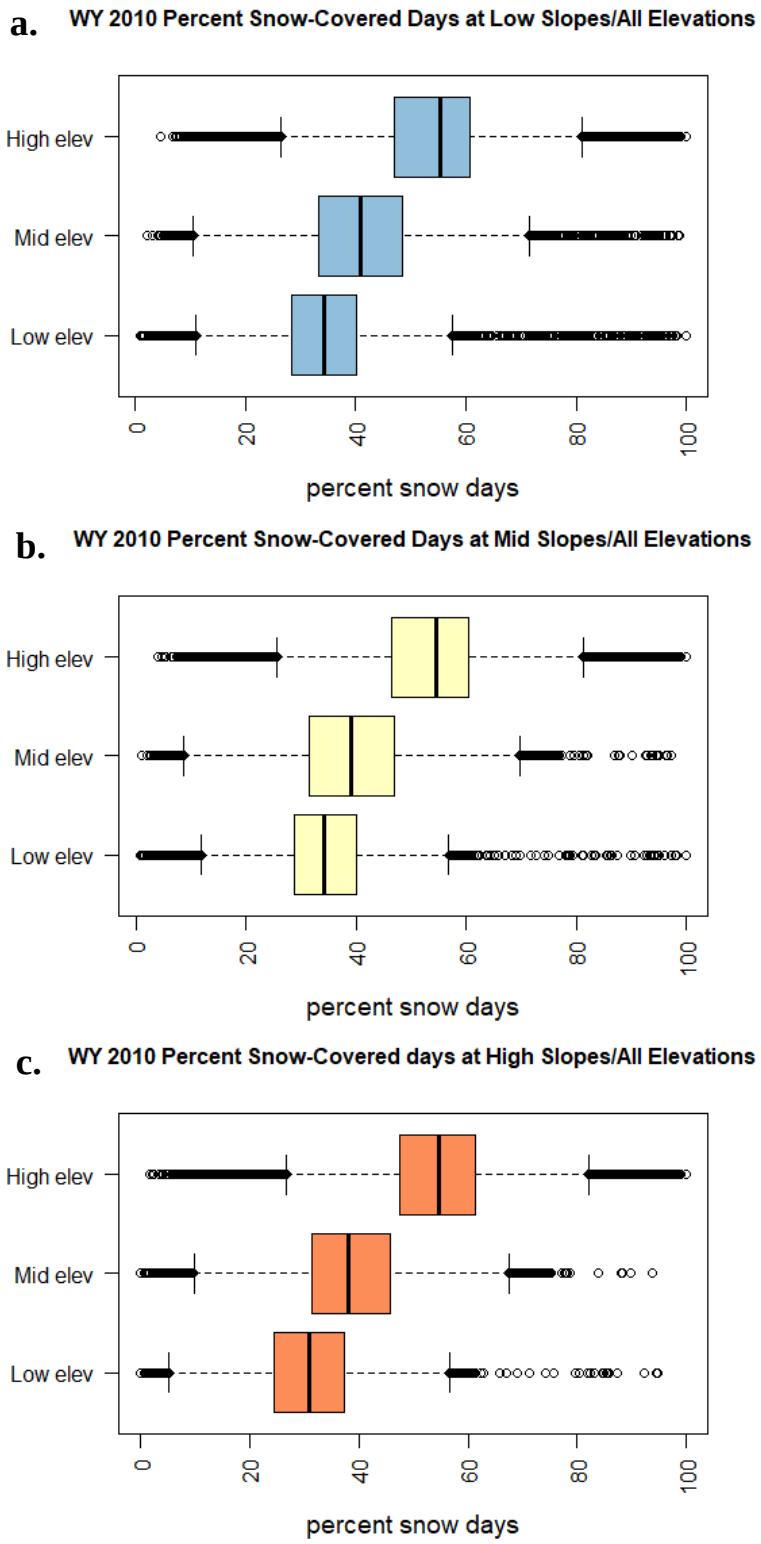


Figure 3.29: Numerical distributions of WY 2010 percent annual snow-covered days by elevation band for a) low slopes, b) mid-slopes, and c) high slopes.

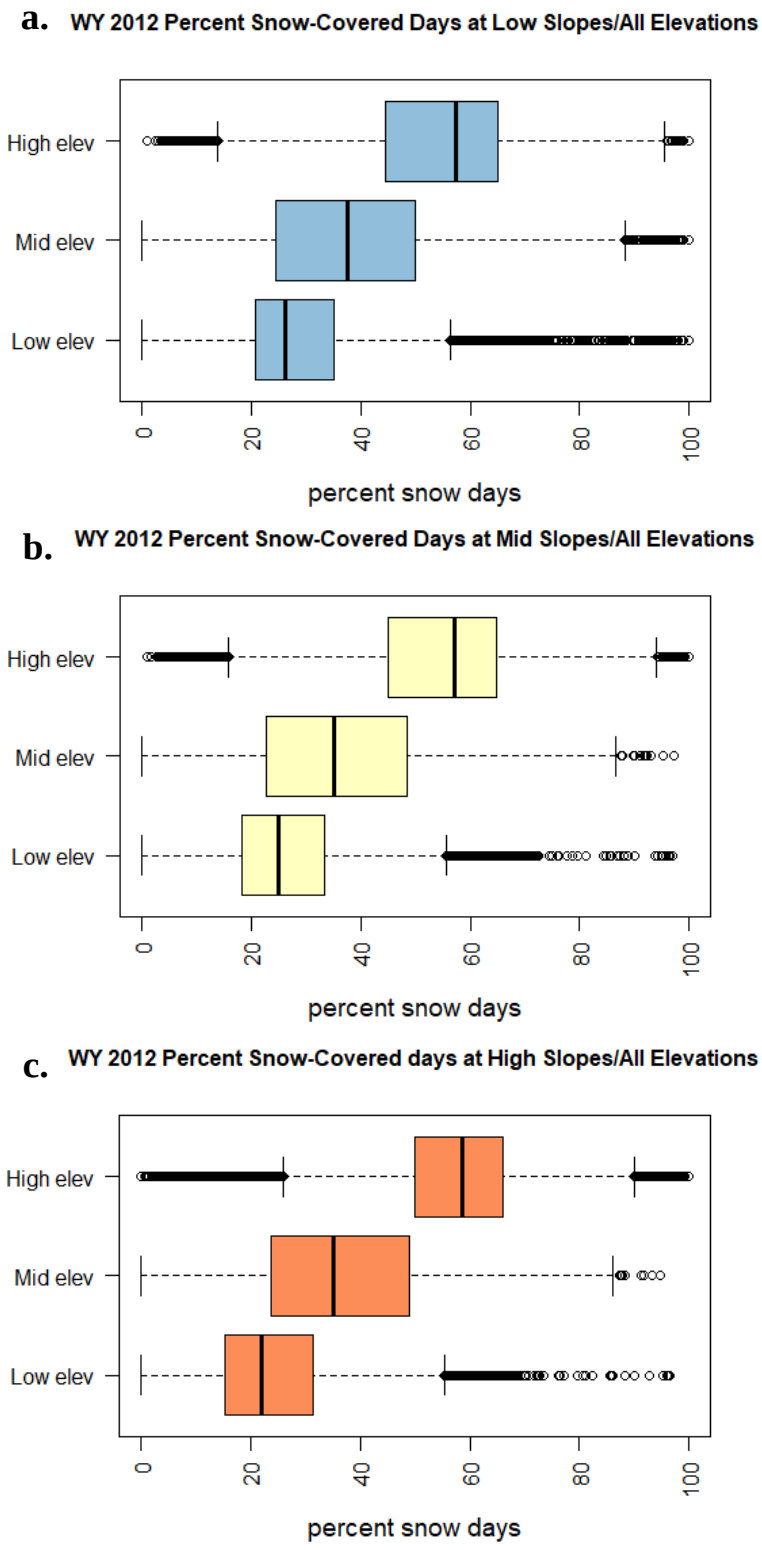


Figure 3.30: Numerical distributions of WY 2012 percent annual snow-covered days by elevation band for a) low slopes, b) mid-slopes, and c) high slopes.

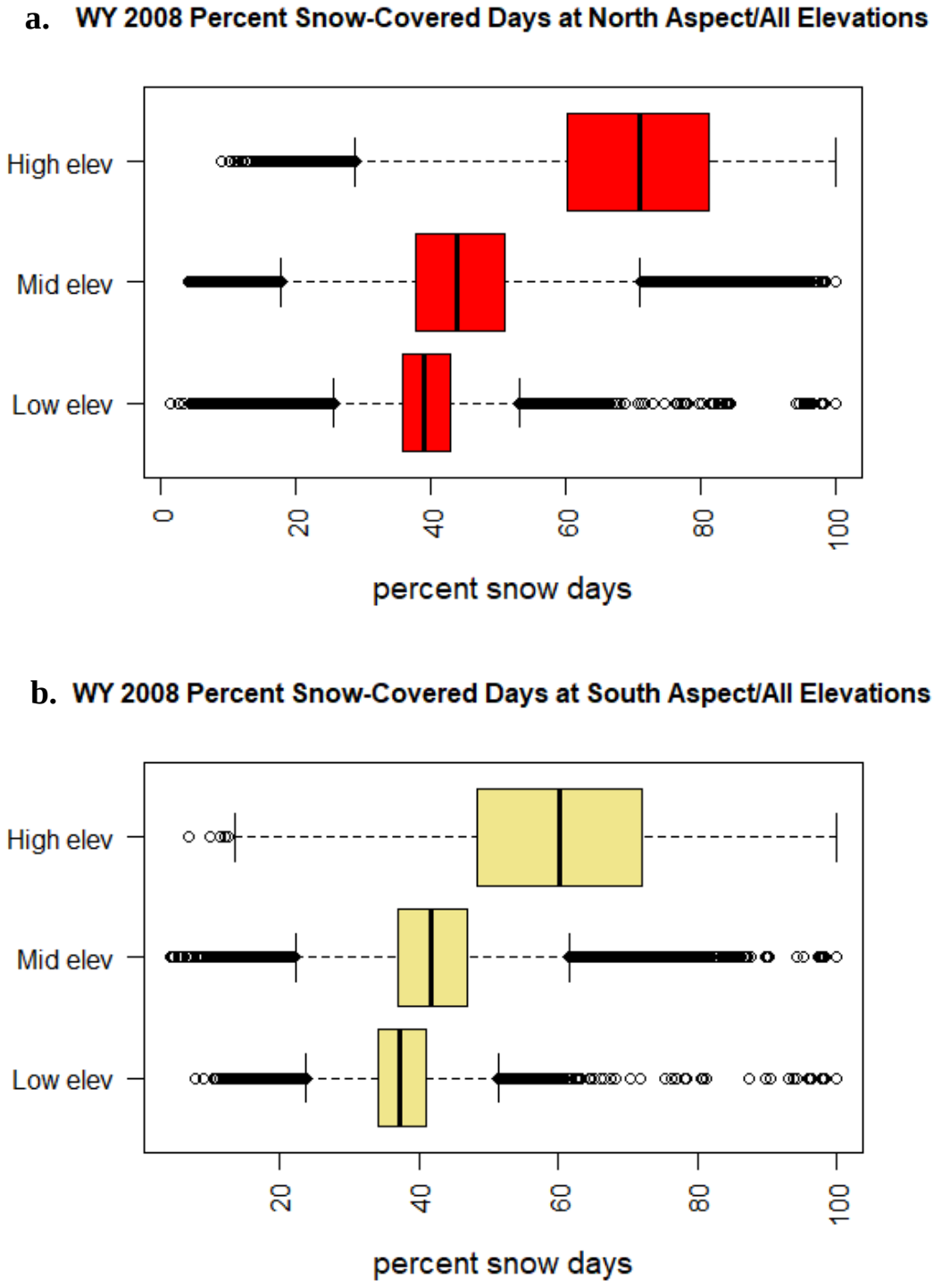
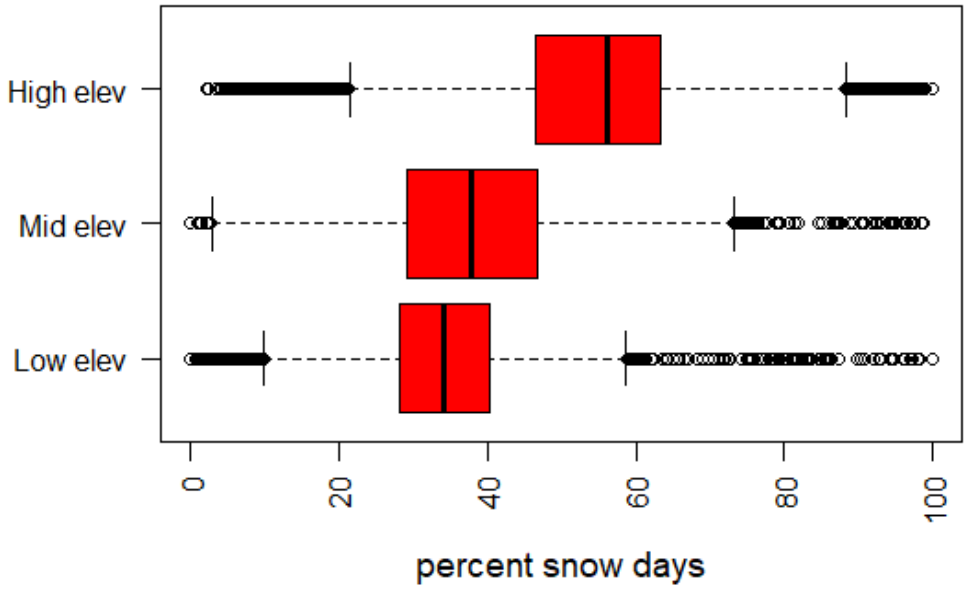


Figure 3.31: Numerical distributions of WY 2008 percent annual snow-covered days by elevation band for a) north aspects and b) south aspects.

a. WY 2010 Percent Snow-Covered Days at North Aspect/All Elevations



b. WY 2010 Percent Snow-Covered Days at South Aspect/All Elevations

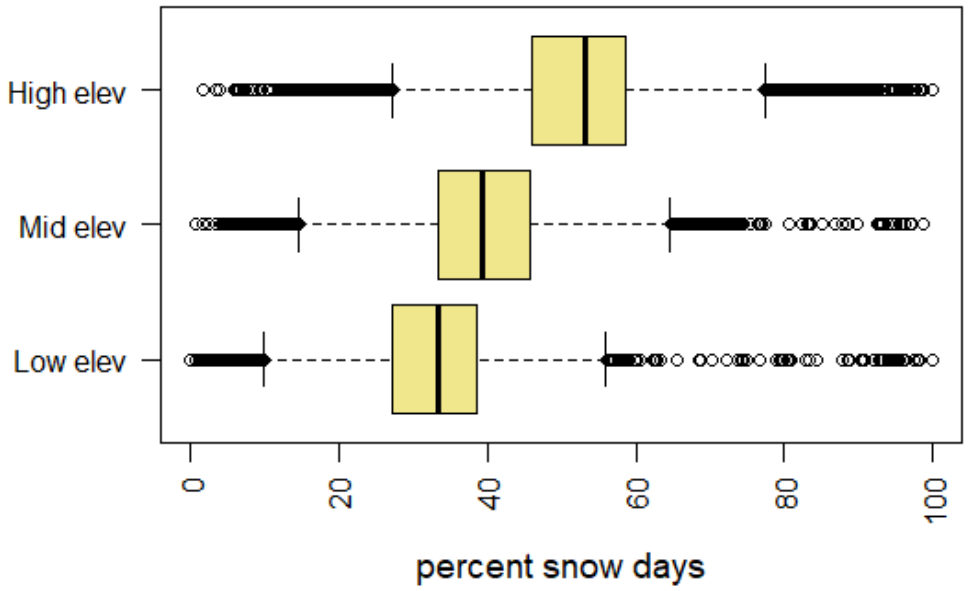


Figure 3.32: Numerical distributions of WY 2010 percent annual snow-covered days by elevation band for a) north aspects and b) south aspects.

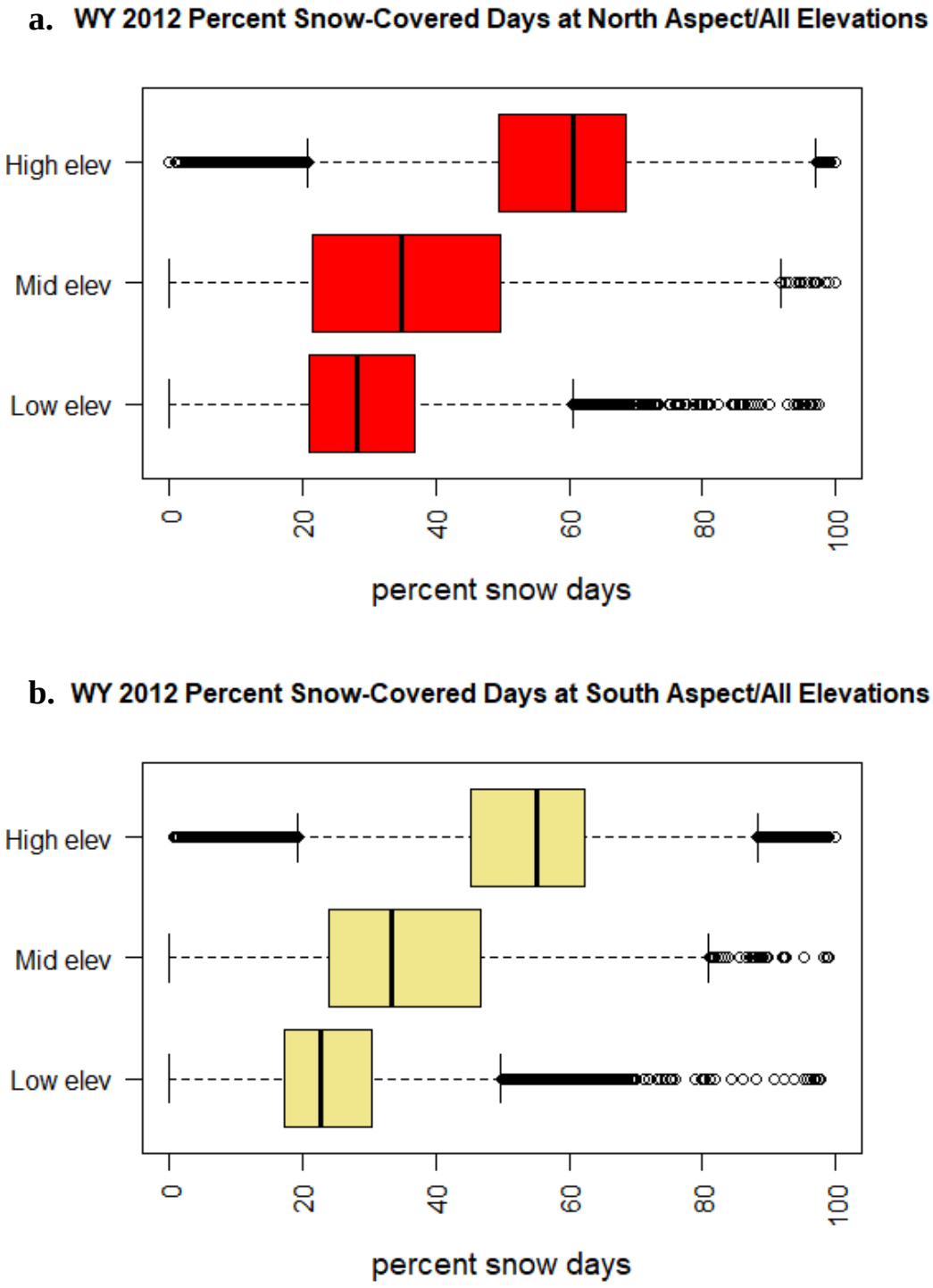


Figure 3.33: Numerical distributions of WY 2012 percent annual snow-covered days by elevation band for a) north aspects and b) south aspects.

We also examined annual percent snow cover distribution for tree and grass areas by elevation band (Figures 3.34, 3.35, and 3.36). These two land cover types had larger variabilities in percent annual snow cover and are most widely distributed among elevation, slope, and aspect categories in the watershed. Additionally, we examined the annual percent snow cover for shrub and clear/unvegetated landcover types (Appendix A). We saw apparent trends of increasing snow cover within all land cover types with elevation regardless of water year.

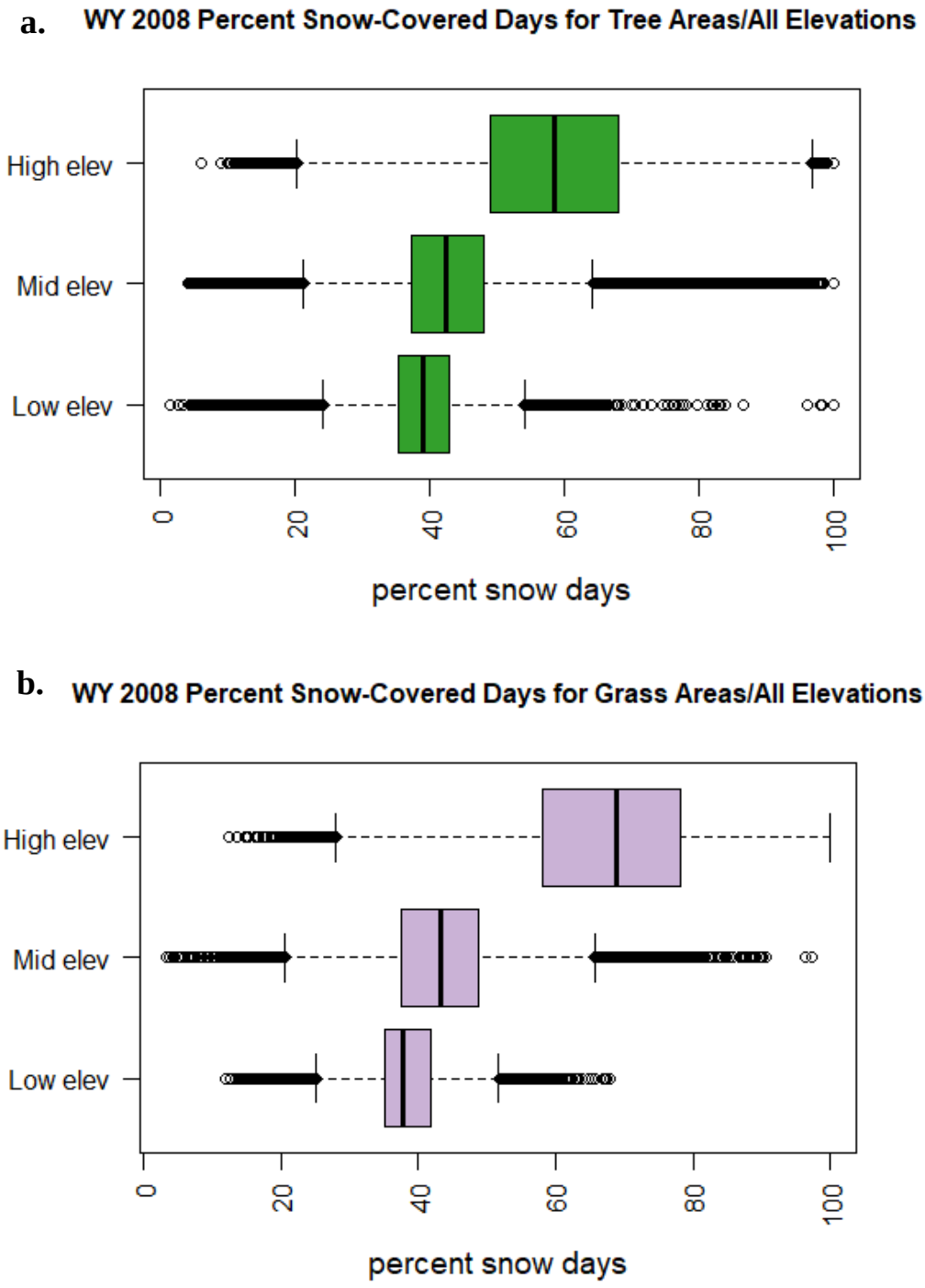


Figure 3.34: Numerical distributions of WY 2008 percent annual snow-covered days by elevation band for a) tree areas and b) grass areas.

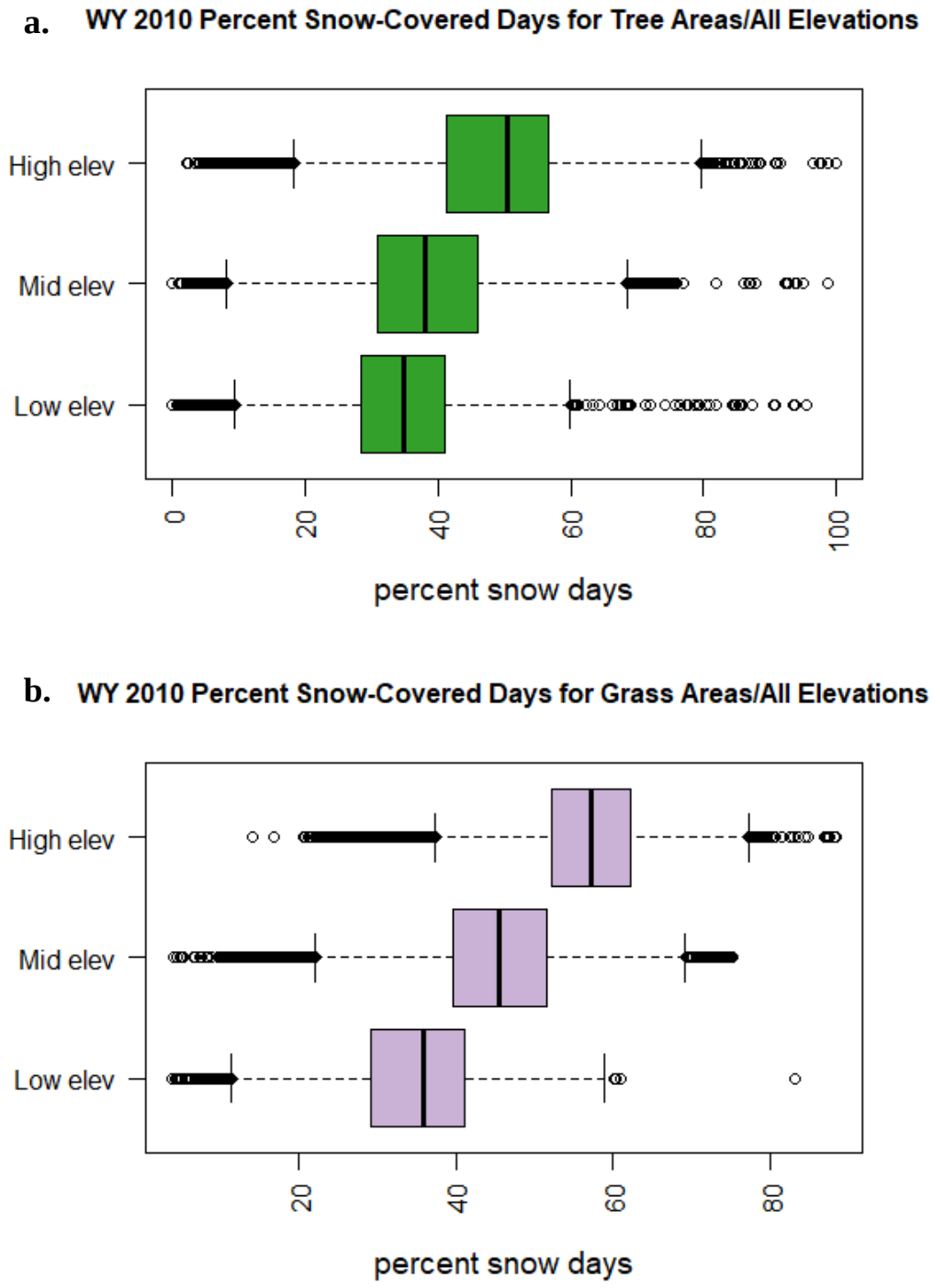


Figure 3.35: Numerical distributions of WY 2010 percent annual snow-covered days by elevation band for a) tree areas and b) grass areas.

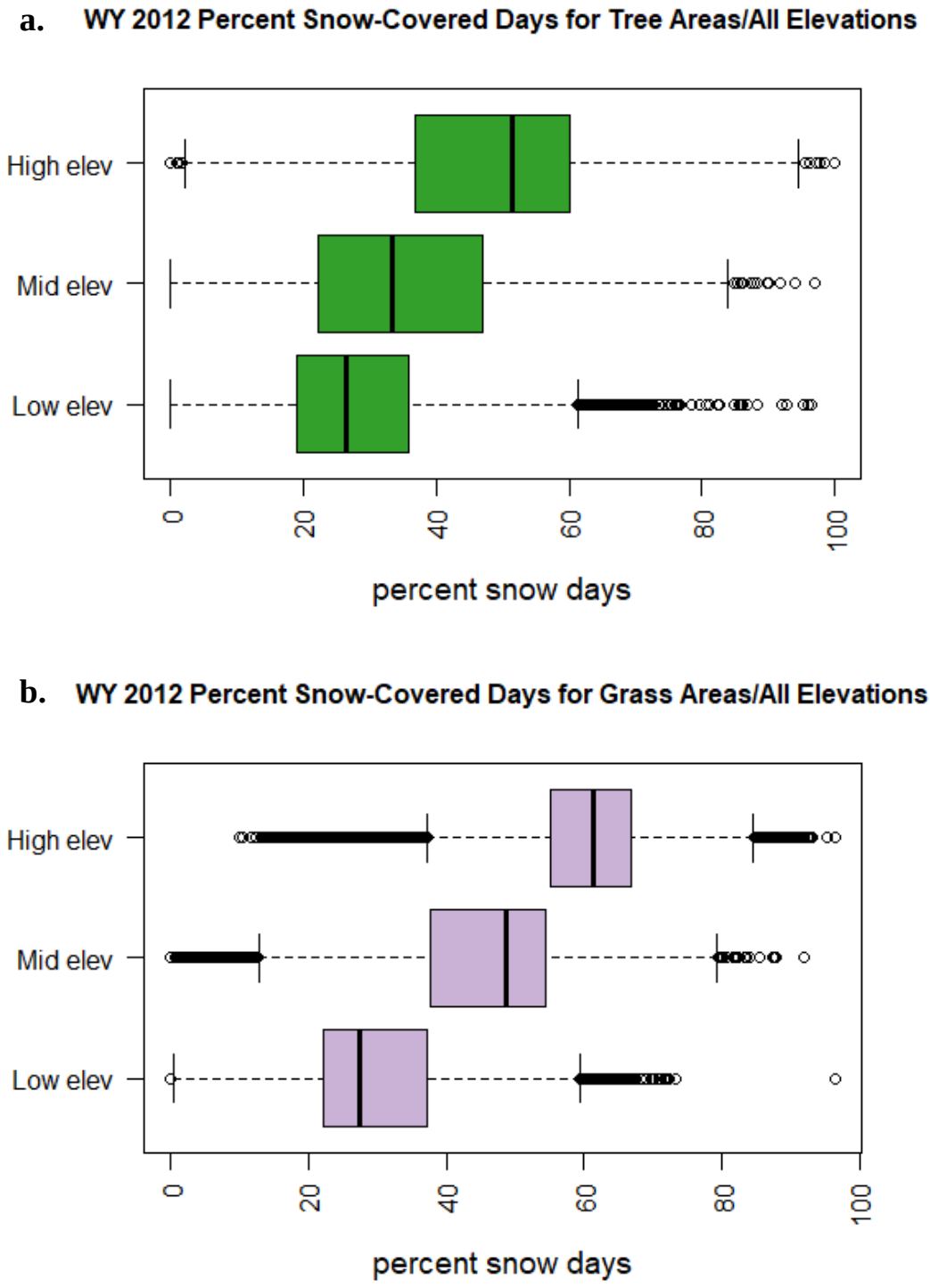


Figure 3.36: Numerical distributions of WY 2012 percent annual snow-covered days by elevation band for a) tree areas and b) grass areas.

3.3.3 SCA and Stream Discharge Data

The patterns of stream discharge at the watershed outlet exhibited a clear relationship to annual peak SWE, and thus fit with our initial categorizations from the SNOTEL annual peak SWE data. The wet water year of 2008 had not only the highest peak discharge at 2710 cfs, but also the widest hydrograph and the highest volume of discharge for the three years at 157,325.8 cfs. Average WY 2010 had a similar peak discharge at 2310 cfs but a much narrower hydrograph, resulting in a total volume of 95,444.9 cfs. The dry WY 2012 had a significantly lower peak discharge at 570 cfs and a total volume of 51,118.1 cfs.

The time lags between the last day the watershed had 50% or higher snow cover and the center of mass of the hydrograph were 45 days for WY 2008, 126 days for WY 2010, and 56 days for WY 2012. We suspect the time lag for WY 2010, however, is artificially high due to increased cloud cover from days 120-190 during this time period.

3.4 Discussion

3.4.1 STARFM 20-Year Dataset

In this study we used STARFM to create a daily, 30-m resolution dataset that can be used to track seasonal snow cover changes over multiple years. To our knowledge, this is the first time STARFM has been used in this manner. The ability to create a multi-decadal, high resolution dataset of this type to monitor snow cover has major implications for snow forecasting and hydrologic modeling, including those hydrologic models that seek to better understand timing and delivery of water to the CZ.

We acknowledge this dataset has limitations that future users must take into consideration. As discussed at length in Section 2.4, cloud cover remains a major

limitation, as does the availability of high temporal resolution data for downscaling. Gaps in our high temporal, low spatial resolution dataset (MODIS) reduced the model's ability to synthesize fine spatial resolution predictions for missing dates. Over longer periods of time, however, the impact that these data gaps have on calculations of mean annual snow cover is reduced. When the period of analysis is longer than 5 years, irregularity of data due to cloud-masking amounts to a semi-random sampling of cloud-free scenes from early, middle, and late times of each month for dry, average, and wet snow years (Selkowitz et al., 2017). This assumes, however, that the cloud mask is not biased towards misidentifying snow as clouds, and that any gaps in data are indeed random, and not concentrated during certain times of the year.

Future work using our dataset or others like it should also be aware of potential errors that may result from inconsistencies involved with using multiple Landsat datasets. Bandwidths and band properties vary between generations of Landsat sensors (Chastain et al., 2019). In addition, there are gaps in Landsat 7 ETM+ imagery after May 2003 due to failure of the Scan-Line Corrector (SLC) (Markham et al., 2004). While methods exist to gap-fill these missing areas in ETM+ imagery (Roy et al., 2008; Chen et al., 2011), we chose not to gap fill our dataset. We instead took advantage of the Landsat 5 TM data overlap from 2003-2011 and used ETM+ data with uncorrected SLC errors only for time periods where it was the only Landsat data available. While not included here due to the scope of the study and the goal of providing proof-of-concept of an "off-the-shelf" method for STARFM, performing data calibration across sensors or choosing to gap fill data are additional steps future users may choose to do depending on their accuracy and data application needs.

Finally, we also recognize the limitations introduced by the methods we used to evaluate seasonal snow cover. As discussed in Section 2.4, we acknowledge that using binary identification for snow cover tends to over or underestimate snow cover within individual pixels (Elder et al., 1998; Rittger et al., 2013) even though we did not explicitly evaluate those errors here. In addition, due to the preliminary nature of our analysis, we only evaluated trends in SCA inter-annually by finding the percent of the year that each pixel was classified as "snow-covered". In the following discussion we offer possible explanations as to the differences observed in snow cover patterns between years, some of which may be due to seasonal variations. Although we did not perform any investigations regarding intra-annual snow cover, this analysis could easily be accomplished in future work with the dataset we created.

3.4.2 Spatial Snow-Cover Analysis of STARFM Results

When we examined relationships between annual percent snow cover and the landscape characteristics of elevation, slope, aspect, and land cover type, we found that elevation had the strongest relationship to annual percent snow cover. There were consistent trends of increasing SCA with elevation for all water years, regardless of the differences in annual peak SWE (Figure 3.1). We conclude that elevation differences are the primary driver of SCA differences in our study area, in part because many other landscape characteristics are themselves highly correlated with elevation (e.g. slope and land cover type). Analyzed slope bands had a spatial distribution that was very similar to the spatial distribution of the elevation bands of our study area (Figures 2.14 and 2.15), and trends in percent snow cover by slope band were similar to those observed for percent snow cover by elevation band (Figures 3.20 and 3.22). When we separately

considered percent snow cover for low, mid-, and high slopes separated by elevation band, we confirmed that elevation does indeed have a strong influence on SCA regardless of slope.

Major land cover types and their snow cover patterns displayed traits that suggested they are also controlled by elevation. For example, clear and unvegetated areas, the land cover type with the greatest percent snow covered days across all water years, were primarily located in the high elevation areas of the watershed. This is consistent with findings from Carroll et al. (2018), who evaluated the spatial distribution of vegetation types in the East River SFA and found that barren conditions dominated at elevations of 3,650-m and above. Conversely, shrub areas, which are located predominantly in the lowest elevations of the watershed (Figures 2.14b and 2.17a), have the lowest percent annual snow cover of all land cover types for all water years.

We also found patterns in the variation of percent annual snow cover by landscape attribute. For example, percent snow cover had consistently larger variation in high slope bands than in low and medium bands. This may be because the majority of the slopes in our study area (66%) that fall into our low and mid-slope bands (Figure 2.17a) are 20° or less. This means that our high slope band encompasses a much wider relative range of slope degrees (20-74°). Slopes at either end of this range behave very differently in terms of snow accumulation. According to McClung & Schaerer (2006), slab-avalanches typically occur on slopes of 25-55°, redistributing snow to flatter areas, and in some cases exposing bare ground on the slope from which the avalanche originated. Above 55°, sluffing (the falling of loose snow) typically keeps snow from accumulating, sometimes to the point where these slopes can remain largely snow-free (Farinotti et al.,

2010). These differences in snow accumulation due to slope degree likely explain the larger variations in our data for high slope areas.

We were surprised to find a relatively relationship between annual percent snow cover and aspect. Though north and east aspects had slightly higher mean annual percent annual snow cover than south and west aspects, the differences in most cases were less than 5%. The standard deviation and interquartile values of percent annual snow cover by aspect indicate relatively high variability in the data, meaning that there is little consistency in snow cover trends due to aspect. When we further examined percent snow cover by aspect and separated each aspect direction by elevation band, we again saw a very clear trend of increasing snow cover with elevation for all aspect directions.

We suspect this lack of a clear relationship between annual percent snow and aspect is largely due to our examination of snow cover at the annual timescale. Differences in solar insolation throughout the winter influences patterns of spatial snow distribution resulting in shallower snow depths (Buhler et al., 2016) and preferential ablation (Anderson et al., 2014) on south/west aspects in the Northern Hemisphere during the melting season. An isolated analysis of STARFM snow cover data from the ablation season would likely reveal these characteristics, and thus display a stronger relationship between aspect and percent snow cover.

3.4.2.1 Spatial Snow Cover by Water Year

A noteworthy observation regarding our annual percent snow cover results was that the amount of variability in the data for each year was consistent relative to the others across all landscape characteristics. Primarily, the data for WY 2012, the driest year, displayed the largest variability in percent snow cover of the three years for all

landscape characteristics. A possible explanation for this variability is that the driest water year is likely to have the lowest snow depths, and thus when melting occurs, areas of bare ground under shallow snow will be exposed earlier. This would create variability by increasing heterogeneity in the landscape earlier in the season, variability that would be captured within the elevation band or hillslope aspect data. This theory agrees with the findings of Anderton et al. (2004), who found that spatial patterns of snow disappearance were largely due to differences in snow depth when elevation range was small. While we did not consider snow depth measurements in our study, we can reasonably assume that lower annual peak SWE measurements also indicate lower annual snow depths. Similarly, snow cover may not appear until later at the beginning of the snow season during a dry year relative to other years, also increasing variation in snow cover data.

3.4.3 SCA and Stream Discharge Data

The measurements of the time lag between the last day the SCA of the entire watershed was measured at 50% or higher and the center of mass of the hydrograph provide initial estimates as to the timescales and potential process pathways by which snowpack becomes discharge at the watershed outlet. In their study of seasonal groundwater and solute fluxes in the East River SFA for WY 2016 (an average water year), Carroll et al. (2018) determined that snowmelt contributed to 69% and 74% of the volume of the rising and falling limbs, respectively, of the annual hydrograph. The longer the time lag between snowmelt and the hydrograph peak, the more likely it is that the water from snowmelt is traveling through longer pathways in the subsurface as opposed to becoming surface runoff that quickly makes its way to the stream. Another recent study in the East River SFA found that lower snowpack years with slower snowmelt

resulted in greater percolation and groundwater recharge, and consequently stream discharge that was more influenced by deeper groundwater flows as opposed to lateral shallow surface flows during wetter years (Tokunaga et al., 2019). Low snow years can also influence streamflow when snowmelt volumes are low and slow enough that the yearly hydrograph peaks due to summer rains rather than snowmelt (Knowles et al., 2015).

These measurements of time lag are also imperfect due to uncertainties in the SCA data from cloud cover. Therefore, using the methods presented here, we are unable to confirm whether the last date observed with 50% or more SCA is the latest date in the water year with this amount of snow cover. Gap-filling methods of various complexity can be employed to fill in missing data due to cloud cover gaps and increase confidence that the observed SCA in the data represents SCA conditions on the ground. Sproles et al. (2018) developed a set of conditions for identifying cloud-covered MODIS images as snow-covered or non-snow-covered depending on the status of the clear images immediately preceding and following the cloud-covered images. Dozier et al. (2008) developed an algorithm that smooths SCA data along the time axis, considering trends in total SCA over time instead of a three dimensional or a neighboring pixel interpolation. This method allows all images in a time series to be considered regardless of cloud cover. Recent advancements in cloud gap-filling methods using machine learning allows for more efficient assessments of multiple criteria that influence the assignment of snow status to a cloud-covered pixel. Criteria such as spectral similarity among pixels, snow distribution continuity within a spatial domain, topographic features, and others have

been evaluated with machine learning methods to gap-fill cloud-covered pixels with success (Hou et al., 2019).

We did not explore the use of any gap-filling methods in this study, as they were beyond the scope of this work. Our results showed that the time lag between the last day of observed 50% SCA and the center of mass of the hydrograph was 11 days shorter for the wettest water year than for the driest water year. This could be due to overland flow reaching the stream more quickly due to greater SWE being released from melting snowpack during the wet year, but we cannot make this assertion with confidence. We suspect that the time lag for the average water year is artificially high due to cloud cover, and that the last day of 50% SCA likely occurred later in the water year than what we found with our data.

Future users of STARFM should explore which gap-filling methods best fit their data and are most appropriate to their research questions. In addition, efforts to address issues with cloud cover via in situ data would greatly improve intra-annual comparison of snow cover patterns. For example, the Surface Atmosphere Integrated field Laboratory (SAIL) campaign scheduled to take place in the East River SFA from 2022-2023 will acquire detailed ground-based observational data of cloud cover, including cloud-over-snow conditions, in the study area (Feldman et al., 2019). These data have the potential to be used to inform models about how to better identify and predict snow cover on the ground when satellite-based observations are obstructed by cloud cover.

3.4.4 Future Work

As climate continues to change, methods that rely heavily on historical snow accumulation and ablation regimes to predict streamflow will become increasingly less

effective for water resource management (Milly et al., 2008). Greater variability in temperature, precipitation, and extreme weather events (IPCC, 2014) will make point measurements at locations that historically served as index sites, especially those in sparsely observed areas, less representative of the regions they are designed to typify. Meanwhile, the demand in the hydrology and CZ communities for high spatiotemporal resolution remote sensing data will continue to grow.

The 20-year, daily NDSI dataset presented here provides future researches with a starting point for answering research questions relating to snow cover in the East River watershed. We have provided some examples of how our data can be utilized to monitor seasonal changes in SCA, but many other possibilities exist. For example, a temporal analysis of snow cover on a daily, weekly, or monthly scale could provide further insights regarding what landscape features have the greatest influence on snow cover during different times of the year. Such analysis could also provide further insight into the greater variability in snow cover observed during WY 2012 which may be common for other dry years.

Conversely, the multi-decadal time span of our dataset provides opportunities for answering questions related to if and how snow cover patterns in the East River watershed are already shifting under the influence of climate change. Our dataset could be used to calibrate/validate climatological models for predicting future snow cover patterns in the East River watershed. For example, Carroll et al. (2018) found that the upper subalpine region of the East River SFA dominates groundwater recharge due to its propensity to retain snowpack later in the year, which is consistent with our findings that these high elevation areas consistently have the greatest percent annual snow cover. Our

results also showed that high elevation snow cover did not show the same decreases during the dry water year as did mid- and lower elevation snow cover. While we did not take temperature differences between our water years of analysis into account, future work may be able to determine which climatological factors have the greatest impact on the retention of this critical upper elevation snow cover, as well as if and how this may change in the coming decades.

Additionally, the ability to monitor snow cover patterns in the landscape on fine spatiotemporal scales with our dataset provides an opportunity to gain a more complete picture of snowmelt driven hydrological processes in the watershed. Previous work in the East River SFA has shown that seasonal subsurface flow regimes directly control concentration-discharge relationships of carbon and other key nutrients exiting the watershed. The depth to which water from snowmelt percolates is directly influenced by the amount of SWE and the timing at which it is released from the snowpack. These processes can control which and how many subsurface weathering fronts the water comes into contact with, and thus the resulting chemical composition of streamflow (Winnick et al., 2017). High spatiotemporal resolution snow cover data has the potential to assist in modeling these subsurface reactions, predicting concentration-discharge relationships before they show up in streamflow, and understanding what impacts this year-to-year variability has on nutrient mobility within the watershed.

3.5 Conclusion

The results of our work provide information that can be used for future research into snowpack and streamflow forecasting, modeling the movement of water through the

CZ, and investigations into both present and future effects of climate change on the East River watershed and other similar snow-dominated mountain catchments.

Specifically, we describe an example of how our 20-year, daily NDSI dataset can be used to investigate differences in SCA between three very different water years. Using daily NDSI data derived through STARFM, we were able to calculate differences in percent annual snow cover across the East River watershed for a wet, average, and dry water year. We further examine percent annual snow cover by the landscape characteristics of elevation, slope, aspect, and land cover type. This analysis suggests that elevation is the primary driver of differences in SCA across all three water years.

Additionally, we provided an example of how our daily SCA data can be used to better understand water delivery to the CZ by examining how percent annual snow cover relates to yearly discharge at the watershed outlet for all three water years. Specifically, by calculating the time lag between the last day of the water year when 50% of the watershed was snow covered and the date of the center of mass of the hydrograph. Although we did find a small difference in the time lag between the wettest and the driest year, this analysis highlighted the limitations that cloud cover introduces into any analysis of optical remote sensing data and the need to address them before overarching conclusions can be drawn from these results.

References

Anderson, B. T., McNamara, J. P., Marshall, H., & Flores, A. N. (2014). Insights into the physical processes controlling correlations between snow distribution and terrain properties. *Water Resources Research, 50*, 4545-4563. doi:10.1002/2013WR013714

Anderton, S. P., White, S. M., & Alvera, B. (2004). Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes, 18*, 435-453. doi:10.1002/hyp.1319

Barnhart, T. B., Molotch, N. P., Livneh, B., Harpold, A. A., Knowles, J. F., & Schneider, D. (2016). Snowmelt rate dictates streamflow. *Geophysical Research Letters, 43*, 8006–8016. doi:10.1002/2016GL069690

Boise State's Research Computing Department. 2017. R2: Dell HPC Intel E5v4 (High Performance Computing Cluster). Boise, ID: Boise State University. doi:10.18122/B2S41H.

Buhler, Y., Adams, M. S., Bosch, R., & Stoffel, A. (2016). Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations. *The Cryosphere, 10*, 1075-1088. doi:10.5194/tc-10-1075-2016

Carroll, R. W. H., Bearup, L. A., Brown, W., Dong, W., Bill, M., & Williams, K. H. (2018). "Factors controlling seasonal groundwater and solute flux from snow-dominated basins". *Hydrological Processes, 32*, 2187-2202. doi:10.1002/hyp.13151

Chastain, R., Housman, I., Goldstein, J., Finco, M., & Tenneson, K. (2019). Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States. *Remote Sensing of Environment, 221*, 274-285. doi:10.1016/j.rse.2018.11.012

Che, T., Li, X., Jin, R., Armstrong, R. & Zhang, F. (2008). Snow depth derived from passive microwave remote-sensing data in China. *Annals of Glaciology 29*, 145-154. doi:10.3189/172756408787814690

Chen, J., Zhu, X., Vogelmann, J. E., Gao, F., & Jin, S. (2011). A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment, 115,* 1053-1064. doi:10.1016/j.rse.2010.12.010

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Hreinsson, E. O., & Woods, R. A. (2011). Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research, 47,* 1-23. doi:10.1029/2011WR010745

Daly, S. F., Vuyovich, C. M., Deeb, E. J., Newman, S. D., Baldwin, T. B., & Gagnon, J. J. (2012). Assessment of the snow conditions in the major watersheds of Afghanistan using multispectral and passive microwave remote sensing. *Hydrological Presses, 26,* 2631-2642. doi:10.1002/hyp.9367

Dozier, J., Painter, T. H., Rittger, K., & Frew, J. E. (2008). Time-space continuity of daily maps of fractional snow cover and albedo from MODIS. *Advances in Water Resources, 31,* 1515–1526. doi:10.1016/j.advwatres.2008.08.011

Elder, K., Rosenthal, W., & Davis, R. E. (1998). Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes, 12,* 1793-1808.

Farinotti, D., Magnusson, J., Huss, M., & Bauder, A. (2010). Snow accumulation distribution inferred from time-lapse photography and simple modelling. *Hydrological Processes, 24,* 2087-2097. doi:10.1002/hyp.7629

Feldman, D., Collins, W., Aiken, A. C., Boos, W., Carroll, R., Chandrasekar, V., Collins, W., Collis, S., DeMott, P., Fan, J., Flores, A., Gochis, D., Kumjian, M., Leung, L. R., O'Brien, T., Raleigh, M., Rhoades, A., Skiles, M., Smith, J… Williams, K. (2019). Surface Atmosphere Integrated field Laboratory (SAIL) in Southwestern Colorado, USA. 135 pp.

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 44*(8), 2207–2218. doi:10.1109/TGRS.2006.872081

Hiemstra, C. A., Liston, G. E., & Reiners, W. A. (2006). Observing, modelling, and validating snow redistribution by wind in a Wyoming upper treeline landscape. *Ecological Modelling, 197*, 35–51. doi:10.1016/j.ecolmodel.2006.03.005

Hou, J., Huang, C., Zhang, Y., Guo, J., & Gu, J. (2019). Gap-Filling of MODIS Fractional Snow Cover Products via Non-Local Spatio-Temporal Filtering Based on Machine Learning Techniques. *Remote Sensing*, *11*(90), 1-24. doi:10.3390/rs11010090

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kerr, T., Clark, M., Hendrikx, J., & Anderson, B. (2013). Snow distribution in a steep mid-latitude alpine catchment. *Advances in Water Resources, 55*, 17-24. dx.doi:10.1016/j.advwatres.2012.12.010

Knowles, J. F., Harpold, A. A., Cowie, R., Zeliff, M., Barnard, H. R., Burns, S. P., Blanken, P. D., Morse, J. F., & Williams, M. W. (2015). The relative contributions of alpine and subalpine ecosystems to the water balance of a mountainous, headwater catchment. *Hydrological Processes, 29*, 4794–4808. doi:10.1002/hyp.10526

Landfire. (2016), Existing vegetation type layer, LANDFIRE 2.0.0, U.S. Department of the Interior, Geological Survey. Accessed 30 September 2020

Lundquist, J. D., Dickerson-Lange, S. E., Lutz, J. A., & Cristea, N. C. (2013). Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling. *Water Resources Research, 49,* 6356-6370. doi:10.1002/wrcr.20504

Markham, B. L., Storey, J. C., Williams, D. L., & Irons, J. R. (2004). Landsat Sensor Performance: History and Current Status. *IEEE Transactions on Geoscience and Remote Sensing, 42*(12), 2691-2694. doi:10.1109/TGRS.2004.840720

McClung, D., & Schaerer, P. (2006). The Avalanche Handbook. 3rd ed. Seattle: The Mountaineers.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Zbigniew, W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? *Science*, *319*, 573–574.

Rittger, K., Painter, T. H., & Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, *51*, 367–380. doi:10.1016/j.advwatres.2012.03.002

Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment, 112*, 3112-3130. doi:10.1016/j.rse.2008.03.009

Selkowitz, D. J., Painter, T. H., Rittger, K., Schmidt, G., & Forster, R. (2017). The USGS Landsat Snow Covered Area Products: Methods and Preliminary Validation CHAPTER 5. In Automated Approaches for Snow and Ice Cover Monitoring Using Optical Remote Sensing (pp. 76). The University of Utah.

Sproles, E. A., Crumley, R. L., Nolin, A. W., Mar, E., & Moreno, J. I. L. (2018). SnowCloudHydro - A new framework for forecasting streamflow in snowy, data-scarce regions. *Remote Sensing, 10*(1276), 1–15. doi:10.3390/rs10081276

Tennant, C. J., Crosby, B. T., & Godsey, S. E. (2015). Elevation-dependent responses of streamflow to climate warming. *Hydrological Processes, 29,* 991-1001. doi:10.1002/hyp.10203

Tokunaga, T. K., Wan, J., Williams, K. H., Brown, W., Henderson, A., Kim, Y., Tran, A. P., Conrad, M. E., Bill, M., Carroll, R. W. H., Dong, W., Xu, Z., Lavy, A., Gilbert, B., Romero, S., Chirstensen, J. N., Faybishenko, B., Arora, B., Siirila-Woodburn, E. R., Versteeg, R.,… Hubbard, S. S. (2019). Depth- and Time-Resolved Distributions of Snowmelt-Driven Hillslope Subsurface Flow and Transport and Their Contributions to Surface Waters. *Water Resources Research, 55*, 9474–9499. doi:10.1029/2019/WR025093

USGS. (2020). Landsat 8 Collection 1 (C1) Land Surface Reflectance Code (LaSRC) Product Guide Version 3.0 LSDS-1368

Veatch, W., Brooks, P. D., Gustafson, J. R., & Molotch, N. P. (2009). Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site. *Ecohydrology, 2,* 115-128. doi:10.1002/eco.45

Winnick, M. J., Carroll, R. W. H., Williams, K. H., Maxwell, R. M., Dong, W., & Maher, K. (2017). Snowmelt controls on concentration-discharge relationships and the balance of oxidative and acid-base weathering fluxes in an alpine catchment, East River, Colorado. *Water Resources Research, 53,* 2507-2523. doi:10.1002/2016WR019724

APPENDIX A

Additional Figures

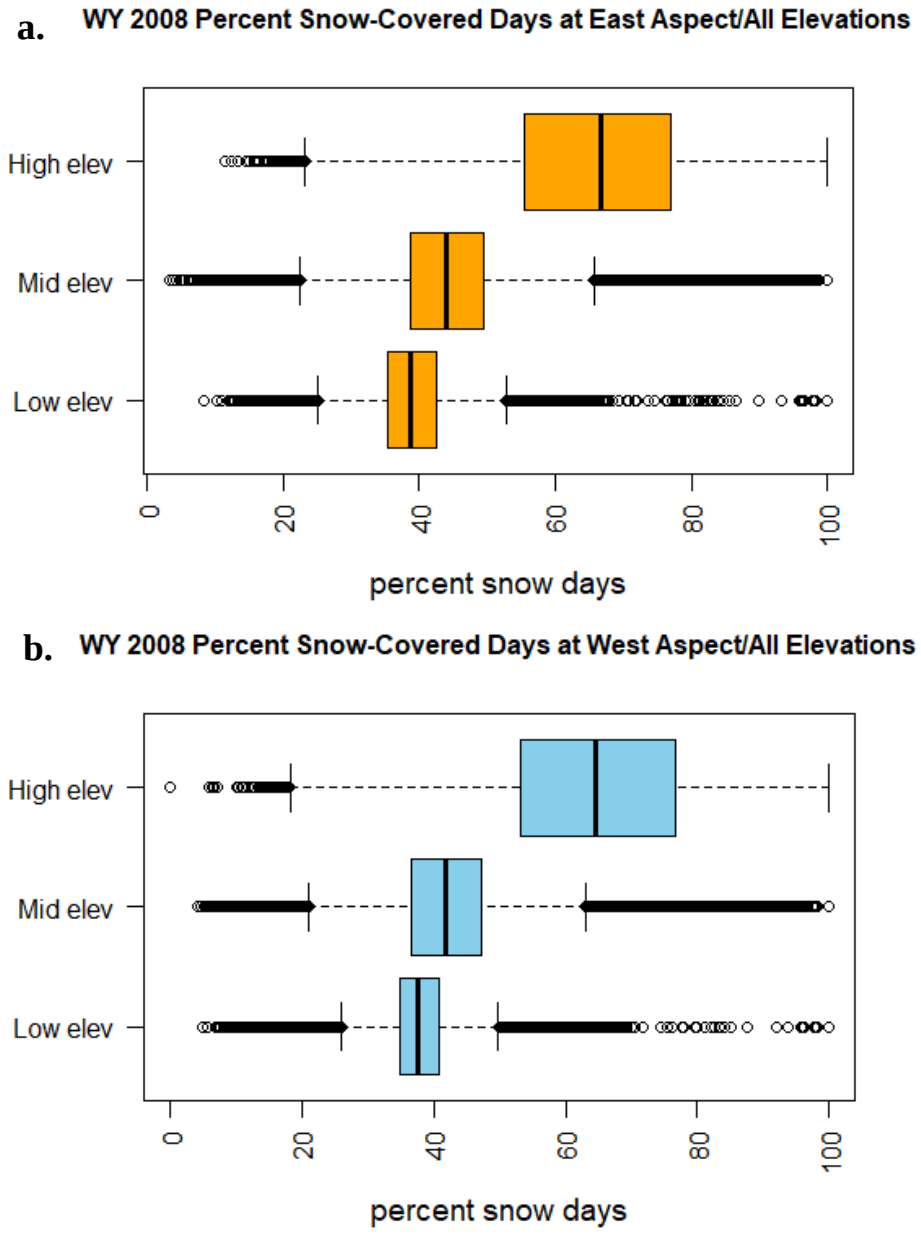


Figure A.1: Numerical distributions of WY 2008 percent annual snow-covered days by elevation band for a) east aspects and b) west aspects.

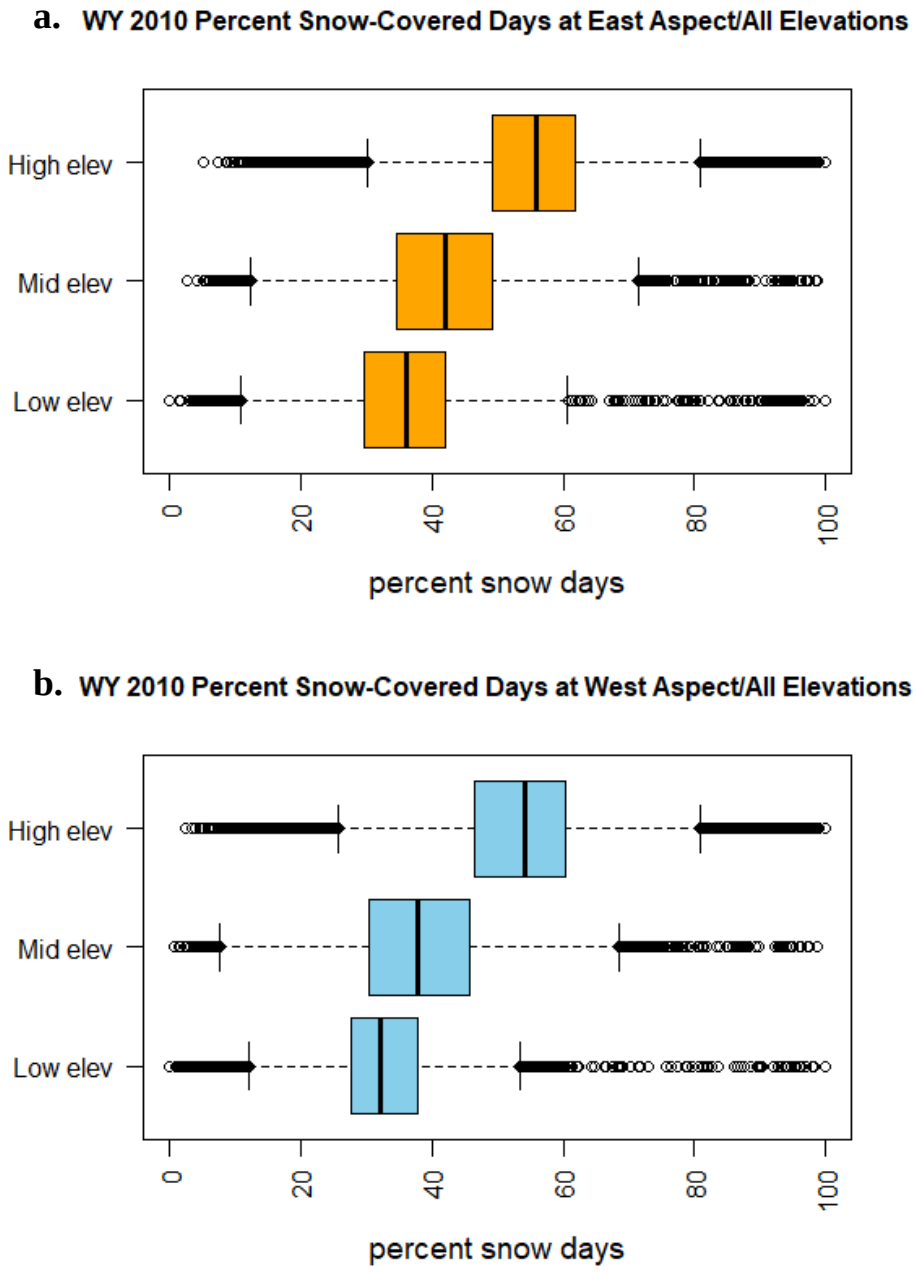


Figure A.2: Numerical distributions of WY 2010 percent annual snow-covered days by elevation band for a) east aspects and b) west aspects.

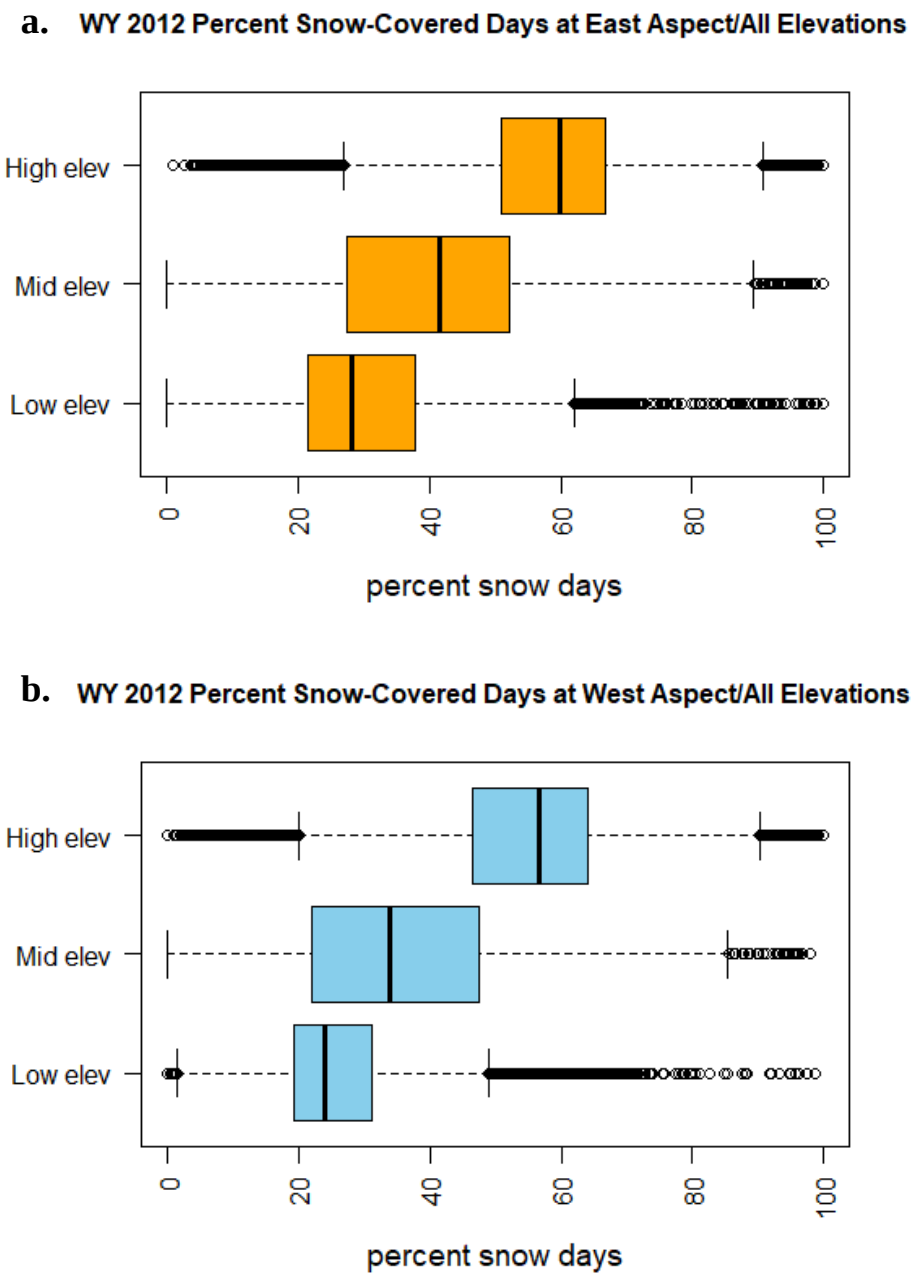


Figure A.3: Numerical distributions of WY 2012 percent annual snow-covered days by elevation band for a) east aspects and b) west aspects.

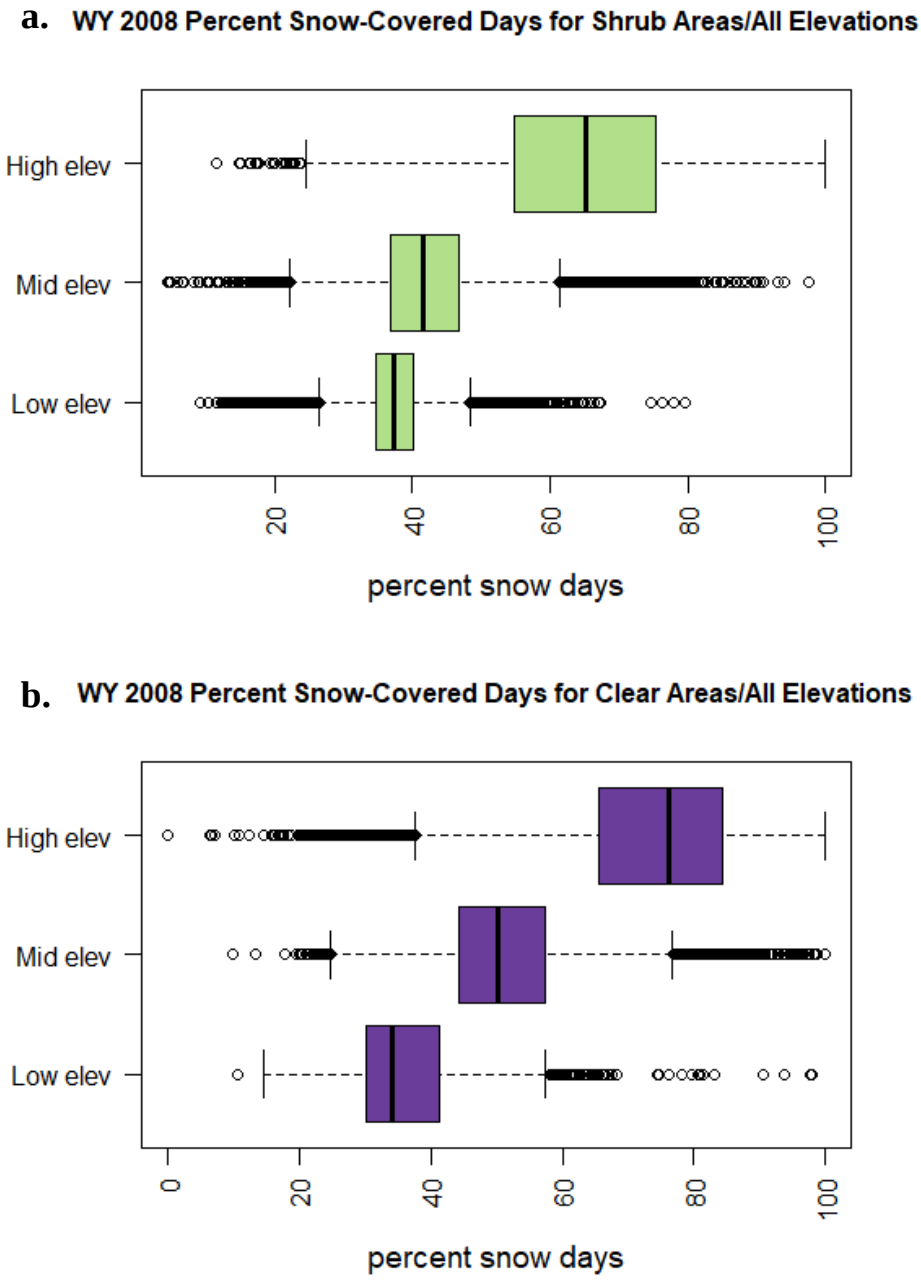


Figure A.4: Numerical distributions of WY 2008 percent annual snow-covered days by elevation band for a) shrub areas and b) clear/unvegetated areas.

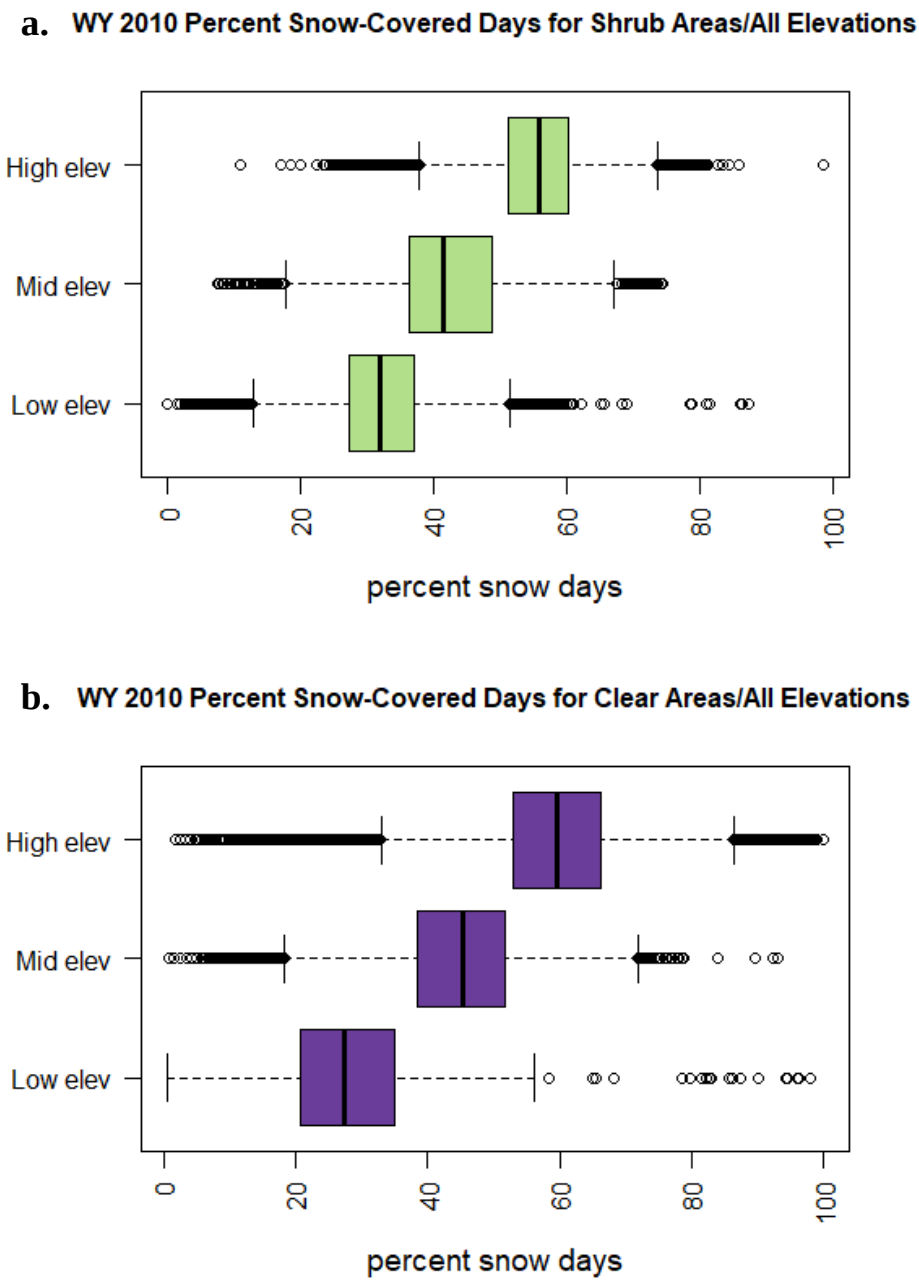


Figure A.5: Numerical distributions of WY 2010 percent annual snow-covered days by elevation band for a) shrub areas and b) clear/unvegetated areas.

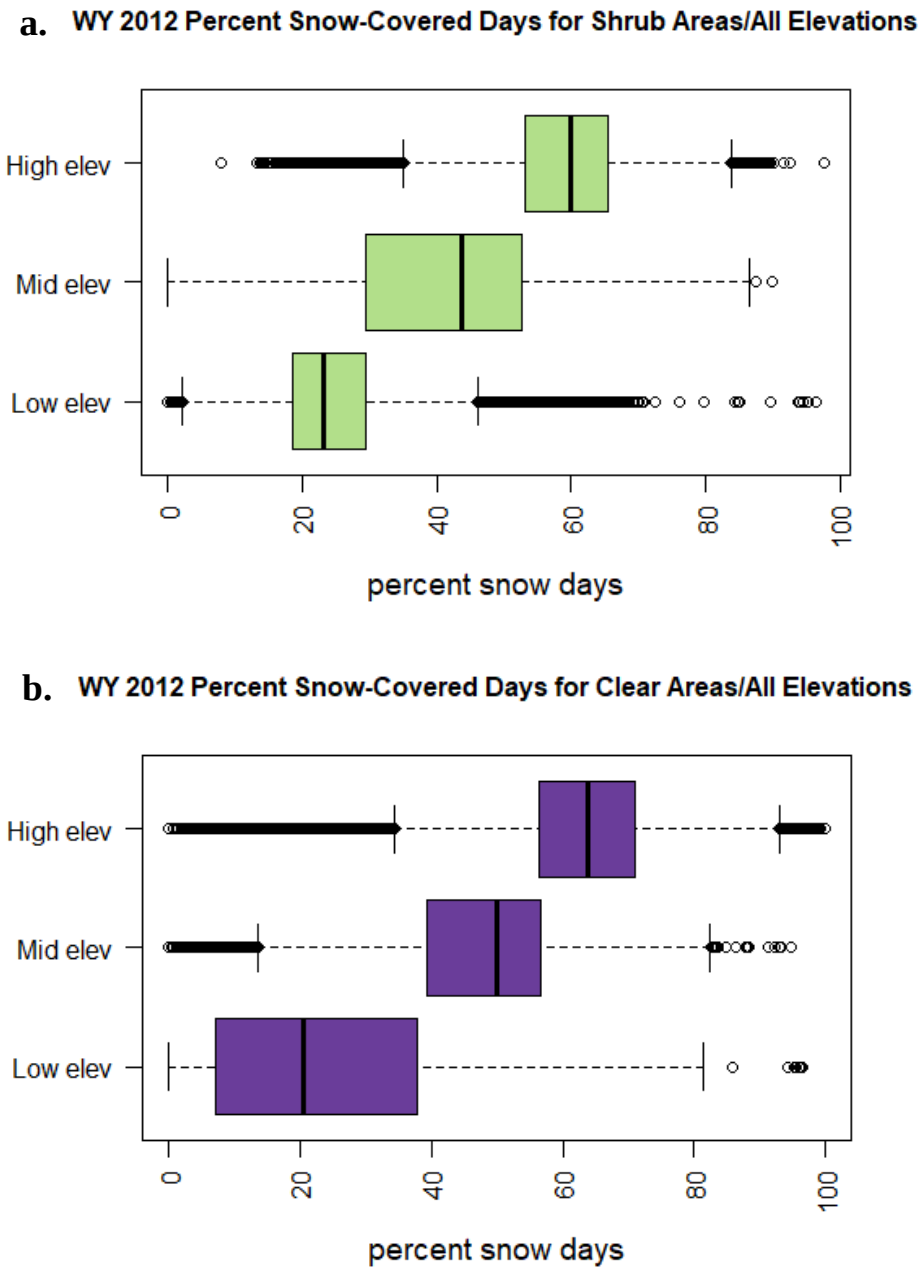


Figure A.6: Numerical distributions of WY 2012 percent annual snow-covered days by elevation band for a) shrub areas and b) clear/unvegetated areas.

APPENDIX B

Github Link

https://github.com/AllisonVincent/StarFM-code

APPENDIX C

MODIS, Landsat, and STARFM File Sizes

Table C.1: The approximate summed size of all Landsat, MODIS, and STARFM files for a single water year of data for the 20-year East River watershed NDSI dataset.

Landsat	MODIS	STARFM
219.5 MB	1180.3 MB	960.1 MB