Aji John

# Understanding climate change effects in alpine meadows: harnessing the power of data science, and remote sensing

Aji John

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Janneke Hille Ris Lambers, Chair

Ka Yee Yeung, Co-Chair

Tom Daniel

Program Authorized to Offer Degree:
Department of Biology

University of Washington

Abstract

# Understanding climate change effects in alpine meadows: harnessing the power of data science, and remote sensing

Aji John

Chairs of the Supervisory Committee:

Professor Janneke Hille Ris Lambers
Department of Biology

Professor Ka Yee Yeung
School of Engineering & Technology - UWT

Montane ecosystems are particularly sensitive to climate warming. This is because Alpine wildflowers that are dominant in these ecosystems are prone to local level extinction (as suitable habitats disappear), as spring and summer temperatures increase. This can either occur when trees (previously limited by climate) invade Alpine meadows or Alpine plants encounter limited habitats at upper range limits into which to expand following warming. Climate-change induced shifts in Alpine wildflower phenology, observed for numerous species, can also have negative impacts by causing flowers to be exposed to damaging climatic conditions (e.g., early frosts) or causing pollinators to become mismatched with flowering. To anticipate these changes in natural areas, we need tools with which to monitor climatic factors influencing Alpine wildflower habitats as well as climatic impacts on Alpine Wildflowers.

Unfortunately, the effects of warming, both on the climatic factors influencing Alpine wildflower habitats and phenology are difficult to quantify at the spatial and temporal scales needed. For example, many studies in Alpine habitats rely on climate stations to detect shifts in warming or snow melt, but these stations are generally few in number, and therefore fail to capture the spatial variability necessary to understand snowmelt patterns. Discernible phases of Alpine plant phenology have been collected by routine field measurements during the growing season, some using in-situ observations and others aided by citizen science efforts (e.g., National Phenology Network). Although such investigations have been instrumental in quantifying phenological shifts, they are challenged by the fact that limited resources often make it difficult to gather observations in remote locations, over large spatial scales and at sufficient frequency to capture phenological transitions.

Recent technological developments in imaging, from the exploding and ubiquitous use of smartphones, increasingly improved UAV devices (drones), and satellites offer great promise in meeting this challenge. Of course, satellite-based imagery (e.g., MODIS, Landsat 8 and Sentinel) has been key for detecting changes in plant phenology using spectral signatures in other systems, but its main limitation in Alpine systems has been its coarse spatial resolution (30 - 1000m) and low temporal frequency (biweekly to once a month). This is insufficient to monitor Alpine systems, which occupy relatively narrow geographic areas (between treeline and rock / ice), and where seasonal dynamics are constrained to very short summer seasons (e.g., 3 months). The advent of CubeSat imagery that provides daily 3 - 5m resolution imagery (e.g. Planet, Digital Globe, etc.) has

opened up the possibility of detecting snow dynamics and multiple phenological changes (e.g., green-up, flowering) at much higher spatiotemporal accuracy and sustainability. Simultaneous advances in quantitative approaches to analyzing images (e.g., machine learning), paired with an increasingly high volume of images available from the general public, similarly provides new opportunities to detect Alpine flowering at increasingly small spatial scales and high temporal frequency.

In this dissertation, I explore the promise and challenge of using images to better understand Alpine systems, particularly in terms of their phenology. In particular, I wished to develop tools that both capture the environmental drivers of that phenology (which is snow in our field sites and many Alpine systems) as well allow us to monitor the progression of phenology at the landscape and individual species level. Specifically, I first (in Chapter one) sought to improve high-resolution (meter-scale) mapping of snow-covered areas in complex and forested terrain. Such mapping is critical for understanding the responses of montane flowering species to climate change. Here, myself and my colleagues used high-resolution imagery from PlanetScope to derive snow-covered areas and build on a previous work that demonstrated snow cover area mapping using only the PlanetScope 4-band (Red, Green, Blue and NIR) reflectance with convolutional neural network (CNN) based machine learning (ML) approach. We augmented the existing CNN model with additional input data including vegetation metrics (Normalized Difference Vegetation Index) and DEM-derived metrics (elevation, slope and aspect) to improve SCA mapping in forested and open terrain (like alpine meadows) and showed that an augmented model that used the Normalized Difference Vegetation Index (NDVI) along with visible (red, green, and blue) and NIR bands was the best performing model. The

NDVI based model with an F-score of 0.89 (Gunnison) and 0.93 (Engadin) was found to be 4% and 2% better than when using canopy height and terrain-derived measures at Gunnison, respectively. We examined the model's performance in forested areas using three forest canopy quantification metrics and found that augmented models better identified snow in canopy edges and open areas but still underpredicted snow cover under forest canopies. The improved high-resolution snow maps in forested environments can support studies involving climate change effects on a broad range of phenomena in mountain ecosystems, including Alpine wildflower phenology as well as evaluations of hydrological impacts in snow-dominated river basins.

In chapter two, my collaborators and I evaluated the use of hyperspectral CubeSat imagery to develop models for detecting peak flowering phenology - hypothesizing that reflectance in visible and NIR bands from CubeSat imagery could be used to distinguish community flowering at the pixel level. I examined Alpine wildflower meadows at Mt. Rainier National Park (MORA) using Random Forest (RF) classification applied to high-resolution (3-m PlanetScope from Planet Labs, Inc., San Francisco, CA, USA) and moderate resolution (10-m Sentinel and 30-m Landsat) imagery. On-the ground citizen science data collected by the MeadoWatch program was used to validate these models. Peak flowering delineation using 3-m resolution PlanetScope imagery resulted in an accuracy of 70% (Cohen's kappa = 0.25). Combining Planet data with other sources of imagery (10-m Sentinel and 30-m Landsat) improved the delineation of peak flowering (accuracy = 77%, Cohen's kappa = 0.39). This approach was also able to identify the timing of peak flowering in a particularly warm year (2015), despite being calibrated on normal climate years. In all, this work suggests that PlanetScope imagery holds promise

for detecting community-level Alpine wildflower phenology, and is therefore particularly useful in global change ecology, where temporal frequency is especially important. Additionally, combining imagery may provide a new approach for cross-calibrating sensors to account for radiometric irregularity inherent in fine resolution PlanetScope imagery.

In Chapter three, I explore the use of machine learning in crowd sourced camera images to quantify flowering species richness in the same meadows. I used three deep learning techniques (Mask R-CNN, RetinaNet and YOLO) to detect wildflower species in images taken during two flowering seasons and found that deep learning techniques can detect multiple species, providing information on flowering richness in photographed meadows. Two-stage detector Mask R-CNN was more accurate than single-stage detectors like RetinaNet and YOLO, with the Mask R-CNN network performing best overall with mean average precision (mAP) of 0.67 followed by RetinaNet (0.5) and YOLO (0.4). The results indicate higher richness just above the tree line for most of the species, which is comparable with patterns found using field studies. These approaches can be used to address many other ecological questions that benefit from automated flower detection, like studies of flowering phenology or floral resources for pollinators. Thus, this approach can complement a wide range of more traditional existing ecological approaches, e.g., observational monitoring using scientists or citizen scientists or experiments. The development of this approach for wildflower phenology predictions provides new possibilities to monitor climate change effects on flowering communities at high temporal scales in locations where regular access is challenging.

In summary, I took advantage of rapid technological and quantitative developments to build tools that allow ecologists to better monitor climate change impacts in sensitive Alpine wildflower meadows in this dissertation. I believe these tools can widely be applied in any ecologically sensitive system where processes of interest are rapid and can be captured in images, but where regular access by scientists to monitor these processes is expensive or difficult (e.g., remote islands, high canopies in trees, etc.).

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

Many thanks to all who helped me in this journey. An unexpected detour that led me to University of Washington and gave me the chance to study under Janneke; for that I would always be grateful to University of Washington and Janneke. The steady support of friends made this journey memorable and helped with every small step I took. Tony Cannistra, Meera Lee Sethi, Stuart Graham helped in every way possible; from reading manuscript revisions to helping with fieldwork, many thanks. Kevin Xu and Angela Xu who were always there alongside, many thanks. Kavya Pradhan who has been the positive force, and a wonderful collaborator, I would like to give deepest gratitude for all the help. Janneke's subtle but persistent feedback on how to conduct research, improve and write concisely made things better. I want to acknowledge Nicoleta Cristea's help in supporting me academically and help conducting cutting-edge science. Many more to thank, but I'm obviously forgetting – academic peers, undergrads, faculty, and friends who made this happen.

The biology department gave me the much-needed support when things did not go that well; many thanks to all who pulled me through those tough times.

Finally, my family took the brunt of it all; my wife Xuemin who undauntedly supported this journey. Our two kids Thea and Marcus who participated with vigor and enthusiasm. I would like to thank my parents and my sister who cheered me on with their unrelenting love and encouragement.

# DEDICATION

To my family – Xuemin, Thea and Marcus

# INTRODUCTION

Alpine environments harbor a wide diversity of plant and animal species but are also extremely susceptible to climate change [1,2]. This is because alpine meadows are limited by a short growing window that is driven by late snowmelt [3], which can make species adapted to this regime sensitive to a warming climate [4]. Additionally, alpine meadows occur in narrow distributional bands squeezed between forests and rock and ice, making them sensitive to warming induced distributional shifts. Unfortunately, monitoring studies in Alpine regions are challenged by limited resources and poor access, resulting in scarcity of datasets that fully capture the spatiotemporal dynamics of these Alpine systems in response to climate change. In this dissertation, I explore a combination of remote sensing and machine learning approaches to fill this data gap, providing the ecological community with tools that allow for a better understanding of how alpine meadows, particularly alpine wildflowers, are responding to climate change.

I focus specifically on the drivers and process of Alpine wildflower phenology (the timing of reproductive stages), because it is highly sensitive to spring and summer temperatures and the timing of snowmelt [5]. Shifts in the seasonal timing of flowering, and fruiting (as well as other phenological events - e.g., germination, migration) have been observed across numerous species in response to changing climates [6,7]. Shifts in the flowering phenology of Alpine wildflowers is concerning, as this could disrupt interactions between Alpine wildflowers and other members of the community (e.g., pollinators [8] as well as other flora and fauna [9,10]). Thus, understanding the spatiotemporal dynamics of snow (a key driver of Alpine wildflower phenology) and spatiotemporal patterns of floral richness can provide insight into the effects of climate

change on Alpine systems [11]. Field studies compiled by scientists and volunteer networks can and have documented shifts in the timing of flowering, but they often lack scalability to the large spatial scale and high temporal frequency needed to fully understand these processes.

Fortunately, recently developed small satellites, so-called CubeSats, have provided new possibilities to monitor landscapes at fine spatial and temporal scales. In comparison to more traditional satellite imagery (e.g., MODIS, Landsat), CubeSats have the advantage of high spatial and temporal resolution. For example, Planet Labs, Inc. [12] uses over 150+ CubeSats (as of 2019) to image the entire land surface of the Earth at a daily time interval, providing 3–5 m resolution images. These images provide information on the reflectance spectra captured in the corresponding electromagnetic spectrum, which are often used to detect plant physiological properties (e.g., NDVI) from various band ratios [13]. Thus, CubeSats provides ecologists with an exciting new opportunity to track a wide array of natural processes that occur at fine spatial scales and high temporal frequency – provided we can extract sufficient information from multispectral images to detect (for example) patterns in snow cover and flowering.

In chapter one, I therefore consider how to combine CubeSat images with information from Lidar (elevation, canopy cover) to predict snow cover. Climate-change induced changes to snow cover are likely to influence alpine meadow and mountain forest phenology, because snow cover in mountains often dictates the start and end of the growing season. Snow cover mapping in mountainous areas remains a complex challenge, since snow accumulation is driven by climate, topography, and vegetation type. The accumulation patterns of snow in forests also vary as a function of distance from the canopy—with canopy, gaps and canopy edges all potentially having different snow

accumulation patterns [14,15]. High-resolution (m-scale) lidar-derived snow depth datasets can provide insight into the main factors controlling snow distribution and melt patterns [16], but these data are expensive to obtain and not available at the frequency necessary to predict year-to-year differences in snow covered area. Combining these data with CubeSat data therefore allows us to combine the fine-scale topography and vegetation information (both drivers of snowmelt) with the high temporal frequency of CubeSat data, to better understand year-to-year variation in snow cover.

Combining data collected at multiple spatial and temporal resolution can be complicated, but fortunately, I leveraged recent advances in statistical approaches to better model snow and wildflower phenology. In particular, machine learning (ML)-based methods have proven to be extremely promising at detecting snow cover [15]; with recent work by Cannistra et al., 2021 [17] demonstrating that snow cover can be successfully mapped from PlanetScope data using CNN to 4-band (R, G, B and NIR) PlanetScope surface reflectance. I augment the Cannistra et al., 2021, CNN snow cover model by using additional predictors including vegetation structure (using lidar-derived canopy height from a canopy height model (CHM) and the PlanetScope-derived Normalized Difference Vegetation Index (NDVI)), and the digital elevation model (DEM or elevation) and its derived attributes (i.e., slope, aspect and northness).

In chapter two, I use a similar approach (combining PlanetScope data with machine learning approaches) to delineate the flowering season of Alpine meadows. Alpine wildflowers are considered good indicators of climatic change, since their phenology is highly sensitive to the timing of snowmelt and climate change projections show a decrease in snow fall that is likely to influence meadow phenology [18]. The ability to detect flowering phenology via remote sensing, therefore has the potential to aid in the

detection of shifts in the peak wildflower season over large areas in remote Alpine wildflower meadows. Therefore, I evaluated whether the spectral bands of PlanetScope (a 3-m, 4-band multispectral image product from Planet Labs) can detect a process of interest – in this case wildflower phenology. We also asked whether combining satellite imagery (PlanetScope – 3-m resolution with Sentinel – 10-m resolution and Landsat 30-m resolution) can improve the detection process. In both cases, we verified the accuracy of these models using existing long-term on-the-ground datasets. These long-term datasets identify peak flowering in wildflower meadows during a period of five years, at multiple sites along an elevation gradient at Mt. Rainier National Park in Washington, USA.

Understanding the timing of the peak wildflower season is important, but many ecologists would also like to decompose community level phenology down to species level phenology and gain information on the phenological flowering patterns of individual species. Thus, in chapter three I focus on combining images captured on the ground with Machine Learning approaches to detect and identify individual flowers, to assess whether this general approach is feasible for capturing phenological patterns in specific wildflower species. We asked whether we could estimate floral richness, which is an important yardstick of the fitness of Alpine wildflowers [19] and could help ecologists quantify the ecological effects of climate change [20–23]. Here, as in chapter one, recent developments in computer vision algorithms provided new opportunities. Specifically, studies have evaluated Machine learning (ML) vision-based algorithms to detect a single phenological stage, which has been successful (e.g. fruit maturity in passion fruit [24], strawberries [25] and tomatoes [26]; ripeness in papayas [27], and flowering in cotton plants [28]). An evolving class of ML algorithms called deep neural networks (dNNs) have been

particularly useful for detecting and segmenting multiple objects in an image. dNN methods can perform object detection (for e.g., a flower) and instance segmentation (for e.g., detect all the instances of a flower). This means flowers can be individually tagged and segmented in a meadow image with dedicated dNNs. These previous studies showed that it is possible to evaluate stages of a plant life event (for e.g., fruiting) from images with ML methods. However, detecting flowers from images in natural settings (such as alpine meadows) was a more complex task because of diversity of meadow species, wide range of floral structures and backgrounds that exist (like leaves, trees, rocks etc.). I addressed this challenge in chapter three by applying dNN methods to camera images of wildflower meadows taken at various elevations at Mt. Rainier National Park during two flowering seasons.

In summary, this dissertation provides ecologists with new approaches for evaluating patterns of snow cover, wildflower phenology and floral richness, which are important for monitoring climate change effects on flowering communities at broader spatiotemporal scales. In chapter 1, I used a machine learning approach (CNN models) to incorporate using additional predictors including vegetation structure (using lidar-derived canopy height from a canopy height model (CHM) and the PlanetScope derived Normalized Difference Vegetation Index (NDVI)), and the digital elevation model (DEM or elevation) and its derived attributes (i.e., slope, aspect and northness) to map snow cover across a variety of landscape types. In chapter 2, I asked whether multi-spectral imagery obtained from Planet Labs, Inc., could be used to quantify peak flowering in sub-alpine meadows. I did so by combining satellite imagery with existing long-term on-the-ground datasets. In chapter 3, I explored the use of dNN methods (YOLO, RetinaNet and

Mask R-CNN) for the detection and localization of Alpine wildflowers in complex natural settings – i.e., montane meadows.

# Chapter 1. HIGH-RESOULTION SNOW-COVERED AREA MAPPING IN FORESTED MOUNTAIN ECOSYSTEMS USING PLANETSCOPE IMAGERY

Aji John [1,*], Anthony F. Cannistra [1], Kehan Yang [2,3],Amanda Tan [3], David Shean [2], Janneke Hille Ris Lambers [1,4] and Nicoleta Cristea [2,3]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA; [2]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA; [3]eScience Institute, University of Washington, Seattle, WA 98195, USA; [4]Plant Ecology, Institute of Integrative Biology, D-USYS, ETH Zürich, 8092 Zürich, Switzerland

**Citation:** John A, Cannistra AF, Yang K, Tan A, Shean D, Hille Ris Lambers J, Cristea N. High-Resolution Snow-Covered Area Mapping in Forested Mountain Ecosystems Using PlanetScope Imagery. Remote Sensing. 2022; 14(14):3409. https://doi.org/10.3390/rs14143409

## 1.1 ABSTRACT

Improving high-resolution (meter-scale) mapping of snow-covered areas in complex and forested terrains is critical to understanding the responses of species and water systems to climate change. Commercial high-resolution imagery from Planet Labs, Inc. (Planet, San Francisco, CA, USA) can be used in environmental science, as it has both high spatial (0.7–3.0 m) and temporal (1–2 day) resolution. Deriving snow-covered areas from Planet imagery using traditional radiometric techniques has have limitations due to the lack of a shortwave infrared band that is needed to fully exploit the difference in reflectance to discriminate between snow and clouds. However, re-cent work demonstrated that snow cover area (SCA) can be successfully mapped using only the PlanetScope 4-band (Red, Green, Blue and NIR) reflectance products and a machine learning (ML) approach based

on convolutional neural networks (CNN). To evaluate how additional features improve the existing model performance, we: (1) build on previous work to augment a CNN model with additional input data including vegetation metrics (Normalized Difference Vegetation Index) and DEM-derived metrics (elevation, slope and aspect) to improve SCA mapping in forested and open terrain, (2) evaluate the model performance at two geographically diverse sites (Gunnison, Colorado, USA and Engadin, Switzerland), and (3) evaluate the model performance over different land-cover types. The best augmented model used the Normalized Difference Vegetation Index (NDVI) along with visible (red, green, and blue) and NIR bands, with an F-score of 0.89 (Gunnison) and 0.93 (Engadin) and was found to be 4% and 2% better than when using canopy height- and terrain-derived measures at Gunnison, respectively. The NDVI-based model im-proves not only upon the original band-only model's ability to detect snow in forests, but also across other various land-cover types (gaps and canopy edges). We examined the model's performance in forested areas using three forest canopy quantification metrics and found that augmented models can better identify snow in canopy edges and open areas but still underpredict snow cover under forest canopies. While the new features improve model performance over band-only options, the models still have challenges identifying the snow under trees in dense forests, with performance varying as a function of the geographic area. The improved high-resolution snow maps in forested environments can support studies involving climate change effects on mountain ecosystems and evaluations of hydrological impacts in snow-dominated river basins.

## 1.2 INTRODUCTION

Mountain ecosystems in the western US are very sensitive to changes in climate [29–31]. Snowpack is melting earlier [32], affecting the fragile alpine systems for, e.g., the peak wildflower season that closely follows the snow disappearance day each year is shifting [33,34]. Unprecedented changes to mountain forests from fires and insect outbreaks [35] are also the result of changing climates, and forest managers are already opening gaps in areas of dense forest to improve ecosystem health and resilience [36]. Ecologists and conservation biologists study these systems using in situ monitoring, which is limited to point observations, precluding the understanding of the ongoing changes in areal extent and magnitude [37].

Changes to snow cover in particular are likely to influence mountain forests and meadows, because snow cover in the mountains dictates the growing degree days (accumulation of warmth). Thus, snow cover can affect the phenological advancement in species, causing the emergence of pollinators (insects etc.). Earlier snowmelt could lead to earlier flowering, which might cause asynchrony with pollinators or make them susceptible to frost events [4,38]. In a warming climate, early spring can also lead to spring runoff causing flooding and drier summers [14,39]. Understanding the heterogeneity of snow cover has implications that could inform biologists about the changes in ecosystems and water managers about the snow cover variability determined by seasonal snowmelt, which is a major component in hydrological water estimates [40].

The analysis of high-resolution (m-scale) lidar-derived snow depth datasets provides insights into the main factors controlling snow distribution and melt patterns, with important drivers shown to be elevation, followed by slope, aspect, and vegetation

cover. With the lidar-derived snow depth datasets, one can observe snow in openings, in forest gaps, and under forest canopies, and how it varies in space and time at high spatial resolutions of 1–3 m [14,41]. Airborne lidar datasets are known to be reliable in their estimates of snow depths in forested areas through comparing airborne lidar snow depths with manual measurements of snow around the trees [14] and are generally found to agree within 5 cm at 0.5–5 m resolutions [42]. High-resolution snow-covered areas can be derived from the lidar-derived snow-depth data using a prescribed threshold [16]. In 2013, NASA started airborne snow depth collections in the Tuolumne watershed in California, USA, weekly during the ablation season under a program called the Airborne Snow Observatory (ASO [43]). The airborne lidar snow depth collections have since been expanded to other areas, but it is cost prohibitive to expand at a global level, and the collections remain limited.

Fortunately, high spatiotemporal resolution imagery is available via Planet Labs, Inc. (Planet) [12] which has the potential to transform the study of earth-processes through remote sensing. Planet operates a constellation of CubeSats—small satellites that are of 3U form factor (10 by 10 by 30 cm) that images the entire land surface area of the earth daily. The “PlanetScope” constellation is composed of roughly 130 satellites that operate in sun-synchronous orbit. The resulting product is orthorectified surface reflectance (SR) that is delivered at 12-bit resolution with 3–5 m resolution covering visible and near infrared bands [44].

Of course, detecting snow from remote-sensing imagery requires a model to translate bands to predictions of snow, and machine learning (ML)-based methods have proven to be extremely promising at detecting snow cover. The majority of studies have used multispectral providers (e.g., MODIS and Landsat-8) because of the band diversity.

Commonly used methods are support vector machines (SVM) and random forests (RF), but deep learning-based methods are being applied increasingly [15]. Along with the multispectral bands, derived indices have also been used (for e.g., Normalized Difference Vegetative Index (NDVI) and Normalized Difference Snow Index (NDSI)), and yield snow cover area (SCA) estimates that have coarser spatial resolutions (e.g., 500 m) but finer temporal resolutions (e.g., daily) [45]. Studies that applied a suite of ML methods to a multispectral dataset combined with ancillary information such as topography found that the relevance of non-spectral attributes is of limited importance in increasing model performance [46,47]. Spatial resolution and accuracy tend to be issues in mountainous areas when using multispectral data (for e.g., MODIS)––coarser data products can capture less of environmental variation in very heterogeneous environments. Newer approaches have begun fusing multispectral satellite data with unmanned aerial vehicle (UAV)-based acquisitions or only using UAV-produced datasets to improve deficiencies over localized areas [48,49]. Another hybrid approach has also been proposed, where a convolutional neural network (CNN) is used to extract features and then a RF-based method is then used to estimate the snow-covered area [50]. Other explorations using satellite spaceborne synthetic aperture radar (SAR) data instead of multispectral have also shown promise in mapping snow-covered areas; SAR-based methods are not as prone to the presence of clouds, but SAR data have limitations in dense forests [51]. Furthermore, recent work by Cannistra et al., 2021 [17], has demonstrated that snow cover can be successfully mapped from PlanetScope data using a machine-learning approach based on a CNN using the 4-band (R, G, B and NIR) PlanetScope surface reflectance data as input. The study highlighted the viability of a CNN-based model trained on lidar-derived snow cover data despite limited radiometric bandwidth and band

placement, but, similar to other satellite-derived data, the performance of the model is lower in forested areas.

Snow cover mapping in forested areas remains a complex challenge as it is driven by vegetation type, topography, and climate. The accumulation patterns of snow in forests vary as a function of distance from the canopy—under the canopy, to canopy edge, and in gaps [14,15]. Dense canopy contributes to shading caused by tall trees and the proximity of varied overstories moderates the interception of snow [52]. Furthermore, the directional variance around the trees has also been found to influence snow accumulation patterns [41]. The relative influence of those factors on snow depth varies as a function of location and the snow season [53]. For instance, Tennant et al., 2017 [54], showed that elevation explained most of the variability in snow depth (16–79%) in forested areas, but the aspect explained more variability (11–40%) in open areas. Cristea et al., 2017, also showed that terrain may matter more in the open than in vegetated areas. Snow–forest interactions also vary as a function of climate and radiation effects, with snow disappearing first in the open or under the forest as a function of local conditions [55]. These studies demonstrated that small-scale variability is being observed with lidar and identified the terrain and vegetation features controlling the spatial distribution of snow depth and, hence, snow cover. Based on these observations, we hypothesized that augmenting CNN-based models with terrain-derived and vegetation predictors is likely to improve predictive models in forested areas (e.g., [54,56]).

Therefore, in this study, we augment the Cannistra et al., 2021, CNN snow cover model by using additional predictors including vegetation structure (using lidar-derived canopy height from a canopy height model (CHM) and the PlanetScope-derived Normalized Difference Vegetation Index (NDVI)), and the digital elevation model (DEM

or elevation) and its derived attributes (i.e., slope, aspect and northness). It is important to note that the NDVI as a reliable predictor for canopy is supported by its sensitivity to vegetation response despite the availability of many other vegetative indices [57]. We then evaluate if these augmentations lead to improvement upon the original band-only model performance using the produced m-scale SCA from PlanetScope imagery. In our assessments we test two hypotheses: (1) terrain-derived predictors and vegetation information improve snow mapping accuracy in both forested areas (FA) and open areas (OA), and (2) terrain-derived predictors such as slope, and aspect are more accurate over open areas than in forested areas where elevation is more important. We evaluate performance across forested areas, near the canopy edge, and in open areas (gaps) using a set of canopy classification metrics.

## 1.3 Study Area

We used three study sites for our analysis: one site for the model training/validation, the Tuolumne River Basin in Sierra Nevada of California, USA (37.89°N, –119.25°W) and two sites for additional model evaluation: Gunnison/East River Basin in the Central Rocky Mountains of Colorado, USA (39.08°N, –107.14°W) and a site near Engadin in Switzerland (46.58°N, 10.03°W) (Figure 1). We selected these sites as they are part of Airborne Snow Observatory (ASO) monitoring [43] and had overlapping PlanetScope imagery. Site characteristics are varied at the three sites; they are dissimilar in elevation, climate, and differing forest cover (sparse vs. dense) and climatic zones (maritime vs. continental). The Tuolumne basin has elevations ranging between 1500–3970 m and is in the central Sierra region of California; the area is a mix of evergreens and shrublands [58]. The Gunnison site is in southwest Colorado and has elevations ranging between

1387–4359 m [59]. The Engadin site is in the Grison region of Switzerland and ranges in elevation between 1700–2000 m; the site is primarily evergreen and composed of spruce and larch forest with homogeneous understory [14].

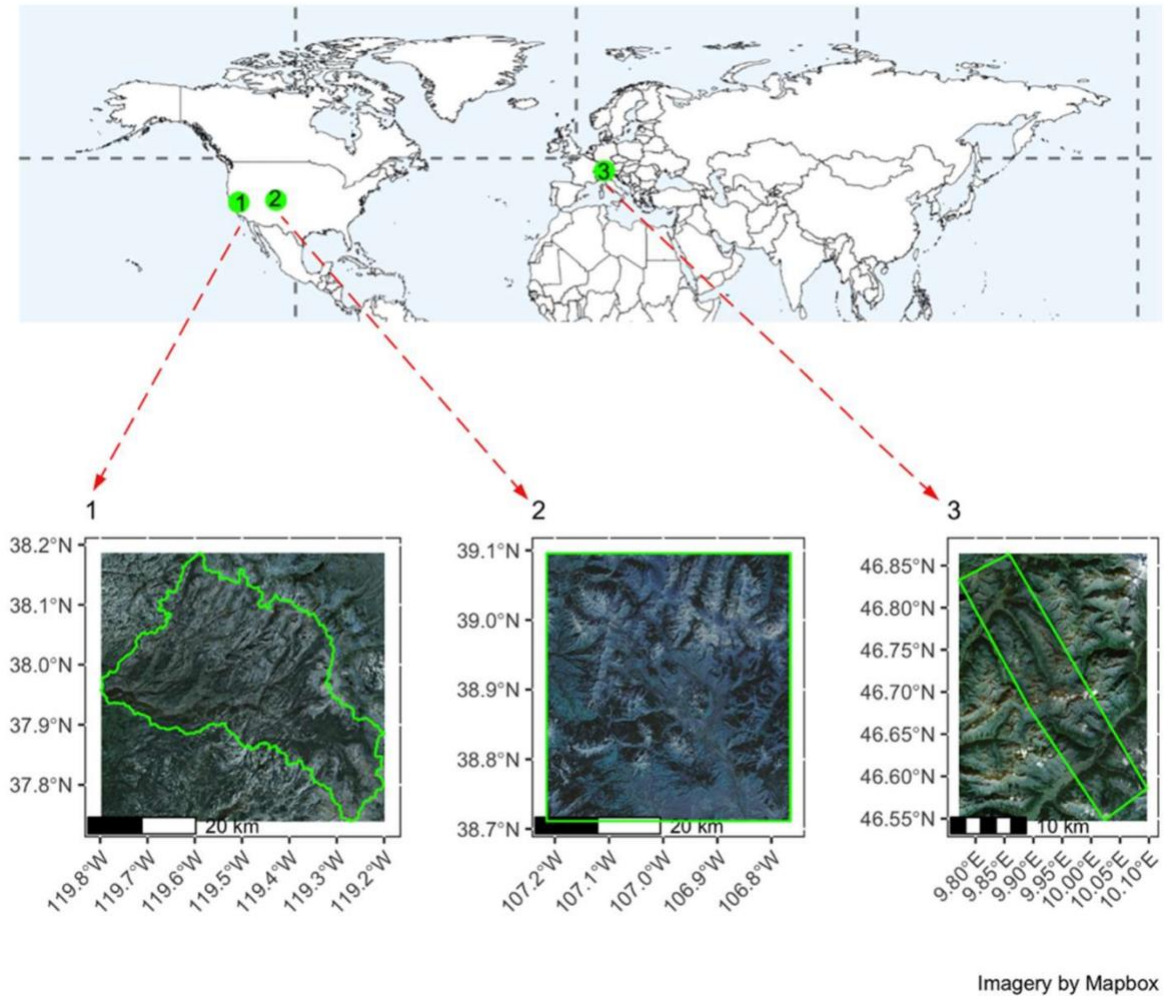


Figure 1-1. Study sites—(1) Upper Tuolumne (37.89°N, −119.25°W, California, USA), (2) Upper Gunnison (39.08°N, −107.14°W, Colorado, USA), and (3) a site near Engadin, Switzerland (46.58°N, 10.03°W, Switzerland).

## 1.4 DATA AND METHODS

### 1.4.1 Data Products

We augmented the training datasets used in Cannistra et al., 2021, to include data from multiple sites for vegetation and terrain. The overlapping PlanetScope imagery used for training and validation is provided in Appendix A. The PlanetScope constellation includes multiple generations of CubeSats starting in 2014. Many of these satellites are in a sun-synchronous orbit with 4 to 8 band radiance products at approximately 3–4 m ground-sampling distance. Earlier generations (Dove Classic) offered 4-band products (RGB and NIR), while newer generations (Dove-R and SuperDove since 2019) provide 8-bands. In this study, we use the 4-band products, which enables longer time series analyses of the full PlanetScope archive.

As terrain features and vegetation predictors are likely to improve predictability in forested areas, we used lidar-derived datasets collected as part of the original NASA ASO campaign and additionally by the Airborne Snow Observatories, Inc. The lidar-derived digital elevation model (DEM), the canopy height model (CHM), and the snow depth product (true snow cover) were provided at 3 m resolution (Figure 2) by Airborne Snow Observatories Inc. (See Appendix A for the list of ASO product IDs used in this study). We followed similar methods outlined in Cannistra et al., 2021, using a snow depth threshold value to derive benchmark snow cover masks. We tested several thresholds (Section 4.3) and determined that a threshold greater than or equal to 10 cm indicated the presence of snow in lidar-derived products [17]. More details on ASO data products and methodology can be found in Painter et al., 2016.

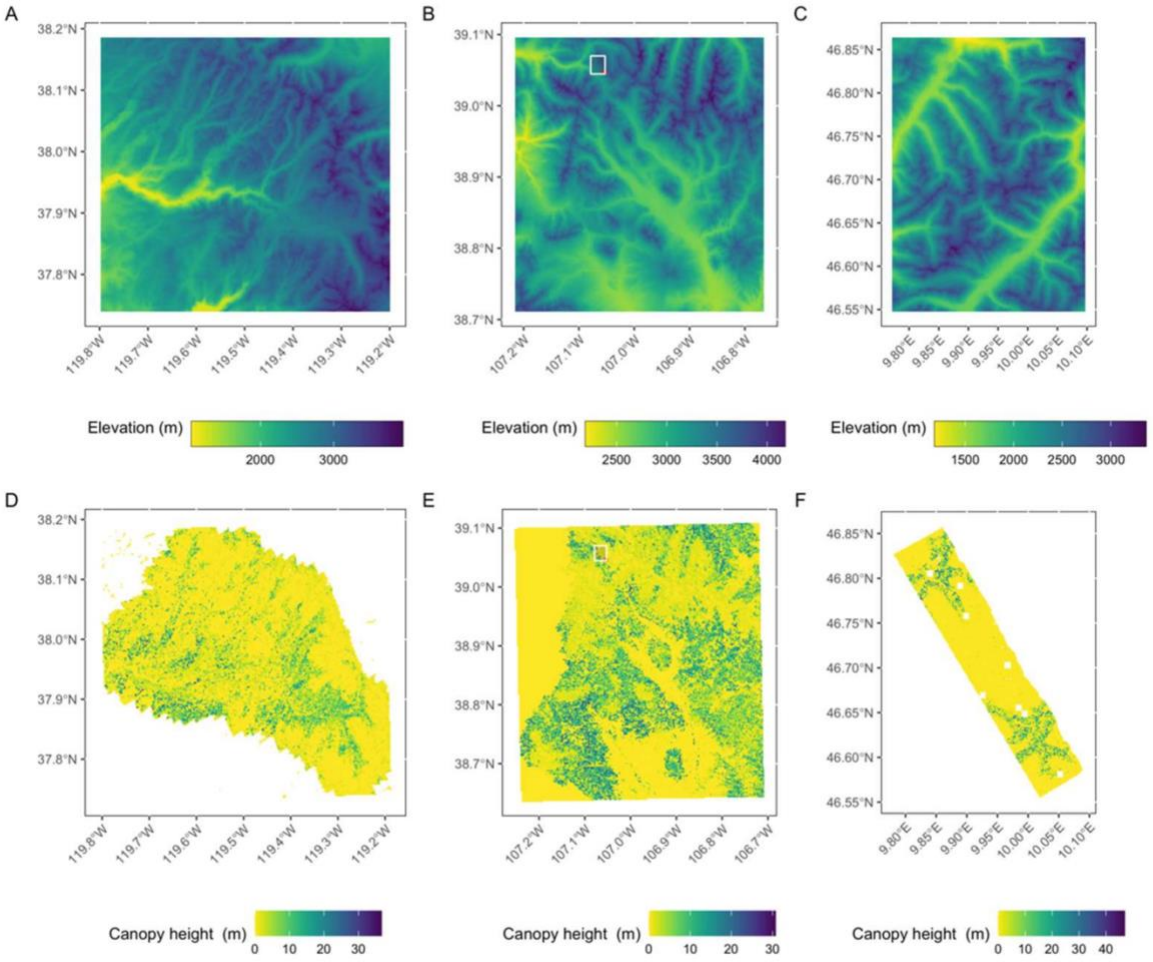


Figure 1-2. Elevation and canopy height for (A,D) Upper Tuolumne, USA, (B,E) Upper Gunnison, USA, and (C,F) Engadin, Switzerland.

A coarser level DEM was also extracted for the study areas via data provided by the Shuttle Radar Topography Mission (SRTM), a product through National Aeronautics and Space Administration (NASA) which is a 1 arc-second (approximately 30 m) product (downloaded via https://dwtkns.com/srtm30m/, accessed on 1 April 2022). The models were primarily trained on the Tuolumne dataset, except for the extension model where the neighboring San Joaquin basin was used to reproduce the original model from Cannistra et al., 2021. All the models were then validated using data from the Gunnison

River Basin, CO, USA (May 2018) and Engadin, Switzerland (May 2017) sites to evaluate model transferability.

### 1.4.2 Cyberinfrastructure

We extended the software and cyberinfrastructure developed by Cannistra et al., 2021 (Figure 3), who used a CNN-based method to identify snow-covered areas using only the 4-band PlanetScope Level-3 Analytic Surface Reflectance (SR) products. The Python-based implementation of the training procedure used PyTorch [60] and is a modified version of the "robosat.pink" software, an open-source set of command-line tools to enable machine learning with satellite imagery via "TernausNetV2"—a deep neural network that was modified after U-net for image segmentation [61]. Cannistra et al., 2021, modified the package to support the use of any N-band multispectral imagery data product and to allow the use of efficient cloud-based data storage and computation infrastructure. We similarly used highly scalable GPU-enabled "P2" instances from Amazon Web Services' (AWS) Elastic Compute Cloud (EC2) service to train the machine-learning models. Model training was performed with 50 epochs using AWS Sagemaker with default hyperparameters (learning rate of 0.000025 and batch size of 7). The code and all processing workflows are available on GitHub https://github.com/ajijohn/planet-snowcove-R (accessed on 1 April 2022).

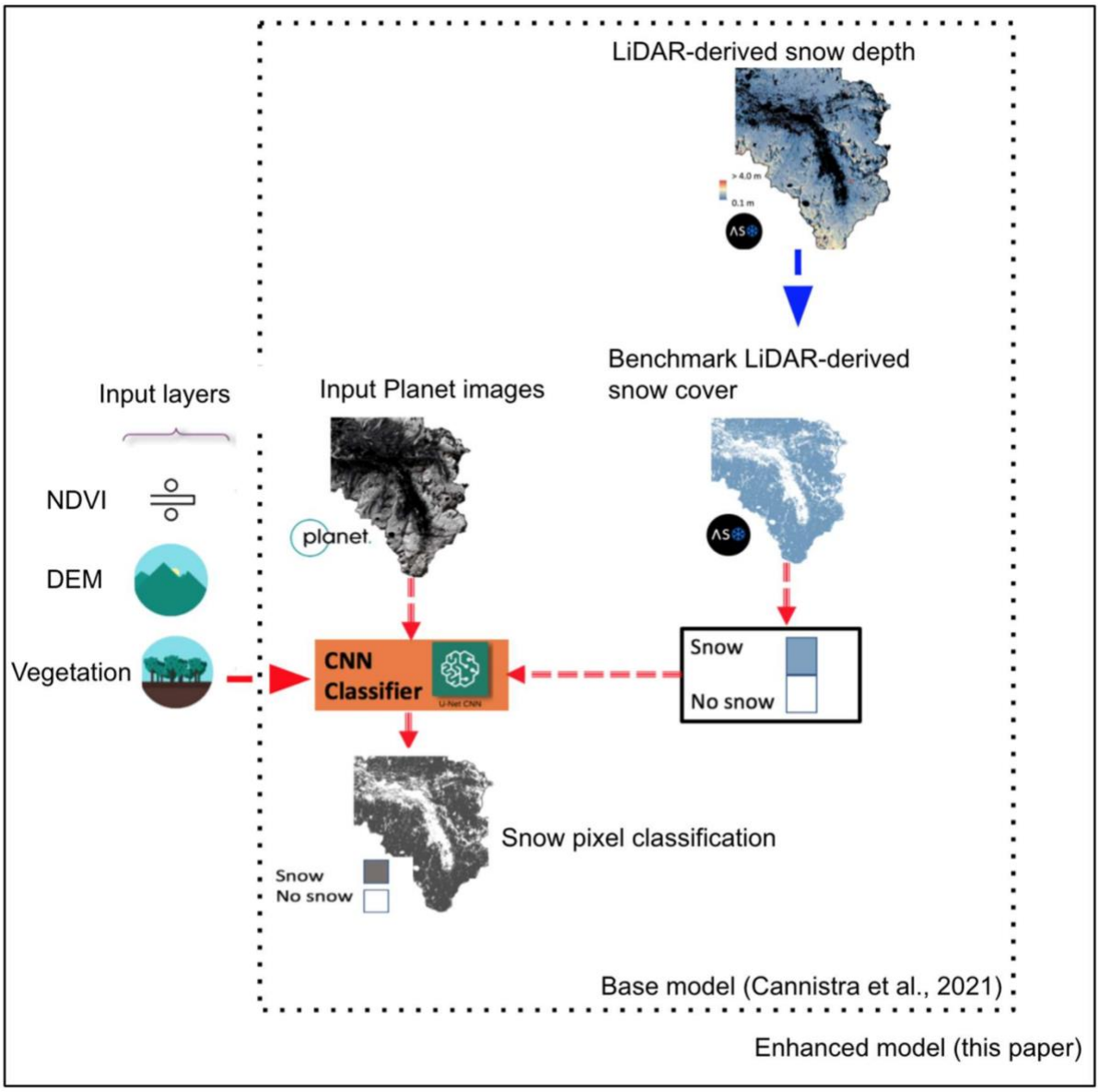


Figure 1-3. End-to-end workflow starting from pre-processing that involves downloading the Planet imagery and matching LiDAR derived snow cover data from ASO. In this study, existing bands are stacked along with DEM-based layers and vegetation proxies (Canopy height and NDVI). Model training is completed using the convolutional neural network model (TernausNetV2) on AWS with lidar-derived SCA as ground truth such as in the base model (Cannistra et al., 2021 [17]).

### 1.4.3 Model Augmentation

We call the Cannistra et al., 2021, CNN-based snow mapping model 'Model 0'. Model 0 was trained on Planet's 4-band PlanetScope data product that contains visible (Red, Green, Blue) and NIR (Near infrared); we refer to these sets of bands as the "BASE". To test hypotheses (that incorporating terrain-derived predictors and vegetation information into the model input data improves snow mapping in both forested and open areas), we enhanced Model 0 to include terrain and vegetation structure. These additional layers included DEM (elevation), slope, aspect, northness as (cosine(aspect) × slope) (e.g., see Tennant et al., 2017), NDVI, and CHM (vegetation height from the Canopy Height Model). To reiterate, we needed to improve the model performance and understand the model limitations in forested terrain. Due to that fact, we considered two types of vegetation representation in the model; through CHM and NDVI, and to draw on the influence of terrain, we used elevation (DEM) along with its derived features (slope, aspect, and northness). Slope, aspect, and northness were calculated using the methods described in Horn (1981). The Normalized Difference Vegetation Index (NDVI) is a normalized difference ratio of the NIR and red band with values ranging between −1 and +1 [62]. NDVI values around 0 generally signify non-vegetative landcover types like, water, snow, clouds, and rocks, and values closer to 1 signify vegetative growth [63]. We normalized CHM, DEM, and its derived attributes (slope, aspect, and northness) using "max–min" normalization whereby for a variable, a base is first established by taking the difference of its maximum and minimum, and then each of its values is subtracted by its minimum and divided by the base.

We tested individual input predictors alone and combinations of predictors to identify the best performing model configuration (Section 3.5 contains more details). The

best performing model from these combinations was then used for further analysis. We term this Model 1. To test if terrain-derived predictors such as slope and aspect or vegetation structure are more accurate over open areas than in forested areas, we produced three grid-level canopy quantification metrics. We describe in detail the canopy quantification metrics in the following subsections. In total, six model configurations were evaluated for complete scene-based analysis and in-depth study based on the canopy metrics.

### 1.4.4 Model Evaluation in Forested Areas

To evaluate the performance of models over different forested terrains, we derived land-cover classifications using the canopy height model (Figure 4).

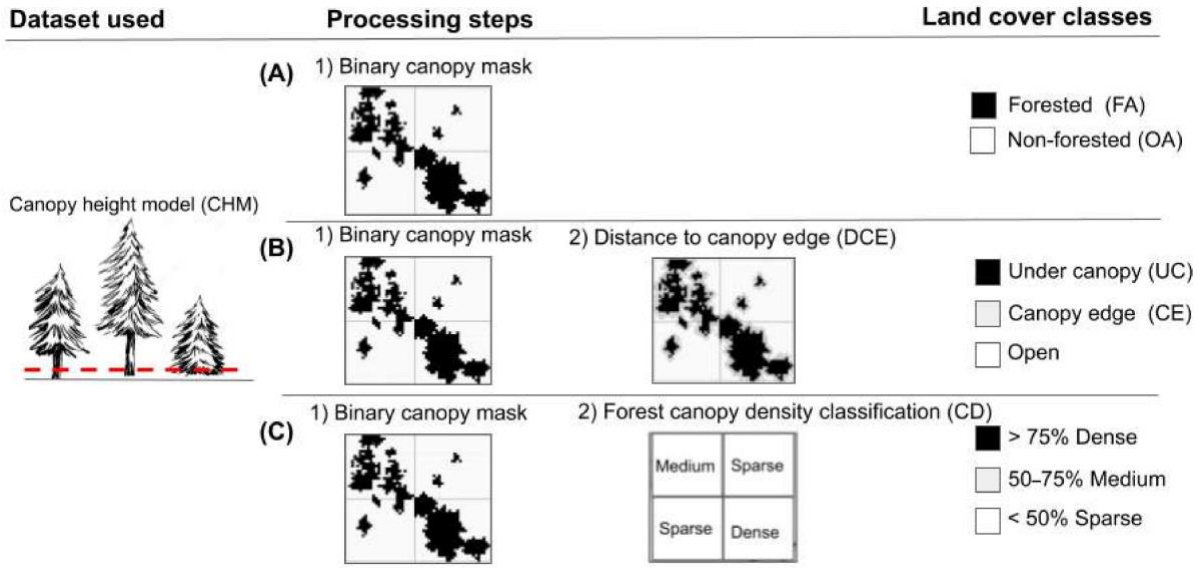


Figure 1-4. Framework for evaluating snow covered area mapping in forested terrain. (A) Classification based on canopy height; greater than 1 m (red line) is classified as forested area (FA), otherwise open area (OA). (B) Classification of land cover area: FA is composed of under canopy (UC) and canopy edge (CE), OA is outside CE; After Mazzotti et al., 2019 [12]. (C) Classification of land-cover area into forest density classes in which forest density is calculated in a larger focal pixel as a function of the vegetation presence (taller than 1 m) derived from the CHM. Forest density class is then assigned using the % of pixels. The coarse grid is 100 × 100 m.

For evaluation using binary maps derived from the canopy-height model, we used the 3 m canopy-height model (the ASO CHM) to derive a canopy-height-based binary grid (Figure 4A). We used a threshold of 1 m to separate forested areas from non-forested areas based on previous works [42] and to be consistent with previous model evaluations [17]. For evaluation using the distance to canopy edge (DCE), we use the Mazzotti et al., 2019, framework for further evaluation of the model's performance, particularly in the vicinity of the trees. For this evaluation we consider three classes as in Figure 4B: under canopy (UC), canopy edge (CE), and open. To delineate these individual classes, we use the distance-to-canopy-edge metric (DCE)—a CHM-derived metric used to study snow distribution in forested terrain [14]. Briefly, the algorithm starts with a CHM that is thresholded at a prescribed height and a binary map is produced. A 3 × 3 running mean window is then applied to identify the canopy and near canopy area and then separate the area under the canopy (UC, black shade) and the canopy edge (CE, grey shade) using a threshold. We used a threshold of 30 m to separate the UC and CE areas. For comparison, Mazzotti et al., 2019, used a threshold in the range of 1 to 8 to categorize gaps, a range of −1 to 1 for edges, and a range of −1 to −3 to signify canopy clusters using a 1-m resolution CHM grid. The CHM used here to generate the DCE metrics in our study had a 3 m spatial resolution; therefore, we scaled the DCE accordingly to account for scale effects. We reclassified the DCE into three classes: values less than 0 were classified as UC—under the canopy, 0 to 30 as CE near the canopy edge, and greater than 30 as OA—open areas. UC refers to small or large clusters of canopies, CE refers to the edge of a canopy, and the Open category signifies gaps or open areas.

For evaluation using canopy density (CD), we used the 3 m canopy height model (CHM) product to produce a canopy density metric (Figure 4C). We first applied a threshold of canopy height greater than 1 m as a binary threshold and then used a moving average filter of 3 by 3 m to smooth it. The density was then calculated using percentage of grid cells with a canopy greater than 1 m within an area of about 100 × 100 m (Figure 4C, second processing step). Classifications were then made into Sparse, Medium, and Dense classes as a function of the percentage of vegetation pixels in the 100 × 100 m window (Figure 4C). More than 75% was considered Dense, between 50% and 75% was considered Medium, and anything less than 50% was labeled as Sparse. For example, a canopy pixel marked as Dense corresponds to 75% of pixels in the grid having vegetation of a height greater than 1 m.

### 1.4.5 Model Performance Metrics

We evaluated the performance of models over entire watersheds. In addition, we examined the models' performance in more detail in a subset of study areas and separately calculated the various land classes as described above. We used the following performance scores: Precision, Recall, F-score, Accuracy, and Balanced accuracy [64,65]. Model-derived SCA was compared with lidar-derived SCA snow cover (binary) at pixel level for all the metrics. "TP" ("true positives") is the number of pixels where snow is correctly classified as snow, and "TN" ("true negatives") is the number of pixels correctly classified as snow-free. "FP" and "FN" ("false positives" and "false negatives") represent the number of pixels that are incorrectly classified as snow and snow-free, respectively.

Precision (1) is the proportion of all true snow classifications (i.e., also true snow classifications in the lidar-derived SCA) of the model divided by all the snow

classifications. Precision tells us how precise a model is, i.e., how good it is at detecting snow where there is in fact snow.

$$\text{Precision} = \text{True Positives} \div (\text{True Positives} + \text{False Positives}) \quad (1)$$

Similarly, Recall (2) is the proportion of all the true snow classifications by the model divided by all the predictions (correctly or incorrectly identified as snow) by the model. Recall tells us how good the model is at detecting all the snow pixels. Usually, a trade-off between Precision and Recall is preferred, but this can depend on the domain of which one is given more preference.

$$\text{Recall} = \text{True Positives} \div (\text{True Positives} + \text{False Negatives}) \quad (2)$$

The F-Score (3) is the harmonic mean of Precision and Recall that provides an overall performance that translates to the usefulness of the model. It ranges between 0 and 1 with higher values meaning a better predictive ability of the model.

$$\text{F} - \text{score} = 2 * \text{Precision} * \text{Recall} \div (\text{Precision} + \text{Recall}) \quad (3)$$

Another overall performance metric of the model we use here is accuracy (4), defined below.

$$\text{Accuracy}=(\text{TP}+\text{TN}) \div (\text{TP}+\text{TN}+\text{FP}+\text{FN}) \quad (4)$$

Balanced accuracy (5) on the other hand, normalizes the "TP" and "TN" predictions when the corresponding binary classes are imbalanced (for example, "TP" are disproportionately larger than "TN"). Furthermore, the metrics were grouped by various canopy quantification metrics in a subset of study areas.

$$\text{Balanced accuracy} = (\text{True positive rate} + \text{True negative rate}) \div 2 \quad (5)$$

### 1.4.6 Role of DEM Resolution in DEM Based Models

In our model development, we use high-resolution, m-scale lidar-derived DEM data, which is only available on select sites. In this step, as an alternative, we evaluate the use of a publicly available global DEM dataset such as SRTM in our models to assess the degree of degradation in performance of our models when using coarser-resolution digital-elevation-model (DEM) products. To test the influence of the DEM quality on the performance of our model, we upsampled the 30 m DEM (using SRTM) and extracted DEM-based attributes (slope, aspect, and northness) for training over the Tuolumne. We then retrained the models with the upsampled DEM data and used the model to validate the SCA at the Engadin and Gunnison sites.

## 1.5 RESULTS

### 1.5.1 Overall Model Performance

We generated high-resolution SCA maps across three geographically diverse areas (Tuolumne Basin, CA, USA, Gunnison, CO, USA, and Engadin, Switzerland) for the various model configurations described above by training only on data from the

Tuolumne basin. The SCA data for the Tuolumne basin were created only for the test dataset and was subsequently used for performance evaluation.

Table 1-1. F-scores of combinations of features (predictors). Metrics were calculated across 19 scenes in Gunnison, and 2 scenes in Engadine. A scene is approximately 24km by 8km.

| **Features** | **Tuolumne** | **Gunnison (n=19)** | **Engadine (n=2)** |
|---|---|---|---|
| BASE[1]+ DEM (Elevation) | .91 | .88 | .91 |
| BASE[1]+ slope + aspect + northness | .67 | .73 | .90 |
| BASE[1]+ DEM (Elevation) + slope + aspect + northness | .90 | .87 | .85 |
| BASE[1] + NDVI | **.92** | **.89** | **.93** |
| BASE[1] + CHM (Canopy height) | .85 | .85 | .92 |
| BASE[1] | .32 | .85 | .90 |
| BASEEXT[1,2] | .76 | .88 | .92 |

[1] refers to 4-bands (Visible and NIR) and tiles from Tuolumne basin only; n represents the number of scenes. [1,2] refers to extended training that includes the San Joaquin basin

Table 1 shows the F-scores of all the predictor combinations. The results for Gunnison and Engadin indicate the model's performance over the entire study area. The least performing model was the one which used only 4-bands (visible and NIR), that we earlier defined as "BASE". The best performing model was the combination of base imagery with the NDVI and was able to perform better at Gunnison and Engadin. The addition of DEM (Elevation) resulted in better scores and derived features that included slope, aspect, and northness and resulted in lower scores than the NDVI-based model. The addition of CHM (canopy height) resulted in better scores than the BASE model and lower scores than the BASE + NDVI- or BASE + DEM-based models. The model that used slope, aspect, and northness was the least performant. In general, across the Engadin site,

all candidate models performed better than the BASE model. The best performing model, BASE + NDVI, is referred to as Model 1 from here onwards. We limited the use of the BASEEXT model that included part of the San Joaquin basin, because we were able to obtain better overall performance in Gunnison and Engadin with a reduced training size and by only using existing bands. BASEEXT is the equivalent of Model 0 developed by Cannistra et al., 2021. Table A2 and Table A3 contain detailed metrics. In addition to the predictor combinations in Table 1, BASE + DEM + CHM and BASE + DEM + NDVI were also evaluated but were dropped from further analysis because of low scores (F-scores of 0.65 and 0.31, respectively).

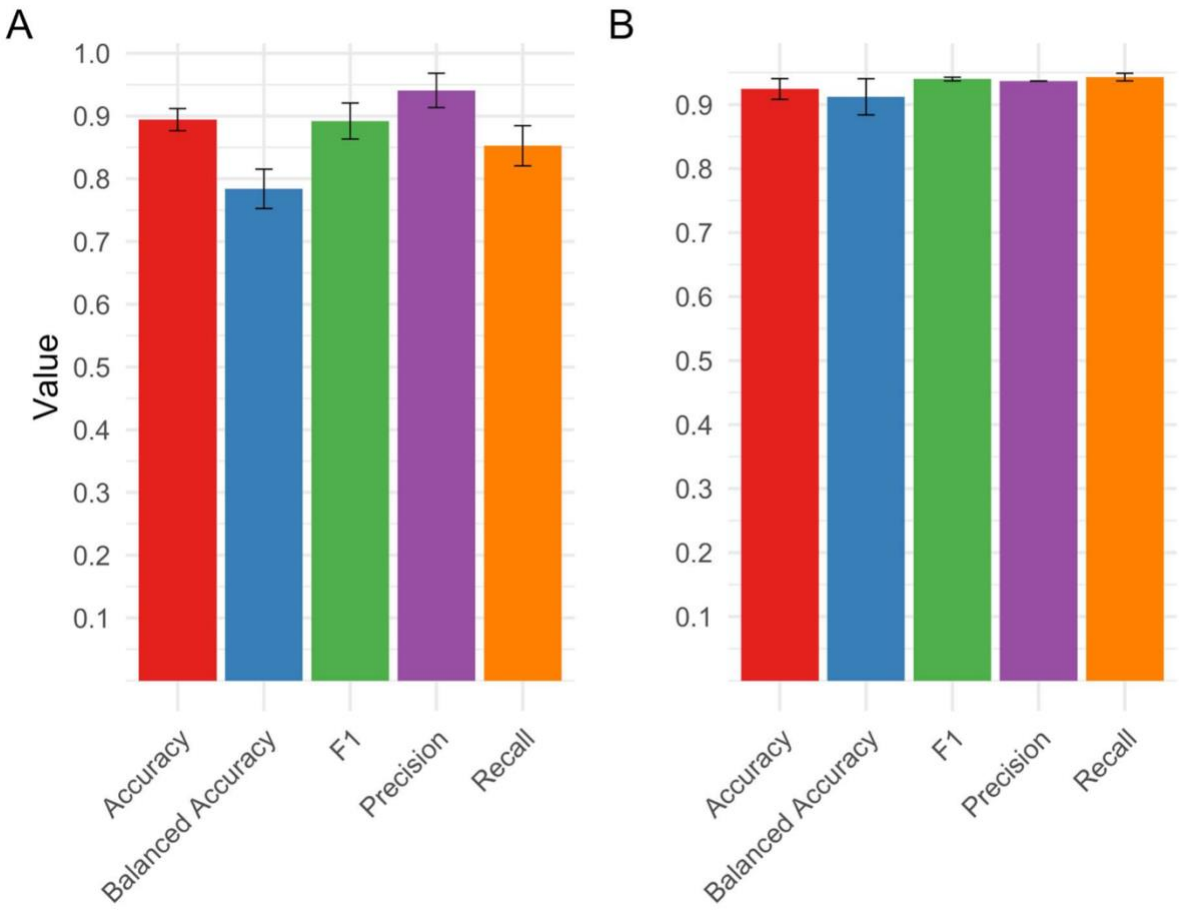

Figure 1-5. Overall performance (BASE + NDVI or Model 1) at Gunnison River and Engadin sites. (A) Mean metrics for Gunnison River basin, CO, USA; (B) mean metrics for Engadin, Switzerland. Bars show standard deviation.

Detailed metrics of the models provide insights into the intricacies of the model's performance (Figure 6). The NDVI-based model has the best-balanced accuracy, evident at Gunnison and Engadin sites (0.78 and 0.91). At the Gunnison site (Figure 6A), we note that even though BASE + NDVI is the best model, its recall is lower (0.85) compared to the slope + aspect + northness model (0.9); a similar pattern is found at the Engadin site as well. The DEM-derived model (DEM, slope, aspect, and northness) and DEM-only models consistently achieved the second-best performance at both sites. The BASE model is found to have the same precision as the best model (0.94) at the Engadin site.

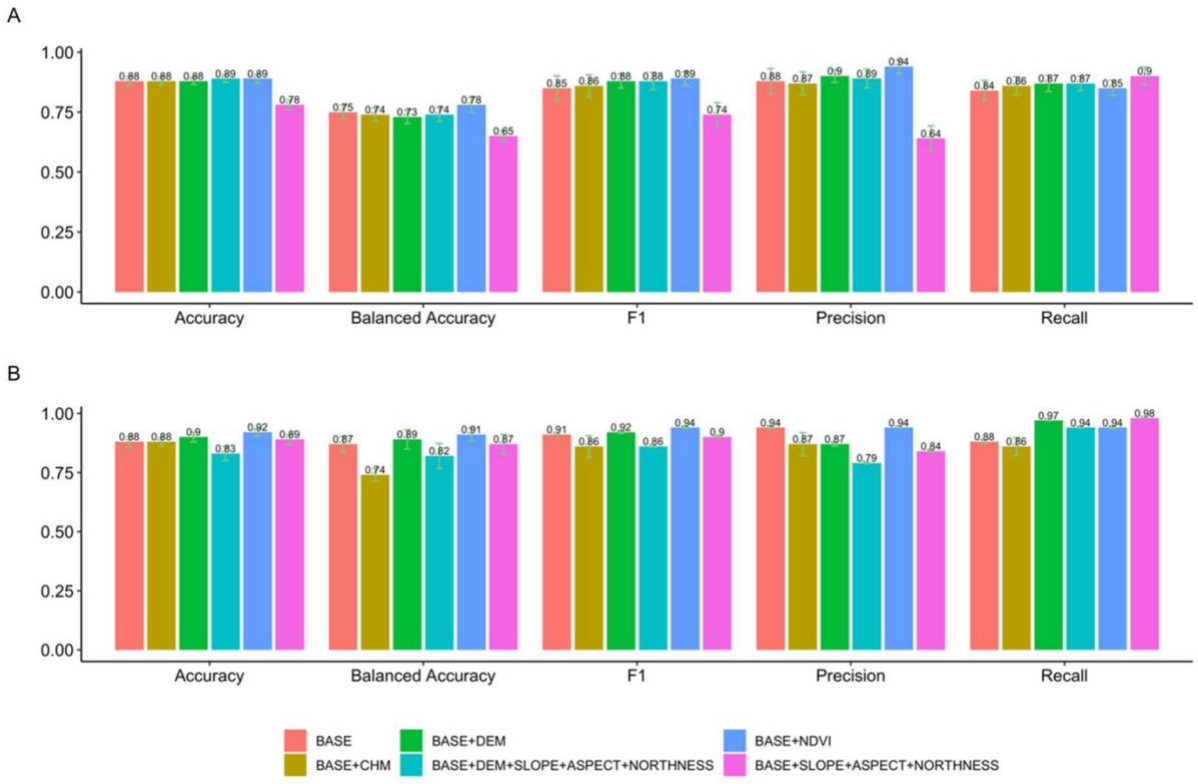

Figure 1-6. Detailed performance metrics of all the models at both the evaluation sites. (A) Complete metrics for Gunnison River basin, CO, USA. (B) Complete metrics for Engadin, Switzerland. Bars show standard deviation.

To examine the spatial distribution of snow in forested terrain more closely, we picked two areas within the larger Gunnison River domain that we further investigated (identified by white and red rectangles in Gunnison (Figure 2B,E)). Figure 7 depicts the evaluation of a larger area, which is mostly dense canopies, and Figure 8 shows the evaluation across a smaller area that is mostly open with sparse trees. The spatial distribution of the mapped snow of a sample site in a dense canopy and an open site at Gunnison show the nuances of the model's performance (Figure 7 and Figure 8). Qualitatively, models generally underpredict snow in dense forest stands (lower left section of Figure 7), but DEM- and NDVI-based models perform comparatively well in comparison to rest of the models (Figure 7D,G). The lidar-derived SCA (Figure 7H) indicates that there is snow under dense canopies. Figure 7A shows the Planet image of the area, in which snow can be observed across the forested area. Therefore, the models, except for one, also predicted no snow, although the lidar dataset shows snow under the canopies. The CHM and BASE model overpredicts the snow in open (see along the diagonal and in the upper right section (Figure 7B,E)). The model with DEM and derived attributes (slope, aspect, and northness) overpredicts in open areas but is able to obtain snow predictions correctly in ridges and valleys (see Figure 7C). The model Base + Slope + Aspect + Northness in general overpredicts snow in open areas and under the canopies (Figure 7F and Figure 8F). The Base + NDVI model performs well across open areas (top right of Figure 7D and Figure 8D). In the ensuing subsections, we describe these results further with detailed canopy quantification metrics.

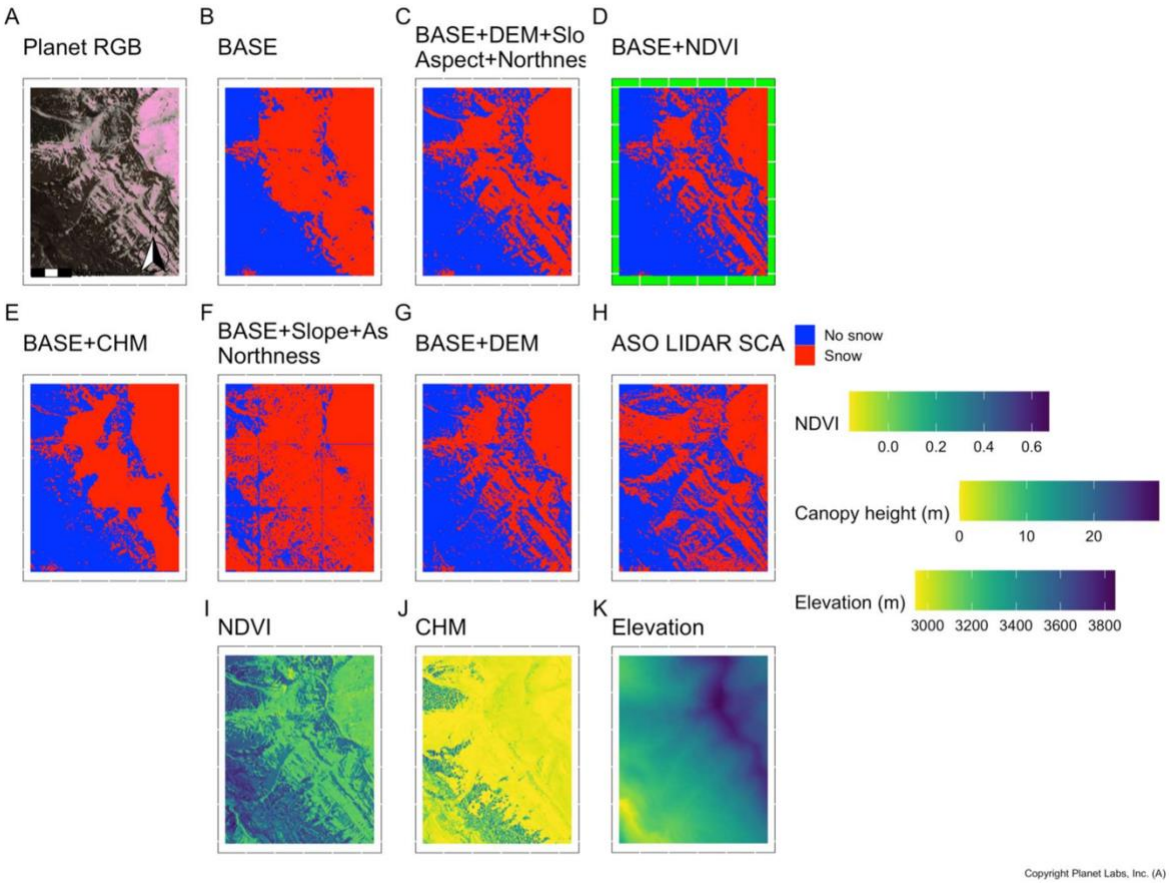


Figure 1-7. Performance of models in a sample area in Gunnison, which is in dense canopy and has canopy edges and open areas. (A) Planet RGB imagery showing the sample area in the Gunnison study site. (B) SCA predicted by the BASE model. (C) SCA predicted by the BASE + DEM (Elevation) + slope + aspect + northness model. (D) SCA predicted by the BASE + NDVI mode, the best performing model. (E) SCA predicted by the BASE + CHM model. (F) SCA predicted by the BASE + slope + aspect + northness. (G) SCA predicted by the BASE + DEM. (H) Lidar-derived SCA. Total study area of approximately 6.25 km$^2$. The highlighted model in green is the overall best model, Model 1. (I) Corresponding NDVI of the large study area derived from PlanetScope image. (J) Canopy height of the study area that shows the density in canopy cover. (K) Elevation of the study area. Note that the tiles were aggregated for this visualization and show some visible artifacts (border lines). The satellite imagery in panel (A) was downloaded from Planet Labs, Inc. All rights reserved.

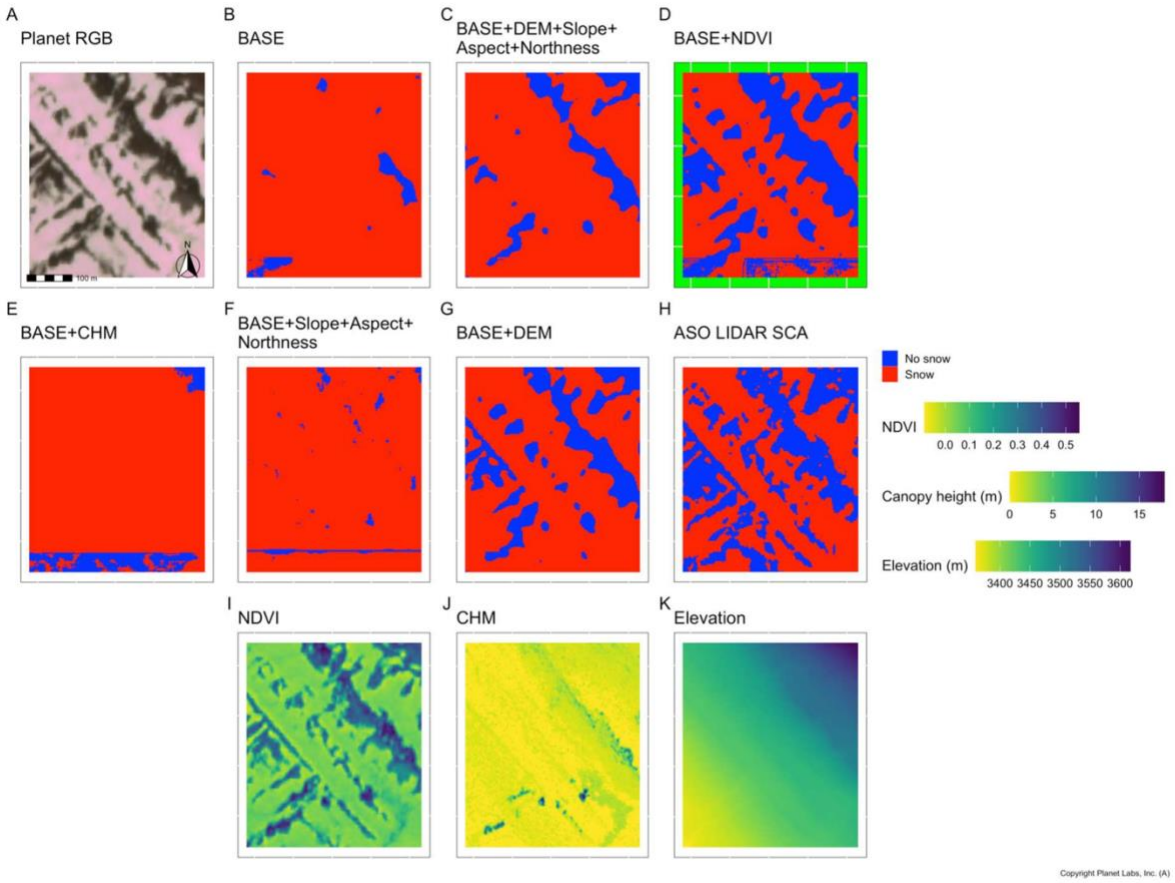


Figure 1-8. Performance of models in a sample area in Gunnison that is predominantly an open area. (A) Planet RGB imagery showing the sample area in the Gunnison study site. (B) SCA predicted by the BASE model. (C) SCA predicted by the BASE + DEM + slope + aspect + northness model. (D) SCA predicted by the BASE + NDVI model, the best performing model. (E) SCA predicted by the BASE + CHM model. (F) SCA predicted by the BASE + slope + aspect + northness. (G) SCA predicted by the BASE + DEM. (H) Lidar-derived SCA. Total sub study area of approximately 0.25 km$^2$. The highlighted model in green is the overall best model, Model 1. (I) Corresponding NDVI of the sample area derived from PlanetScope image. (J) Canopy height of the sample area that shows sparseness in canopy cover. (K) Elevation of the study area showing little variation. Note that the tiles were aggregated for this visualization and show some visible artifacts (border lines). The satellite imagery in panel (A) was downloaded from Planet Labs, Inc. All rights reserved.

### 1.5.2 Evaluation of SCA over Open and Forested Areas Using Canopy Metrics at Gunnison

At the Gunnison site, snow-mapping performance in the open and forested areas improved when DEM and vegetation-based predictors were used, and performance was better in open areas and at canopy edges than under the canopies. Table 2 shows the results using the various canopy metrics and land-cover classifications. In the open areas all the models but one, which included only slope, aspect, and northness, performed better than the BASE model. Using the CD metric, the vegetative models have better performance in sparse and mediumly forested areas compared to densely vegetated areas. In dense areas, the models with NDVI and DEM show similar performance. Using the DCE metric, we see better results in open and canopy edges across all the models except in the case of slope, aspect, and northness. The model with DEM and its derived attributes was found to be better under the canopy compared to all other models. Using a cut-off such as the CH metric, NDVI and DEM models were comparable in performance across open and forested areas.

Table 1-2. Performance metrics (F-score) in forest classes delineated using different forest canopy metrics within the Gunnison site across all the models. CD = canopy density, DCE = distance to canopy edge, CH = canopy height, CE – Canopy Edge, UC – Under Canopy. The results were derived from 19 PlanetScope scenes for the Gunnison area.

| Canopy Metrics | CD | | | DCE | | | CH | |
|---|---|---|---|---|---|---|---|---|
| Model | Sparse | Medium | Dense | Open | CE | UC | OA | FA |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| BASE + DEM | .89 | .88 | .86 | .84 | .89 | .86 | .89 | .87 |
| BASE[1]+ slope + aspect + northness | .69 | .74 | .79 | .67 | .69 | .78 | .70 | .74 |
| BASE[1]+ DEM (Elevation) + slope + aspect + northness | .88 | .87 | .85 | .86 | .87 | .92 | .88 | .87 |
| BASE[1] + NDVI | .90 | .88 | .86 | .89 | .90 | .87 | .90 | .88 |
| BASE[1] + CHM (Canopy height) | .85 | .86 | .85 | .85 | .86 | .85 | .86 | .85 |
| BASE[1] | .85 | .85 | .84 | .82 | .85 | .85 | .84 | .86 |

[1] Refers to Visible + NIR bands

### 1.5.3 Evaluation of Lidar-Derived SCA Threshold

The threshold that is used to produce a binary snow-covered area dataset from lidar-derived snow depth measurements might affect the performance, because snow depth is sensitive to the location of canopy edges. To assess the impact of this threshold on performance, we evaluated the change in the models' performance after varying this threshold. The threshold used to derive the lidar-derived SCA was varied across a range of thresholds at the Engadin site. We varied the threshold from 3 cm to 20 cm and recalculated the performance scores. The evaluation was carried out using the DEM-derived model (DEM, slope, aspect, and northness). The performance slightly decreased when the threshold was lower than 8 cm and above 10 cm (Figure 9).

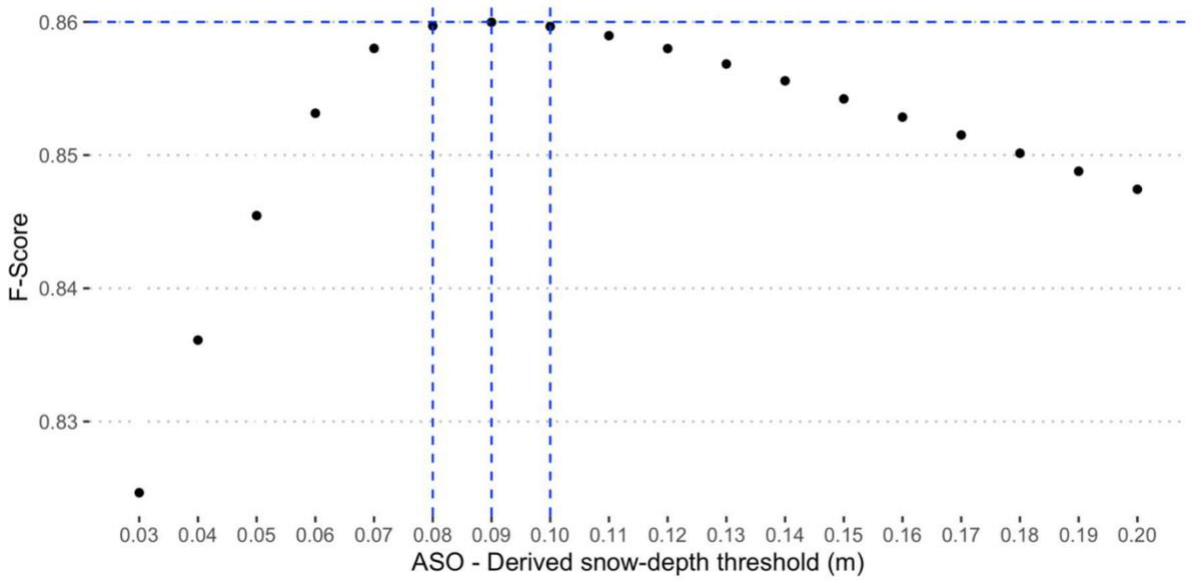


Figure 1-9. Model performance of DEM with derived attributes model when the lidar-derived threshold for snow cover was varied from 3 cm to 20 cm.

### 1.5.4 Sensitivity of the DEM Resolution for Training

Model performance at both the evaluation sites was comparable when the 30 m SRTM based DEM was used in the model training. The model using DEM with derived attributes (DEM, slope, aspect, and northness) was chosen for this assessment as it was the ideal model to evaluate the nuances of DEM resolution. At the Engadin site, the models trained using 3 m and 30 m DEM data had comparable performances, but across the Gunnison site the performance was different across these two model types. In particular, the model trained using the 3 m DEM data had better recall and accuracy. At the Gunnison site, the SRTM-based model had lower precision than the 3 m model in the Gunnison evaluation, but the F-scores were nearly the same (Figure 10). The SRTM-based model was calculated over 18 scenes versus 19 scenes for the 3 m model, this was because the coarser SRTM model failed to distinguish the snow-vs.-no snow in one of the scenes at Gunnison.

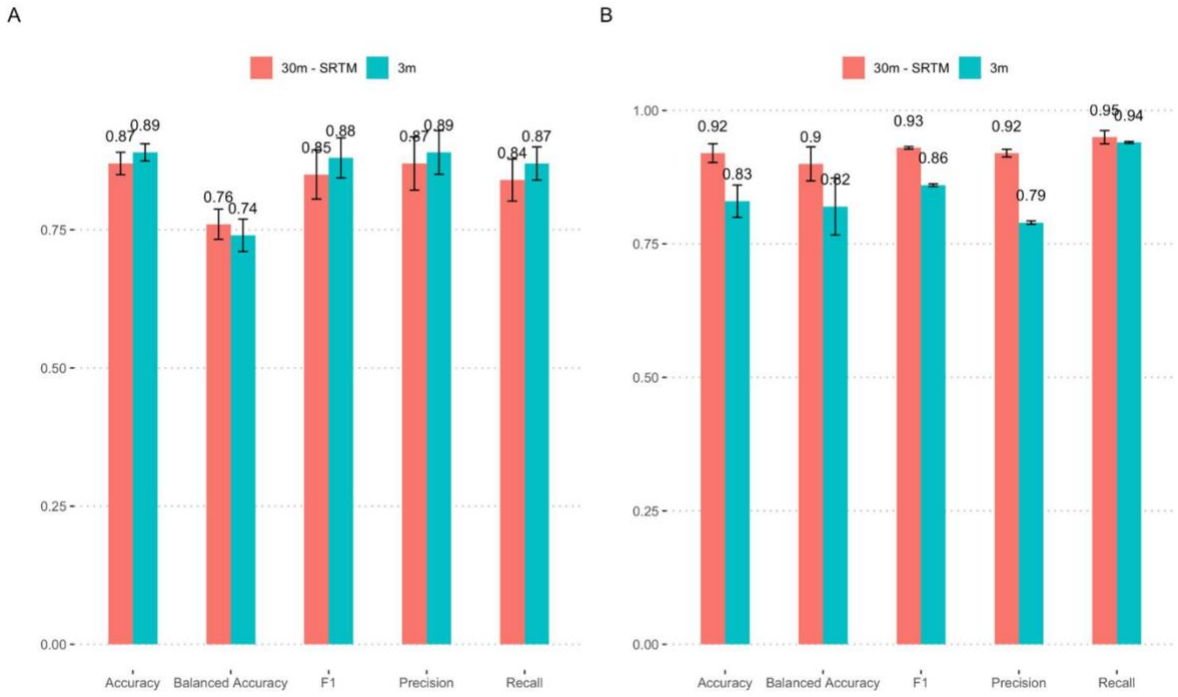


Figure 1-10. Model performance of DEM with derived attributes (DEM, slope, aspect and northness) model when 3 m was evaluated against a 30 m DEM (via SRTM). (A) Performance metrics over the Gunnison site. (B) Performance metrics over Engadin site. Bars show standard deviation.

## 1.6 DISCUSSION

Overall, our results suggested that combining Planet with other landscape- or satellite-derived landscape features (especially elevation and vegetation) will improve predictions of snow cover, especially under dense canopies. However, modeling approaches, site locations, and data availability added complexities to predictions, implying that higher-quality remote-sensing products and further methodological improvements could further improve snow-cover predictions from Planet data. In our study, we evaluated the addition of new features to an existing CNN-based model originally based only upon unaltered inputs of Planetscope band data to map snow-covered areas from Planetscope imagery at m-scale spatial resolution. These additional features included a vegetative index (NDVI),

canopy height, DEM, and topographic features (slope, aspect, and northness, the variables derived from the DEM).

Our hypothesis that terrain and vegetation structure improve the mapping of snow in forested areas was substantiated by the more accurate predictions of models using NDVI (proxy of canopy greenness) rather than terrain-based features (DEM, slope, aspect, and northness). The NDVI-based model was an overall best-performing model, and the DEM-derived model was better for pixels right under trees. We disproved our second hypothesis (where we suggested that the addition of slope, aspect, and northness features would improve performance more in forested areas than in open areas), with our finding that models including these factors overestimated snow in open areas contrary to our expectation.

### 1.6.1 NDVI Model (Preferred) Performance

We see in general that all the models perform better with band-based models (models which only use the spectral bands, which includes NDVI as well) still outperforming DEM and canopy-height-based models. Given that the F-scores across all our predictor combinations are in near proximity (except for the model with slope, aspect, and northness), all of them are suitable candidate models. The band-based models capture the dynamics of snow cover, but the derived vegetation index (NDVI) is more precise (Precision of 0.88 vs. 0.94). Furthermore, the NDVI-based model has the highest balanced accuracy (0.78), a metric valued in case of imbalanced classes. The DEM-based models (except for slope-, aspect-, and northness-based model) have similar performances (both have F-scores of 0.88), suggesting that they would have similar predictive abilities, but that the DEM only model might be a more parsimonious choice.

Canopy-based models are comparable to "BASE" (F-score of 0.86 and 0.85), suggesting that adding canopy height on top of the spectral bands does not drastically improve the model performance. The slope, aspect and northness model is shown to be less precise and less accurate (F-score of 0.65), but it does show high Recall. This is likely because of a high proportion of true positives (pixels correctly classified as no snow) compared to a significantly smaller number false negatives (pixels incorrectly marked as snow) but a higher number of false positives (pixels incorrectly marked as no snow).

We show that multispectral bands alone are sufficient in mapping snow-covered areas in forested areas. It is well known that band indices work well for deriving SCA, for example, MODIS-based snow cover product uses a combination of NDVI and NDSI to determine snow cover [66] and also more recent CNN-based models have found spectral bands to be sufficient at determining snow cover [17,46]. The NDVI-based model (Model 1) was the best model (F-score of 0.89) in our study. In this case, the model is likely associating the range of NDVI with presence or absence of snow, i.e., the higher the value of NDVI (closer to 1 for trees) and lower the value (close to −1) is getting associated as no snow (Figure 7E). However, this result is contrary to some other studies where NDVI was found to be limiting as a predictor [45,46]; however, this could be because of the coarser resolution of the product [67] that was used (MODIS has a 500 m multispectral resolution vs. the 3–5 m multispectral resolution of PlanetScope). Several studies using ML affirm the importance of multispectral bands in determining snow cover regardless of spatial resolution [45–47,66]. We find that other models were also relatively good, F-score varied by 15% across the remaining models; the elevation-based models being the second best, elevation is an important driver to the presence of snow and hence is better at delineating

snow cover. NDVI is calculated using the existing bands, so no additional datasets are required.

### 1.6.2 Effect of DEM (Elevation) and Its Derived Attributes

Models using the DEM-derived attributes were also better at classifying snow in forested and open areas. This is consistent with other studies where contributions of northness, slope, and aspect have been shown to influence the presence of snow [47,52,54,68]. We found, in general, that the DEM-based models show comparable performance across different geographic areas (Figure 6A,B). Although they still underestimate snow in forest understories, they have better performance than the NDVI and CHM-based models. At Gunnison, the performance of the DEM-derived model (especially the one with slope, aspect, and northness) is better (F-score of 0.92 in Gunnison site using DCE) under canopies than rest of the models, but the same model is comparable to rest of the models in Engadin (to note, we had less snow on the ground, see Appendix B).

### 1.6.3 Effects of the DEM Resolution on the Training Performance

We show the accuracy of SCAs generated using PlanetScope imagery in forested complex terrain is also subject to DEM resolution. The use of the fine resolution DEM (3 m) is found to improve the detection of snow (Figure 7 and Figure 8) over the 30 m resolution DEM, the 3 m model provides us with an accuracy of 0.85 vs. 0.78 from the SRTM-based model (Figure 10). However, we posit that the performance achieved in this study using coarser-level DEM might be reproducible in other geographic areas, because of the ready availability of coarser-resolution DEMs with global coverage [69].

### 1.6.4 Effects of the DEM Resolution on the Prediction Performance

We find that the coarser DEM was a reasonable replacement for the 3 m DEM. Specifically, the F-scores were comparable whether models incorporated the 3 m or the coarser DEM as the input (Figure 10). Over the Gunnison site (n = 19), the accuracy score with the 3 m DEM model with derived attributes as its input was 0.89 in comparison to a similar model that was trained using a coarser DEM, which gave a score of 0.87—a loss in accuracy of 2%. Moreover, with the coarser DEM, the model was not able to distinguish snow in 1 out of the 19 scenes. The difference in accuracy is not substantially different, but it does highlight the limitations in the use of coarser DEM when using DEM-based models. However, we deem that using DEM features from existing publicly available DEMs such as SRTM will ensure reliable model performance at other sites. Moreover, fine-resolution DEMs are not available everywhere, so the promise of a coarser DEM with a slight decrease in accuracy suggests that the model is theoretically applicable in many more geographic areas than it would be if only fine-resolution DEMs were used.

### 1.6.5 Applicability of Explored Models

Models can represent the SCA across forested terrain, but their skill in doing so varies as a function of forest-cover density. We note that the NDVI-based model can map the snow in open areas, gaps, and areas with sparse trees. The use of canopy quantification provided insights into model performance across canopy classes. Generally, models (including the best performing NDVI based model) performed better at canopy edges and in open areas than in the under-canopy areas. Models also showed significantly better performance in the under-canopy areas at the Engadin site than Gunnison. The F-scores

across Gunnison ranged between 0.78–0.90 and were 6% higher than when using the BASE model.

The use of different canopy classes allowed us to benchmark model performance as a function of land-cover classifications and identify the critical thresholds at which the models succeed or fail. The addition of NDVI improves the performance (F-score of 0.87) of mapping snow cover in forested areas (Under Canopy metric in Table 2) and is better than CHM (Canopy height)- and DEM (Elevation)-based models in these areas [66,70] , but the inclusion of DEM with slope, aspect, and northness is far more accurate (F-score of 0.92). The use of DEM derived attributes improves the prediction under the canopies where optical methods clearly have disadvantages.

We also note model improvements in different land-cover classes—open, sparse medium, and forested areas, with the performance varying by geographical area by 25% in F-score across the Gunnison site. Snow-covered tall vegetation would have a relatively higher albedo than snow-free shorter vegetation (<1 m) that has lower albedo; hence, optical collections such as PlanetScope are able to detect snow in low vegetative and open areas.

We document geographic differences in snow predictability in all our model evaluations. The F-score performance metric varies between 0.73 and 0.93. The snow characteristics at the two test sites inherently differ because of geomorphic and climatic differences. The characteristics of snowmelt dynamics are complex, as the effects of physiographic features play vital roles in snow accumulation and melt [52,71,72]. This is particularly relevant to our study which used training data from the snow ablation period, where snow distribution is most heterogeneous. We also caution that this might limit the model's transferability in a different geographical area (e.g., the Tibetan plateau [73]). A

multi-site composite training might alleviate this issue and could be explored as a next step [46].

The vertical accuracy used in thresholding snow-depth was found to be reliable. The threshold value used for lidar-derived snow-classification played a minimal role in the differences we observed in the model's performance (Figure 9). We find that the performance is similar for thresholds from 8 cm to 10 cm and is lower for below 8 cm and above 10 cm. This also suggests that any threshold between 8 and 10 cm should suffice in model development, a finding echoed in other works as well [17,43].

### 1.6.6 Model Feature Selection and Training Volume

Adding more features increased the training time proportionally, for example, a BASE model took an average of 4 h compared to a DEM-derived model (DEM along with slope, aspect, and northness) took five times longer. In machine-learning models, a simple model is preferred, as it prevents overfitting [74,75]; increasing the number of dimensions in the model input makes it more likely that the model captures both real and random effects.

### 1.6.7 Limitations

Firstly, we chose a limited set of predictors based on our physical understanding of snow dynamics in the system, and we believe it is likely that the models we ultimately selected would also perform well at predicting snow under trees in other snowy and forested systems. However, we suggest that those wishing to build similar predictive models in their own system consider other predictive variables to optimize their model to their local circumstances.

Secondly, the NDVI model captures the terrain dynamics as shown by the canopy quantifications but still misses snow in dense canopies. This is likely limited because of correlated reflectance in PlanetScope bands [76], variable radiometric quality, and the general difficulty in capturing the reflectance of snow via optical-methods in forest understories [70]. We speculate similar performance when other fractional band measures (e.g., the use of Green index that is different from the NIR and Green bands that are linked to forest canopy [77]) are considered, again, because of narrower bands by PlanetScope and the high degree of correlation between the bands [78]. We expect that the improved radiometric resolution and addition of new bands (such as the red edge) will help improve the predictions of snow in forest understories in the near future. Moreover, in general, snow under a forest canopy is challenging to observe via optical methods because of forest canopy cover and the resulting low signal-to-noise ratio [13,79]. Additional availability of bands (e.g., shortwave infrared (SWIR)) would also enable use of the Normalized Difference Snow Index (NDSI) that might better the predictability of the model in forests [46,80] and further gives the ability to mimic MODIS-type explorations that utilize broader band availability [66,70].

## 1.7 Conclusions

We used predictors such as vegetation structure (using canopy height and NDVI), digital elevation models (DEM), and DEM-derived attributes to produce 3 m snow-covered area from PlanetScope imagery to evaluate improvements in two representative important river basins, Tuolumne and Gunnison (in USA), and at Engadin (Switzerland). Overall, we find that the inclusion of NDVI into the model increases the model transferability more significantly than DEM and DEM-derived attributes such as slope, aspect, and northness.

Our best model that used NDVI along with visible (red, green, and blue) and NIR bands captures the influence of vegetation on snow distribution. Specifically, the use of vegetation proxies (NDVI and canopy height) and terrain-derived measures was found to improve the accuracy of detecting snow in forested areas. The use of slope, aspect, and northness improves the ability of predicting snow in forest understories. The best model with an F-score of 0.89 (Gunnison) and 0.93 (Engadin) was found to be 4% and 2% better than when using canopy height and terrain derived measures at Gunnison, respectively. The NDVI-based model results in the best snow-identification performance in both forested and open areas compared to other models. Furthermore, adding only DEM and its derived attributes was also found to be transferable in test areas. The use of slope, aspect, and northness was found to overpredict snow in open areas. Even though optical methods are known to have shortcomings in observing snow in dense forest understories, our model's improvements, along with the detailed canopy-based evaluation metrics such as those presented here, can be used to improve model performance regarding various types of forest feature (i.e., gaps, canopy edges, and dense overgrowth). Climate change projections show hydrologic changes to many mountainous basins; the approaches used in our study could be beneficial in mitigation efforts regarding climate change uncertainties. Improving snow-covered area identification in forested areas could improve hydrologic modeling accuracy and help to estimate late-season snowpack distribution. Our model holds promise, as it can better predict snow in forested areas that is in sync with captured imagery.

## Author Contributions

Conceptualization, A.J., A.F.C. and N.C.; methodology, A.F.C. and A.J.; software, A.J. and A.F.C.; validation, A.J. and N.C.; formal analysis, A.J.; investigation, A.J.; resources, N.C.; data curation, A.J. and N.C.; writing—original draft preparation, A.J.; writing—review and editing, A.J., A.F.C., K.Y., A.T., J.H.R.L., D.S. and N.C.; visualization, A.J.; supervision, N.C.; project administration, N.C.; funding acquisition, N.C. All authors have contributed to manuscript review and editing. All authors have read and agreed to the published version of the manuscript.

## Funding

This research was funded by the Data Science Environments project award from the Gordon and Betty Moore Foundation (Award #2013-10-29) and the Alfred P. Sloan Foundation (Award #3835) to the University of Washington eScience Institute, NASA (Award #80NSSC21K1151), NSF EAR—1947875 and NSF OAC—2117834.

## Acknowledgments

We acknowledge the support from the University of Washington eScience Institute, UW Biology and UW Civil Engineering. We would also like to thank Kat Bormann from ASO, Inc. for providing data, and Giulia Mazzotti and Clare Webster for helpful discussions. Comments from three anonymous reviewers greatly improved the manuscript.

## Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## 1.1 APPENDIX A

Table A.1. ASO collection IDs for the sites

| Site | ASO ID[1] |
|---|---|
| Tuolumne, CA, USA | ASO_3M_SD_USCATE_20180528 |
| Gunnison River, CO, USA | ASO_3M_SD_USCOGE_20180524 |
| Engadine, Switzerland | EUCHDB20170517_SUPER |

[1] 3m gridded product

Table A.2. Performance metrics in Gunnison River Basin, CO, USA [1]

| Metrics | BASE+ NDVI | BASE | BASE EXT[2] | BASE + CHM | BASE + DEM | BASE + DEM (Elevation) + slope + aspect + northness | BASE + slope + aspect + northness |
|---|---|---|---|---|---|---|---|
| Accuracy | .89 (.01) | .88 (.01) | .88 (.01) | .88 (.01) | .88 (.01) | .89 (.01) | .78 (.02) |
| Precision | .94 (.02) | .88 (.88) | .89 (.02) | .87 (.04) | .87 (.02) | .89 (.03) | .64 (.02) |
| Recall | .94 (.03) | .84 (.04) | .87 (.02) | .86 (.03) | .86 (.03) | .87 (.02) | .90 (.03) |
| F-score | .89 (.02) | .85 (.05) | .88 (.02) | .86 (.04) | .88 (.02) | .88 (.03) | .74 (.04) |
| Balanced Accuracy | .78 (.03) | .75 (.02) | .73 (.02) | .74 (.02) | .73 (.02) | .74 (.02) | .65 (.02) |

[1] All the scenes in the study area (n=19); Scores represent mean values, and values in brackets are standard deviation.
[2] Scores are over 16 scenes as model was not able to predict snow over 3 scenes.

Table A.3. Performance metrics in Engadine, Switzerland [1]

| Metrics | BASE+ NDVI | BASE | BASE EXT[2] | BASE + CHM | BASE + DEM | BASE + DEM (Elevation) + | BASE + slope + aspect + |
|---|---|---|---|---|---|---|---|

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | slope + aspect + northness | northness |
| Accuracy | .92 (.01) | .87 (.02) | .90 (.005) | .90 (.02) | .90 (.02) | .83 (.03) | .88 (.02) |
| Precision | .93 (.0009) | .93 (.006) | .90 (.001) | .91 (.003) | .86 (.007) | .79 (.003) | .83 (.003) |
| Recall | .94 (.005) | .87 (.007) | .93 (.017) | .93 (.013) | .97 (.0001) | .93 (.001) | .98 (.0003) |
| F-score | .93 (.002) | .90 (.001) | .92 (.007) | .92 (.004) | .91 (.004) | .85 (.002) | .90 (0.001) |
| Balanced Accuracy | .91 (.02) | .86 (.03) | .88 (.02) | .89 (.03) | .88 (.04) | .81 (.05) | .87 (.04) |

[1] All the scenes in the study area (n=2); Scores represent mean values, and values in brackets are standard deviation.

## 1.2 APPENDIX B

Table of PlanetScope Scenes

Table B.1. 19 Planet Scene IDs (Dove Classic constellation) for Gunnison River Basin, CO, USA

| PlanetScope Scene ID |
|---|
| 20180524_172142_103d_3B |
| 20180524_172143_103d_3B |
| 20180524_172144_103d_3B |
| 20180524_172145_103d_3B |
| 20180524_172146_103d_3B |
| 20180524_172147_103d_3B |
| 20180524_172148_103d_3B |
| 20180524_172326_0f51_3B |

| |
|---|
| 20180524_172327_0f51_3B |
| 20180524_172329_0f51_3B |
| 20180524_172330_0f51_3B |
| 20180524_172331_0f51_3B |
| 20180524_172332_0f51_3B |
| 20180524_172333_0f51_3B |
| 20180524_172632_0f2d_3B |
| 20180524_172633_0f2d_3B |
| 20180524_172634_0f2d_3B |
| 20180524_172635_0f2d_3B |
| 20180524_172637_0f2d_3B |

Table B.2. 19 Planet Scene IDs (Dove Classic constellation) for Engadine, Switzerland

| PlanetScope Scene ID |
| --- |
| 20170516_092514_1028_3B |
| 20170516_092515_1028_3B |

# Chapter 2. DETECTING MONTANE FLOWERING PHENOLOGY WITH CUBESAT IMAGERY

Aji John [1,*], Justin Ong [2], Elli J. Theobald [1], Julian D. Olden [3], Amanda Tan [4] and Janneke Hille Ris Lambers [1,5]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA; [2]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA; [3]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, USA; [4]eScience Institute, University of Washington, Seattle, WA 98195, USA; [5]Plant Ecology, Institute of Integrative Biology, D-USYS, ETH Zürich, 8092 Zürich, Switzerland

**Citation:** John A, Ong J, Theobald EJ, Olden JD, Tan A, HilleRisLambers J. Detecting Montane Flowering Phenology with CubeSat Imagery. Remote Sensing. 2020; 12(18):2894. https://doi.org/10.3390/rs12182894

## 2.1 ABSTRACT

Shifts in wildflower phenology in response to climate change are well documented in the scientific literature. The majority of studies have revealed phenological shifts using in-situ observations, some aided by citizen science efforts (e.g., National Phenology Network). Such investigations have been instrumental in quantifying phenological shifts but are challenged by the fact that limited resources often make it difficult to gather observations over large spatial scales and long-time frames. However, recent advances in finer scale satellite imagery may provide new opportunities to detect changes in phenology. These approaches have documented plot level changes in vegetation characteristics and leafing phenology, but it remains unclear whether they can also detect flowering in natural environments. Here, we test whether fine-resolution imagery (<10 m) can detect flowering and whether combining multiple sources of imagery improves the detection process. Examining alpine wildflowers at Mt. Rainier National Park (MORA), we found that high-resolution Random Forest (RF) classification from 3-m resolution

PlanetScope (from Planet Labs) imagery was able to delineate the flowering season captured by ground-based phenological surveys with an accuracy of 70% (Cohen's kappa = 0.25). We then combined PlanetScope data with coarser resolution but higher quality imagery from Sentinel and Landsat satellites (10-m Sentinel and 30-m Landsat), resulting in higher accuracy predictions (accuracy = 77%, Cohen's kappa = 0.39). The model was also able to identify the timing of peak flowering in a particularly warm year (2015), despite being calibrated on normal climate years. Our results suggest PlanetScope imagery holds utility in global change ecology where temporal frequency is important. Additionally, we suggest that combining imagery may provide a new approach to cross-calibrate sensors to account for radiometric irregularity inherent in fine resolution PlanetScope imagery. The development of this approach for wildflower phenology predictions provides new possibilities to monitor climate change effects on flowering communities at broader spatiotemporal scales. In protected and tourist areas where the flowering season draws large numbers of visitors, such as Mt. Rainier National Park, the modeling framework presented here could be a useful tool to manage and prioritize park resources.

## 2.2 Introduction

Shifts in seasonal timing of biological events in plants, such as germination, flowering, and fruiting, in response to climate change have been observed across numerous species [6,7]. For example, many species demonstrate earlier onset of development and advances in other key life-history events as the climate warms [81]. Alpine wildflowers are considered good indicators of climatic change, as their phenology is highly sensitive to spring and summer temperatures and the timing of snowmelt [5].

Such shifts in the timing of flowering of Alpine wildflowers is concerning, as it could disrupt interactions between Alpine wildflowers and other members of the community, such as pollinators [8]. Understanding and anticipating the impacts of climate change on flowering is therefore paramount to help inform conservation efforts to ensure the preservation of wildflower meadows in the future [82].

Mounting evidence from field studies on wildflower phenology compiled by scientists and volunteer networks document shifts in the timing of flowering [83–85] that seem linked to climate warming at those locations [11,34,86]. However, whether it is possible to extrapolate information gained at smaller spatial scales to larger spatial scales remains uncertain [87]. Quite simply, monitoring flowering phenology across broader spatial and temporal domains is challenged by limited monetary and human resources. The ability to detect flowering phenology via remote sensing, if possible, therefore has the potential to aid in the quantification of fine-scale flowering phenology at large spatiotemporal scales. Hyperspectral and multispectral imagery has proven to be useful when analyzing vegetation type and leaf green-up [88] in crop phenological studies [89–91] but, to the best of our knowledge, has not been used to assess other plant phenological stages in natural systems.

Recent developments of small satellites, so-called CubeSats, have introduced new possibilities to monitor land cover at fine spatial and temporal scales. CubeSats have the advantage of high spatial resolution (in comparison to more traditional satellite imagery) but are limited to few spectral bands and usually have a small form-factor (dimensions in the order of 10 cm by 10 cm by 30 cm). The low costs associated with their deployment have resulted in several companies using CubeSat technology to provide high spatial and temporal coverage of the earth. One such provider is Planet Labs, Inc. [12], which uses

over 150+ (as of 2019) CubeSats to image the entire land surface of the Earth at a daily time interval and 3–5 m resolution. This imagery provides ecologists with the exciting opportunity to track a wide array of natural processes that occur at fine spatial scales. Radiometric quality remains an issue of this data [92]; however, this shortcoming may be improved by cross-calibration using highly calibrated coarser resolution imagery, like Landsat and Sentinel [13].

Multispectral bands have long been used to detect plant physiological properties, and CubeSats offer the possibility of doing so at finer spatial resolutions than other commonly used satellite imagery. Multispectral products are produced by examining the reflectance spectra that are captured in the corresponding electromagnetic spectrum and are often used in calculating band ratios [13]. For example, a commonly used band ratio (or radiometric index) that is used to measure vegetation state is NDVI (Normalized Difference Vegetation Index—ratio of red and near infrared bands), which has been found successful when applied to crop type estimation, e.g., in maize, soybean, and wheat, because of the inherent uniformity in the vegetation [93]. In addition, it has been shown that yellow flowers can decrease NDVI values in alpine meadows [94], and a cut-off of 0.4 NDVI has been reported as indicative of full flowering in a study involving sunflower crops [95].

For this paper, we determined whether multi-spectral imagery obtained from Planet Labs, Inc., could help to quantify peak flowering in sub-alpine meadows. We did so by combining satellite imagery with an existing long-term on-the-ground datasets. The long-term datasets identify the growing season phenology of wildflower meadows during a period of five years along an elevation gradient at Mt. Rainier National Park in Washington, USA. We evaluated whether phenological stages are observable via the

spectral bands of PlanetScope (a 3-m, 4-band multispectral image product from Planet Labs) by using reflectance in the red and near infrared (NIR) bands. Additionally, we evaluated whether PlanetScope-derived NDVI can be used to delineate flowering. We also combined finer resolution imagery from Planet with coarser resolution imagery from Sentinel (10 m) and Landsat (30 m) to assess whether more accurate phenological detection is possible when combining high these two types of imagery. Our objective involved the two following questions. First, can peak flowering be detected via either PlanetScope NIR and red bands or a normalized measure, like NDVI? Second, is detectability of flowering improved when fine-resolution PlanetScope is combined with coarser resolution imagery (with higher quality images), like Landsat and Sentinel?

## 2.3 Materials and Methods

### 2.3.1 Study site

Study sites are located in Mt. Rainier National Park (46.8529° N, 121.7604° W; summit elevation 4392 m), part of a stratovolcanic mountain range in the Cascade Range of Washington State, USA. The regional climate is maritime with dry summers and wet winters. The vegetation is dominated by coniferous forests at lower elevations (<1450 m), and sub-alpine wildflower meadows at mid-elevations (1450–1900 m). We restricted our analyses to meadows spanning an elevational gradient, where phenology has been quantified in a long-term study from 2010–2015 (Theobald et al., 2017). Spatial polygons encompassing meadows were manually drawn to include all sampling locations in this study. These meadows are adjacent to a trail where additional long-term monitoring of flowering phenology is ongoing (2013–present) through a citizen science project called

MeadoWatch (http://www.meadowatch.org/), providing additional information on flowering phenology after 2015. The selected meadow sites were on the south slope of Mt. Rainier, along Mazama Ridge, approximately 100 m apart (Table 2-1 and Figure 2-1).

Table 2-1. Elevation and area of meadow sites.

| Meadow | Elevation (m) | Area ($m^2$) |
|---|---|---|
| 1 | 1487 | 877 |
| 2 | 1595 | 1328 |
| 3 | 1678 | 4694 |
| 4 | 1779 | 1357 |
| 5 | 1894 | 1136 |

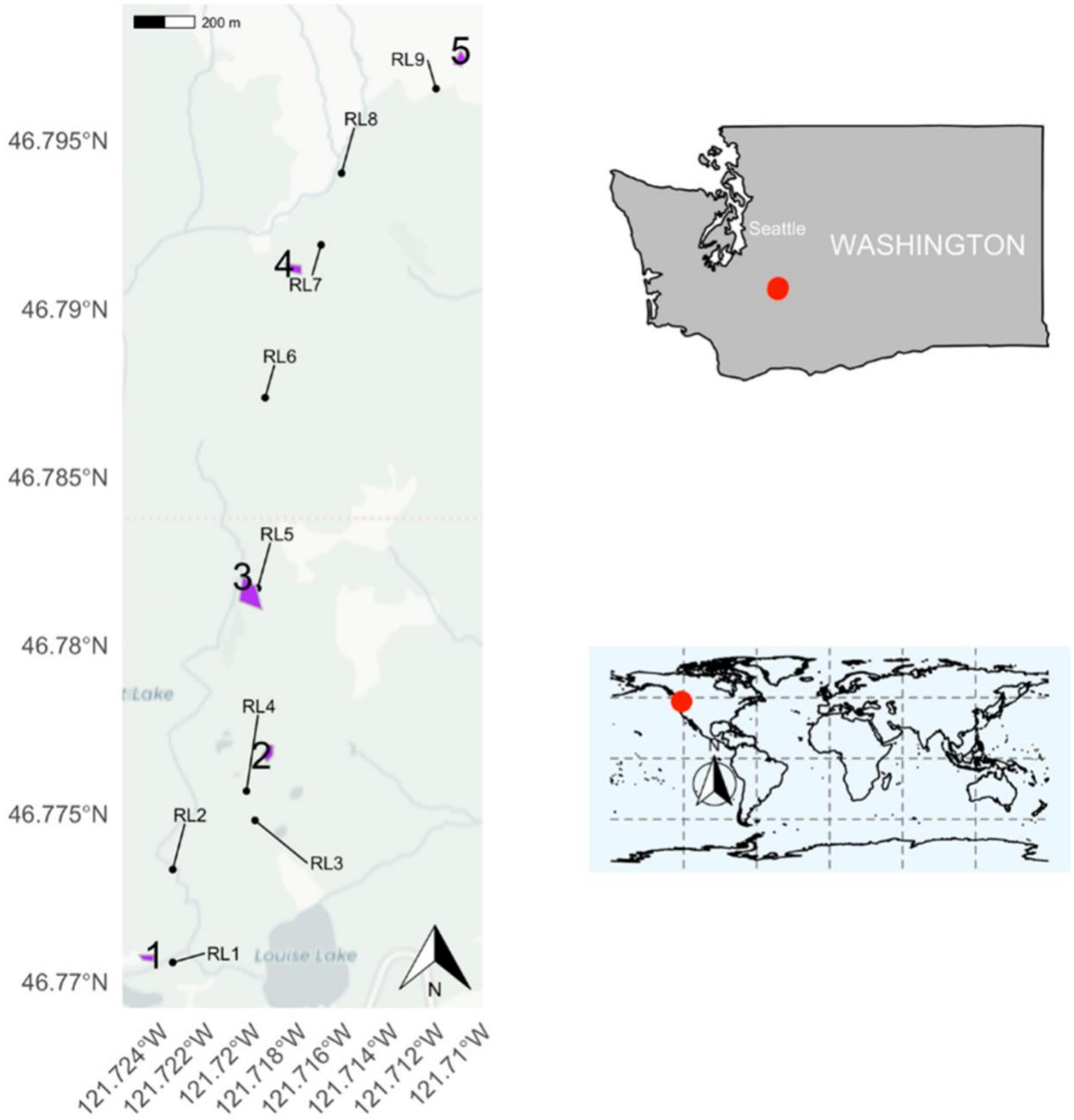


Figure 2-1. Meadow study sites at Mt. Rainier National Park; five sub-alpine meadows colored in purple are approximately 100 m apart in elevation and cover an elevational gradient from 1490 m to 1900 m. Polygons are the sites from the study Theobald et al. (2017), and the dots annotated with RL1-9 are plots from MeadoWatch.

### 2.3.2 Remote Sensing Data

Remote-sensed imagery were collected from Landsat, Sentinel, and Planet Labs. PlanetScope imagery from Planet Labs became available only after 2016, so we extracted

two years of PlanetScope imagery (2017–2018), 6 years of imagery from Landsat (2013–2018), and 2 years of imagery from Sentinel, which was initiated in 2015 (2017–2018). We excluded data from 2016 because of quality issues with Sentinel imagery (see Appendix D for imagery acquisition and resolution). Thousands (n = 2893) of images of meadows were analyzed over 5 years of study across providers Landsat 8, Sentinel-2, and Planet. For all three providers, we extracted summary reflectance values across red, green, blue, and NIR bands for each meadow area polygon (see Appendix C for band wavelengths). Additionally, we calculated summary reflectance in short-wave infrared (SWIR) and green chromatic coordinate ($g_{cc}$) for Sentinel imagery for each meadow. We briefly describe the three providers below.

The Landsat 8 OLI (Operational Land Imager) platform offers 30-m data with high-quality spectral calibration. The Landsat-8 satellite is one of a series that were launched by the National Aeronautics and Space Administration (NASA)/U.S. Geological Survey (USGS). Landsat Collection 1 Level-1 data products are used in the Landsat analysis (here onwards referred as L8).

Sentinel-2 is owned by the European Space Agency (ESA) and offers 10-m data for red, green, blue, and NIR bands. Officially called the Copernicus Sentinel-2 mission, Sentinel-2 consists of two polar-orbiting satellites in the same sun-synchronous orbit, phased at 180 degrees to each other (Sentinel-2, 2019). Sentinel-2 Level-1B products are used in the Sentinel workflow (here onwards referred as S2-1B).

Planet Labs PlanetScope constellation consists of 120+ CubeSats which orbit in two near-polar, sun-synchronous orbits of ~8° and ~98° inclination angle at an altitude of roughly 475 km. The CubeSats acquire both visible (RGB) and NIR data with 12-bit radiometric resolution that images the entire land surface on Earth daily. Specifically, we

used PlanetScope item PSScene4Band (type analytic_sr) in the Planet workflow (here onwards referred as PS).

### 2.3.3 Training and validation data: Peak flowering from on-the ground observations

To associate the satellite images to peak flowering, we developed a method to determine peak flowering window by elevation: first, we deduced flowering windows from field-based MeadoWatch observations (collected on the trail adjacent to meadow polygon) observations for 2017 and 2018 when PS, L8, and S2-1B imagery was available. Then for validation, we determined flowering time for 2013–2015 from observations from MeadoWatch and Theobald et al. (2017) when only L8 imagery was available. Example kernel density curves of observed flowering of top 10 recorded species for 6 years from MeadoWatch (2013–2018) are presented in Figure 2; 2015 showed a considerable shift in peak flowering day due to unprecedented warm temperatures (in contiguous United States average temperature was 2.4 °F above the 20th century average). See Appendix D for imagery and study period details.

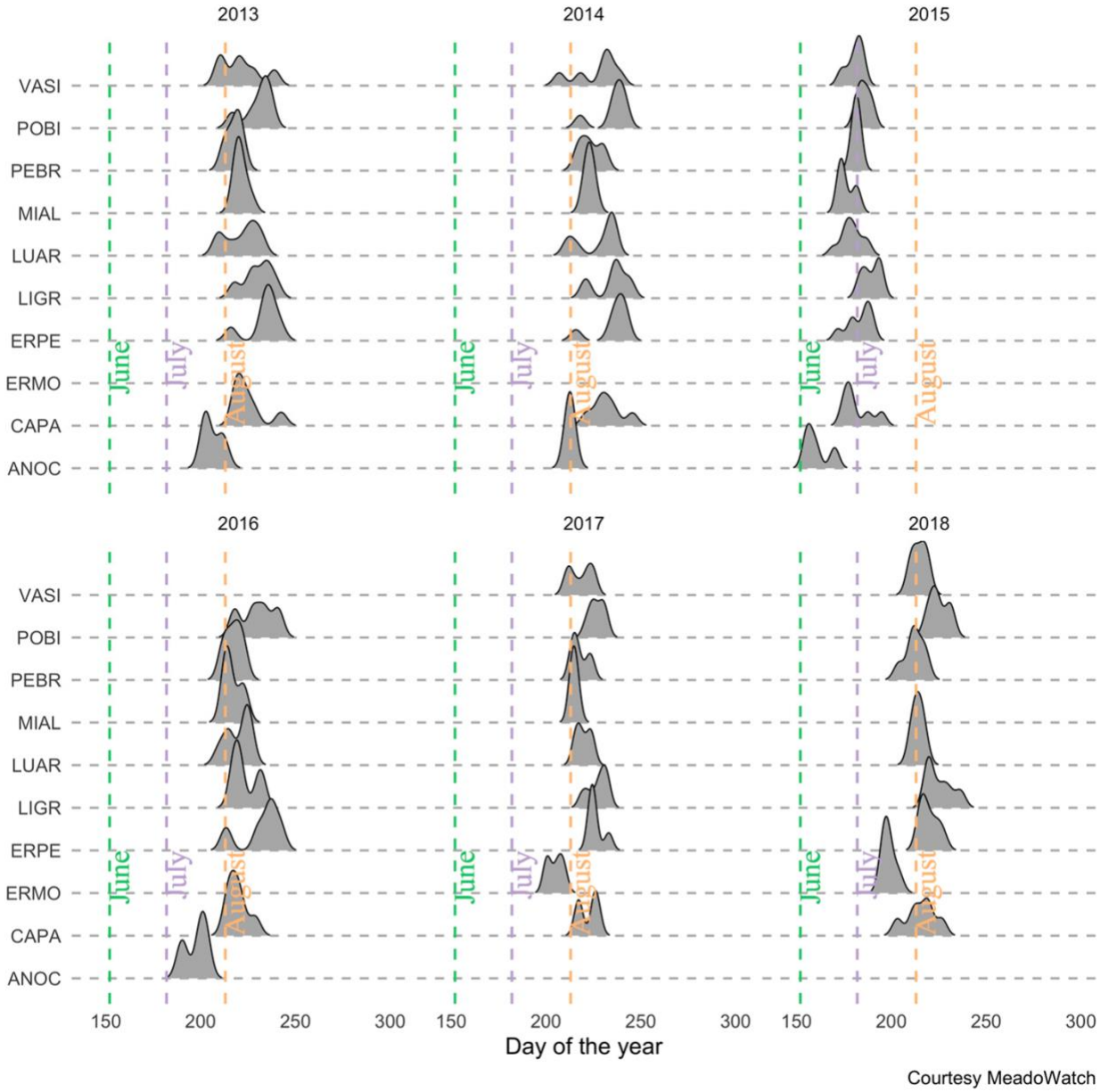


Figure 2-2. In-situ observed flowering of 10 of the most abundant meadow species across a 6-year period (2016 shown for completeness). The dotted lines indicate start of the month. Species include Valeriana sitchensis(VASI), Polygonum bistortoides (POBI), Pedicularis bracteosa (PEBR), Microceris alpestris (MIAL), Lupinus arcticus (LUAR), Ligusticum grayi (LIGR), Erigeron peregrinus (ERPE), Erythronium montanum (ERMO), Castilleja parviflora (CAPA), Anemone occidentalis (ANOC).

Flowering windows were estimated using flowering observations by the MeadoWatch program and those reported in Theobald et al. (2017). In both the cases, observer(s) recorded the species, the date (day of the year) and flowering status ('yes' or

'no') in multiple plots (plots measured 1 m × 1 m in Theobald et al. 2017 and were estimated at ~2 m × 1 m in MeadoWatch). The main difference between these datasets is that multiple 1 m × 1 m plots were sampled within one of five meadows sites by Theobald et al. (2017), whereas sampling occurred at nine single MeadoWatch plots, along a prominent hiking trail. MeadoWatch plots were close to the five Theobald sites and spanned the same elevation gradient. Plots within the Theobald et al. (2017) sites were combined to delineate an area (purple polygons in Figure 2-1) to fetch the satellite imagery. For the purposes of this study, the MeadoWatch plots were grouped by elevation using the shortest distance to nearby Theobald site (Figure 2-1). Once grouped by elevation, the observations were then separated by year to calculate mean and standard deviation (SD) of peak flowering day (Figure 3A). The days in one ±SD of the mean was considered to be peak flowering window for that year and elevation (Figure 3B). Our determination of flowering windows assumes that all 10 focal species are found in all plots in all years. This is not a valid assumption but making it does not bias our conclusions as we are making inference at the community-level. In other words, we are determining peak flower as the date at which there is the highest probability of seeing any species in flower and are using the 10 most prominent and abundant species to make this assessment.

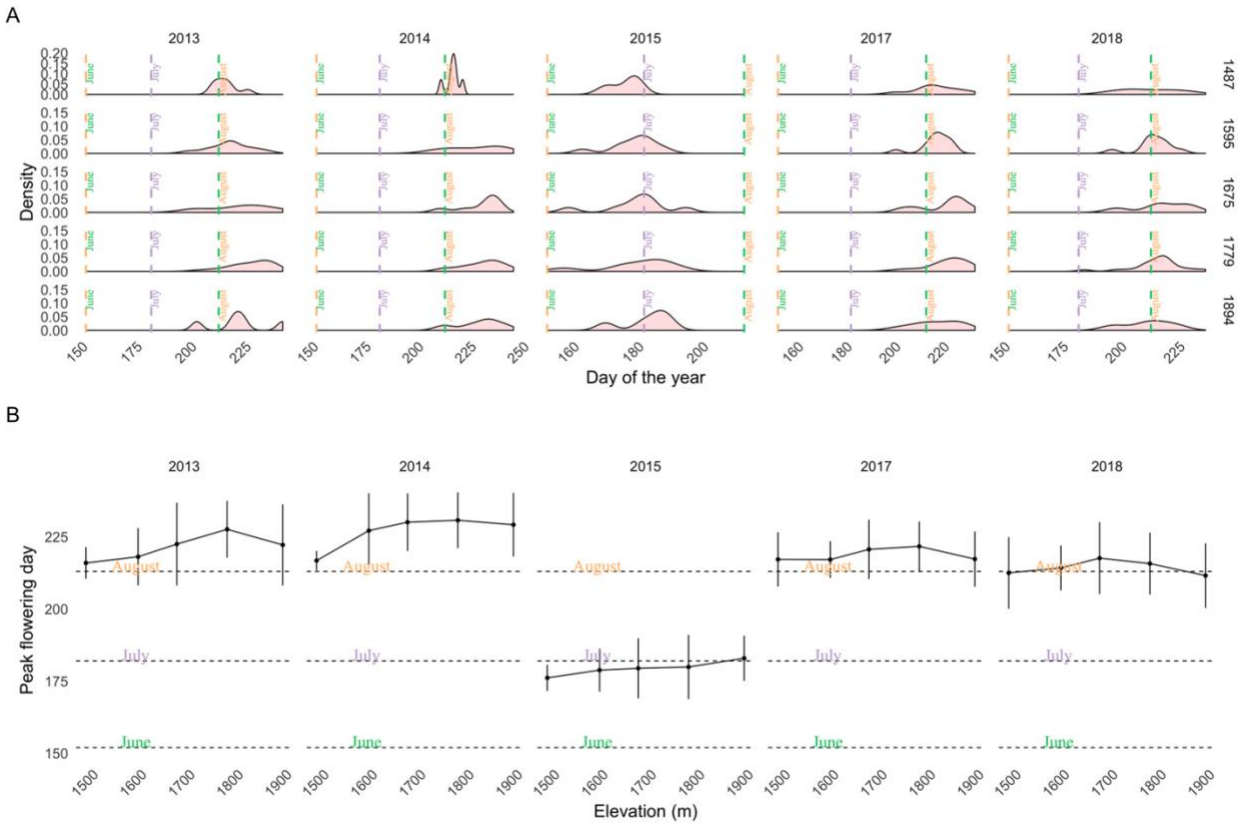


Figure 2-3. Peak flowering windows using MeadoWatch observations. (A) Kernel density of flowering observations by elevation (in meters) and year. (B) Estimated flowering windows calculated using mean flowering day and ±1 SD; error bars indicate length of the flowering window with the dot signifying mean peak flowering day. Note the shift in flowering phenology in 2015, a historically warm year.

### 2.3.4 Satellite Data Processing

All processing of satellite imagery data was performed using SWEEP [96], a workflow management platform for distributed execution in cloud infrastructures. SWEEP workflows for acquiring and analyzing imagery were developed for PS, L8, and S2-1B and are briefly outlined in the Appendix B.

### 2.3.5 Analysis

To assess the relevance of the NIR and red spectral bands to flowering phenology, we employed Principal Component Analysis (PCA) to summarize dominant patterns of

variation in the spectral band data and relate it back to flowering phenology. PCA, a multivariate ordination approach, is a preferable way to account for correlations among spectral bands [97,98]. We hypothesized that a shift in reflectance signature in NIR and red bands would be observed in known flowering months and would differ from the signature when meadows green-up (early summer, after snowmelt) or when the meadows are covered in snow (early summer and spring). Therefore, for each satellite image we extracted the minimum, maximum and average pixel value by spectral band for each meadow. We created a spectral matrix where rows (or objects) corresponded to satellite images (from PS), and columns (descriptors) included summary measures (minimum, maximum, and mean) of all the bands and NDVI (calculated using the corresponding red and NIR band). PCA was conducted using PS imagery only because a fine resolution of 3 m has more spectral representation for an average meadow (typically 30 m by 30 m) and is more likely to capture the spectral variability in a meadow, i.e., 9-pixel values for an average meadow. We used PCA in part because it could potentially help identify the reflectance bands that are related to months when flowering is typically observed. Hence, we characterized the relationship between the flowering meadows and spectral reflectance to understand which band captures flowering.

We used a Random Forest (RF) classifier [99] to predict the occurrence of flowering (binary yes or no) as a function of the spectral variables for years 2013–2015. Random Forest (RF) classifiers are a model-averaging or ensemble-based approach in which multiple classification tree models are built using random subsets of the data and predictor variables. This approach uses a recursive partitioning algorithm to repeatedly partition the data set according to the predictor variables into a nested series of mutually exclusive groups, each as homogeneous as possible with respect to the response variable.

It requires fewer parameters to be fit, has been shown to be less biased than other machine learning methods, and is known to be effective in classifying vegetation in remote sensing applications [100,101]. We used Cohen's kappa statistic to quantify model performance (comparing expected vs. observed error), and Gini importance to obtain the predictive contributions of the spectral features in the Random Forest (RF) classifier. We picked the threshold for determining the predicted probability of flowering by evaluating the receiver operating curve (ROC) and the distribution of true positive rate (sensitivity) and true negative rate (specificity) with respect to threshold.

We compared how finer resolution imagery compared to the combined resolution imagery by comparing their Random Forest (RF) performance metrics. We qualitatively assessed peak flowering using L8 imagery from the years 2013–2015 with two surveys; in-situ observed flowering survey by Theobald et al., 2017 and MeadoWatch.

All the statistical analyses were performed in R (R Development Core Team 2008).

## 2.4 RESULTS

### 2.4.1 Importance of Spectral Bands in Flowering

The PS data showed that reflectance in the NIR band is comparatively higher in July and August, with the other bands showing a high degree of correlation (Figure 4C,D). There is a steady decrease in reflectance from June to July during snowmelt, and then a rapid decrease in reflectance during July and August, aligned with the timing of green-up and flowering.

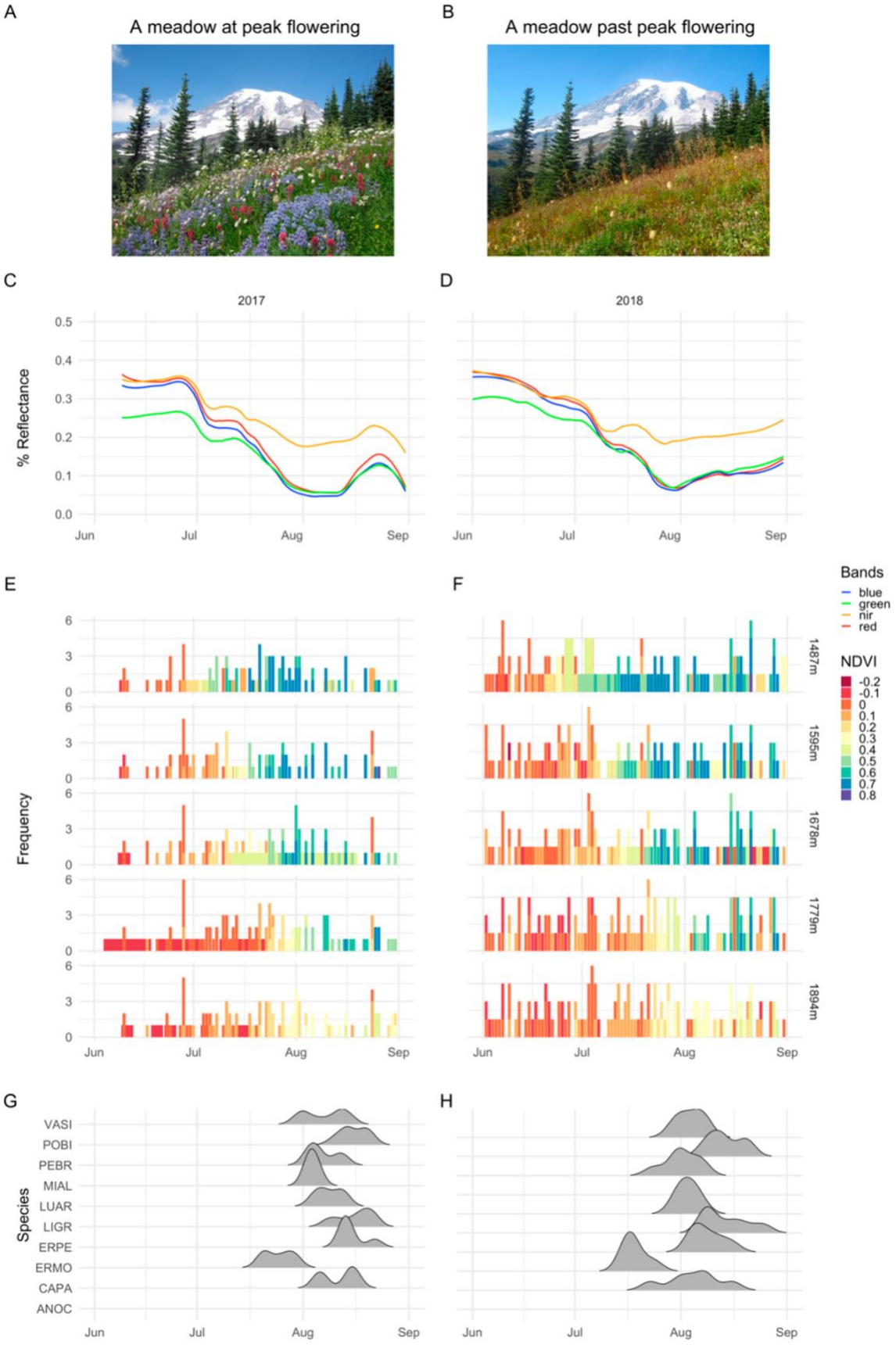
A
A meadow at peak flowering
B
A meadow past peak flowering
C
2017
D
2018
% Reflectance
0.5
0.4
0.3
0.2
0.1
0.0
Jun
Jul
Aug
Sep
E
F
Bands
blue
green
nir
red
NDVI
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Frequency
6
3
0
1487m
1595m
1678m
1779m
1894m
G
H
Species
VASI
POBI
PEBR
MIAL
LUAR
LIGR
ERPE
ERMO
CAPA
ANOC

Figure 2-4. (A,B) A typical flowering meadow at peak, and a meadow past the peak flowering. (C,D) Reflectance profile of all the meadows across two years (2017 and 2018) using PlanetScope item PSScene4Band (type analytic_sr) in the Planet workflow (PS). (E,F) Normalized Difference Vegetation Index (NDVI) profile of all the meadow sites by elevation for two years (2017, 2018) using PS. On the alternate y-axis is the elevation in meters, and y-axis shows the number of PS captures/meadow having the corresponding NDVI threshold. The NDVI thresholds are determined by taking the average of NDVI metric across the entire meadow. (G,H) Flowering observations for years 2017 and 2018 from MeadoWatch; showing dominant 10 flowering species.

Strong patterns are evident in red and NIR with respect to temporal patterns in observed flowering. Reflectance in green band is lower when compared to other bands until early July, and then close to zero reflectance in visible bands (i.e., red, green and blue; also means strong absorption). Increasing reflectance in the blue band reflects the snowmelt period until early June, followed by months characterized by decreasing reflectance that indicate greening-up of vegetation. Green-up causes a steady decrease in reflectance in all bands until an inflection point around late August, evident in 2017 and 2018 (Figure 4C,D). Higher absorbance in visible bands is observed in flowering months except for NIR, which is correlated with other bands until June, but then shows higher reflectance in flowering months (Figure 4).

The NDVI profile of the meadow sites by elevation for two years (2017, 2018) show alignment with flowering months (Figure 4G,H). NDVI is most elevated in late July and August (Figure 4D,E); coincidently, there is a ramp-up late June followed by a decline in greenness index after August (Figure A5). NDVI values between −0.1 to +0.1 signifies snow, which is evident in months until June (Figure A2 and Figure A3). Additionally, NDVI variability is lower at the highest elevation, where vegetation is sparse, than at the lower elevations; also evident is the lag in peak NDVI by elevation (Figure A4A,B). We found that average NDVI was linked to flowering months July and August (Figure A4C), also, NDVI drops after plateauing in late July–early August (PS imagery, Figure A6).

A large proportion of the variability in the spectral bands of PS was captured by the first two principal components (PCs) (Figure 5A,B). In peak flowering months (July and August), PC1 was positively correlated to mean NDVI, and negatively correlated to mean NIR reflectance values (Figure 5). Additionally, in peak flowering months, the reflectance in visible and NIR bands are negatively correlated to PC1. In non-flowering months, PC1 is positively correlated to reflectance in visible and NIR bands. Seasonal trends are evident, as snow starts declining in the meadows in May and continues with vegetative growth towards late June and flowering peaking around August. In typical flowering months July and August, NDVI values are strongly clustered with correspondingly higher positive mean values of NDVI in the range between 0.6 and 0.8 (Figure 5A, B). Spatial variability of flowering across elevations is also evident with lower elevations strongly tied to increased NDVI than higher elevations (Figure 5C).

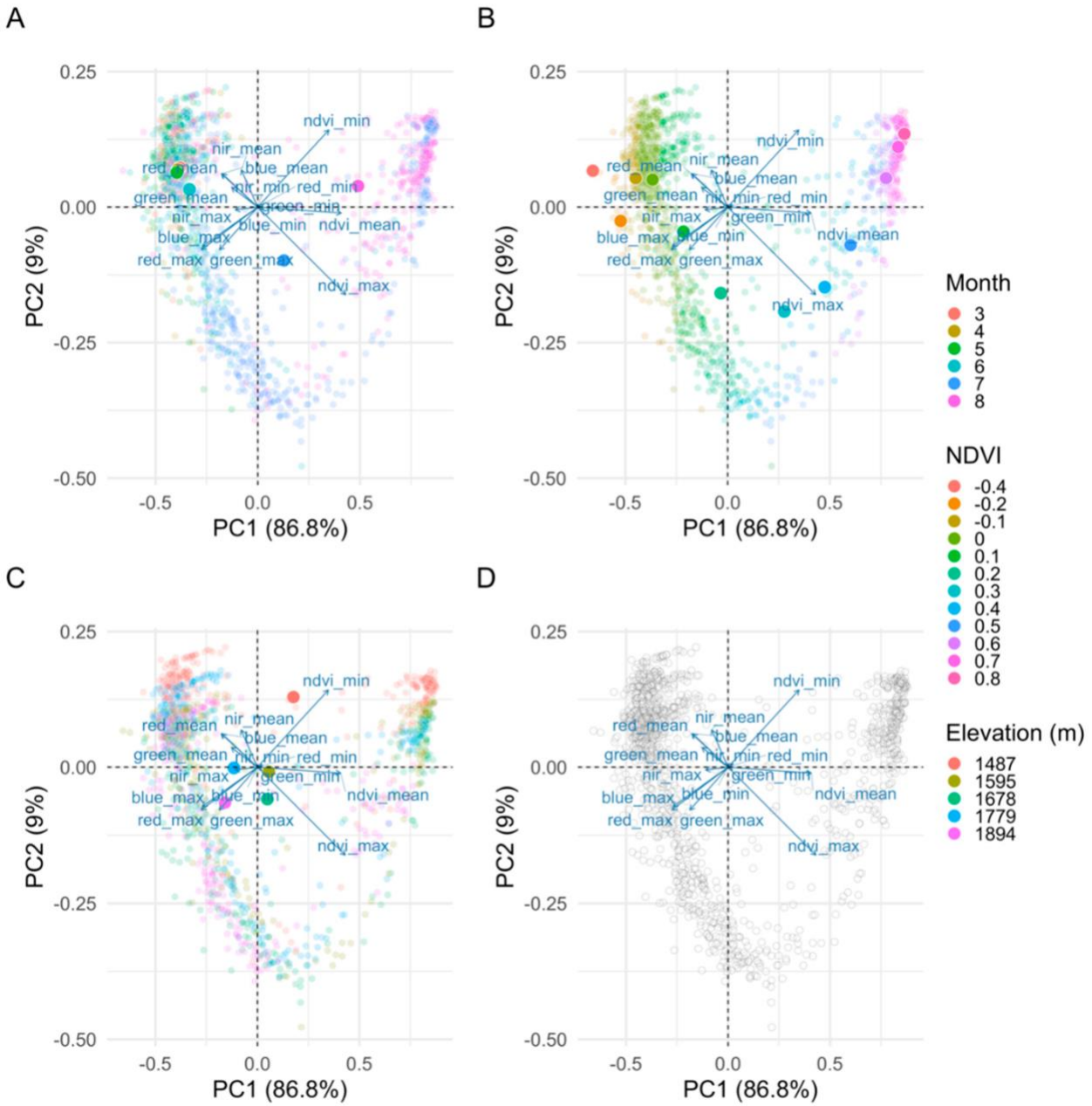


Figure 2-5. Biplot ordination from the principal component analysis (PCA) for 2 years (2017 and 2018) of spectral data from PS for the 5 meadow sites. (A–C) PC1 and PC2, overlain with the centroid (filled large dots) and average summary reflectance's captured in visible and near infrared (NIR) bands colored by month, NDVI, and elevation (blurry dots); (D) the ordination without any highlighting of the individual summary reflectance's. (A) Seasonal trends are evident; snowmelt in the meadows to flowering from upper left to upper right in the panel; blurry dots are colored by month. (B) Flowering months (late July and August) demonstrate higher positive mean values of NDVI in panel; blurry dots are colored by NDVI. (C) Flowering at sites showing the spread by elevation explained by increase in NDVI in the flowering months; blurry dots are colored by elevation of the meadow. (D) Strong correlation between visible bands, and NDVI metric that is orthogonal to NIR/visible bands.

### 2.4.2 Flowering Predictions Using Random Forest (RF)

Random Forest (RF) was trained to estimate the flowering window from PS and PS combined with coarser resolution satellite imagery. The results showed that the accuracy, i.e., correct classification rate when only PS was used was 70% (Kappa 0.25), when combined with coarser providers was 77% (Kappa 0.39). However, when only coarser imagery was used, accuracy was 72% (Kappa 0.37). PS explained the most variation at 55% and combined imagery without PS was the lowest with 29% (Table 2-2). We used a threshold of 0.25 to determine the predicted probability of flowering.

Table 2-2. Metrics when different types of imagery was used. Model combining the PS imagery with coarser providers yielded better results than PS only based model.

| Metrics | PS + L8 + S2-1B | L8 + S2-1B | PS |
|---|---|---|---|
| Accuracy (%) | 77 | 72 | 70 |
| Median CV[1] RMSE | .29 | .27 | .31 |
| Median CV[1] Variation (%) | 50 | 29 | 55 |
| Kappa | .39 | .37 | .25 |

[1] 99 cross-validations with .10 proportion withheld at each run

In our Random Forest (RF) variable importance analysis, NIR was less relevant than NDVI for identifying peak flowering. Mean values of reflectance in NDVI were most important in the RF model when only fine resolution imagery was used (Figure 6A), whereas, when using combined imagery, the blue band was most important (Figure 6B). However, NIR is integrated into NDVI as part of being a normalized measure with the red band, and NDVI does stand out in both the datasets, but NIR does not stand out as much as the red band.

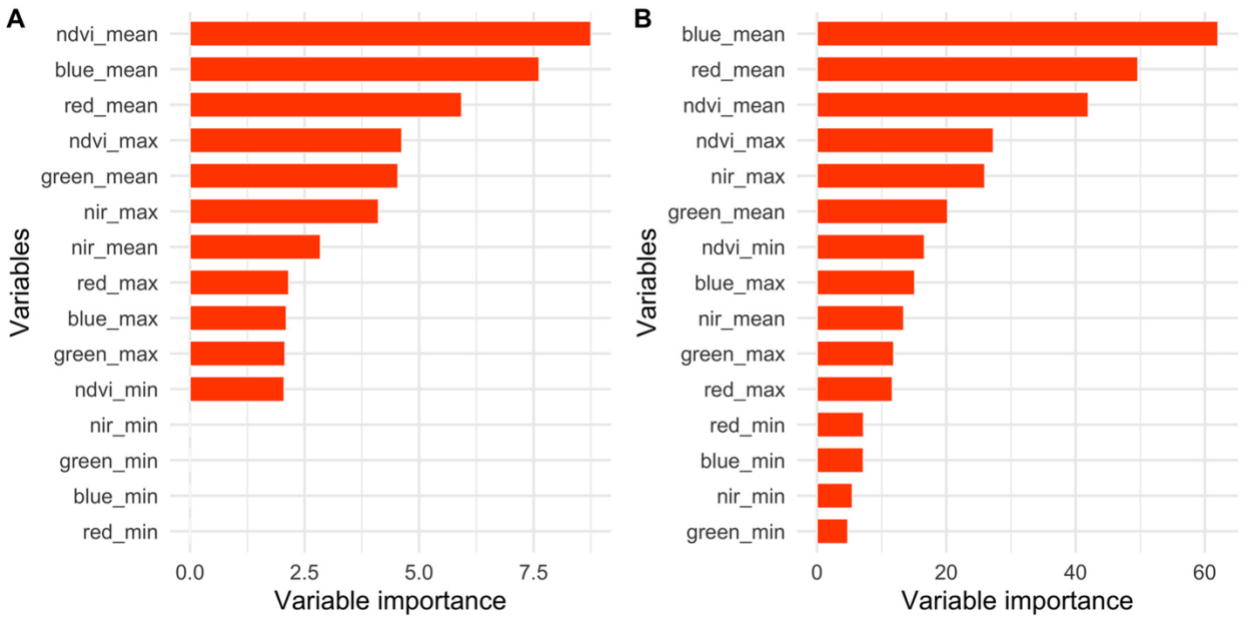


Figure 2-6. Relative importance of predictor spectral variables related to flowering when fine resolution imagery was used versus when fine-level imagery was combined with coarser resolution imagery. (A) Using PS only highlights NDVI as the top contributor. (B) PS, along with L8 and S2-1B, highlights the blue band as the top contributor.

Predicted peak phenology windows from spectral images aligned with observed phenology in some years than the others. The RF model captures the middle (or median) of the flowering window when compared to in-situ but tends to be longer, and the predicted window does not align in most cases when compared to MeadoWatch (Figure 7A,B). In comparison with both in-situ and MeadoWatch observations, there is over- and underestimation, which is evident in misaligned start of peak flowering window in certain years/elevations and overlapping intervals in some years/elevations. For 2015, an exceptionally warm year, both Random Forest (RF) models predict with less overlap but exhibit longer flowering windows that show delayed onset of flowering. Both the models overpredict the start of the flowering in average years (2013 and 2014) than in a warm year (2015). The predicted flowering windows aligned better when combined imagery was used than when only the finer imagery was used.

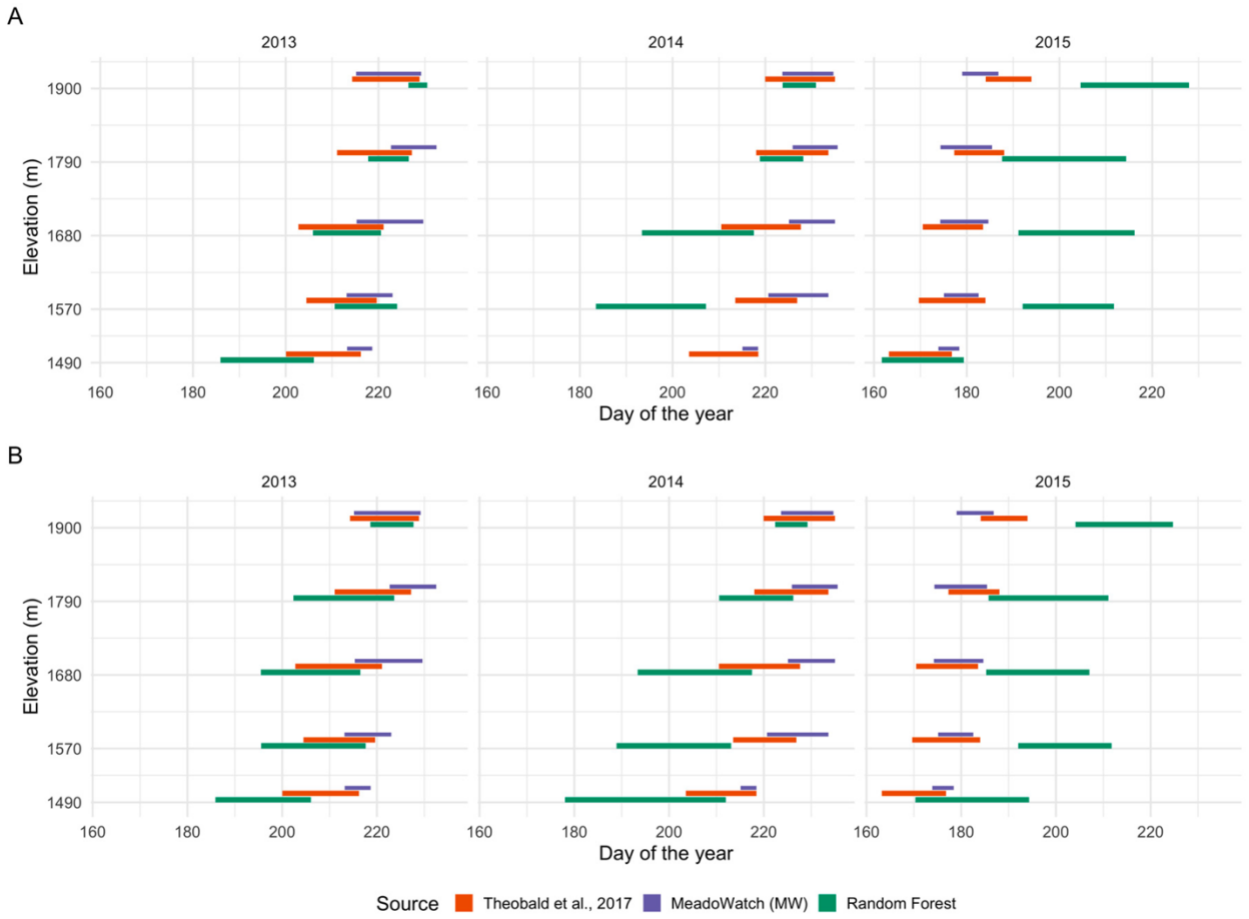


Figure 2-7. Qualitative comparison of observed window from the in-situ observations of (Theobald et al., 2017), MeadoWatch program, and Random Forest (RF) based flowering window. (A) Predicted and Observed peak flowering window when only 3-m (PS) resolution data was used for training. (B) Predicted and Observed peak flowering window when 3-m (PS) resolution, along with 10-m (S2-1B) and 30-m (L8) data, was used for training.

## 2.5 Discussion

Our study demonstrates that the timing of peak flowering in alpine meadows is detectable using fine resolution CubeSat imagery, and that the addition of coarser resolution imagery also substantively improves model accuracy. We also found that NDVI (a metric often used to quantify vegetative phenology) better predicted flowering phenology than did the NIR spectral band. The flowering window predicted from our model overlapped with the observed window in many site-year combinations but not in others, suggesting that accuracy is still an issue when using remote sensing imagery to detect flowering in

meadows. Finally, more years where fine-resolution imagery overlaps with on-the-ground phenology data would likely have allowed for improved model and model assessment, as we only had 2 years of PS imagery to relate to our MeadoWatch observations.

We found that higher reflectance was observed in NIR and red bands when flowering occurred and can potentially be combined with reflectance captured in other bands for improved flowering detection. Meadow sites exhibited visibly higher reflectance in the NIR when flowering as compared to meadow green-up, an expected result as flowers decrease absorption after the initial green-up [89,102]. Strong absorbance was also found in red and green bands during flowering, as well equal absorbance in the blue band when comparing to NIR. The blue band is typically associated with green-up time; it is a likely indicator of plant photosystems being active as allocation to vegetative growth is needed to support the energy required to produce flowers [103]. We found that NIR was correlated to flowering as is NDVI (which is a function of NIR). We also explored other indices, like green chromatic coordinate ($g_{cc}$) index, which is a proxy for greenness that has been found to be useful for detecting flowering other studies [104]. Our findings suggest that there is a decrease in greenness (i.e., $g_{cc}$) past peak green-up prior to flowering (Figure A5).

Use of visible and NIR bands are instrumental in phenological explorations but might be better when coupled with additional bands. We show that using visible bands (red, green, and blue) and NIR bands were sufficient in narrowing down the flowering window, but use of finer bands at the red-edge region might help further augment the flowering signal [105,106]. For example, use of SWIR has been found to be useful in accentuating signals of senescent vegetation. In addition, a recent announcement by Planet to provide

additional bands (between 5 to 8) on next generation CubeSats could improve the phenological assessment at finer resolutions.

NDVI was found to be a significant predictor of peak flowering phenology in alpine/subalpine meadows. NDVI normalizes NIR and red bands and has been shown to be associated with green-up [107]. Furthermore, flowering must be preceded by peak green-up, which is evident by plateauing of NDVI (Figure A4A,B), signifying saturation because of chlorophyll accumulation [108]. Although reflectance in NIR visually shows the strongest pattern during peak flowering (Figure 4C,D), we actually found that a normalized measure like NDVI (which incorporates NIR) better predicts flowering (Figure 6A,B). The NDVI-based metric is usually applied in crop classification studies [109], but we also found it useful for detecting flowering in alpine meadows—an important result as it indicates this metric may also be applicable in studies which are looking to differentiate other phenological stages.

Despite the fact that NDVI was a reliable metric in our study, we recognize that it also has a number of shortcomings. Remotely sensed NDVI has been acknowledged as differing from true NDVI because of atmospheric effects and varying soil brightness, both of which can differ from image to image [63]. Although soil blotches are not common in alpine meadows except at very high elevations, the largest challenge is associated with canopy background limitations [110]. For this reason, Enhanced Vegetation Index (EVI) is often used as an alternative to NDVI to address the soil and atmospheric limitations [90]. We therefore suggest that EVI and the other related indices should be explored for their utility in predicting phenological stages beyond vegetative ones, including the combination of several indices to improve measurement accuracy.

We show that combining several types of satellite imagery leads to enhanced predictions of flowering phenology. Specifically, by combining 3-m spectral imagery with 10-m and 30-m imagery, our model predictions showed better overlap with observed flowering windows when compared to just the finer resolution model. However, the combined model tended to overpredict the start of the flowering windows (except in the anomalously warm year—2015). Similar studies have shown that multi-resolution data analysis improves results: "fuzzier" low resolution data can provide "big picture" information, and, when combined with the lower-resolution data, finer details can be revealed [111].

Our results also suggest that coarser 30-m Landsat imagery can be useful to infer peak flowering. The model parameterized with combined imagery (including finer scale Planet imagery) was able to infer peak flowering when applied to Landsat in years where only coarser scale imagery was available. Additional improvements to model predictions are possible by refining atmospheric corrections used for Landsat and Sentinel; in this study, we uniformly applied corrections across each individual band in order to account for atmospheric absorption, such as the effect of haze (Appendix A). However, these effects are not likely to be uniform. Future work should examine the effect of band-specific atmospheric corrections on phenology prediction accuracy.

Despite having an amalgam of satellite imagery of the meadows, one fundamental question is that are we able to distinguish the flowering signal from the background, i.e., soil, rocks or green vegetative growth [112,113]. PS comes with only 4 bands and has overlapping bands that might lead to pixel quality issues [114,115] , and other coarser providers have constraints of resolution and frequency of captures. However, meadow wildflower phenology progresses seasonally in a predictable manner. In the springtime,

plants are covered with snow; when snow melts, they green-up, and then peak-flowering happens, and then snow falls again, covering all vegetation. This cycle culminates in about 4 months. We see this same 4-month cycle in the satellite signal through RGB composite and NDVI (Figure A6). Our method is bound by these constraints but having more on-the-ground observations in time and even better satellite resolution (e.g., <3 m from DigitalGlobe) might be a significant next step. One option of combining imagery of various resolutions as we showed here is a promising approach. We plan on exploring the possibilities of augmenting the model training with more on-the-ground data and/or using the sub-meter imagery to create the training data.

Our work highlights the challenges to detect flowering from satellite imagery in heterogeneous plant communities, such as those found in alpine meadows. Previous studies reporting promising results were mostly conducted in simplified systems (e.g., monoculture crops) where there is greater uniformity in landscape features and phenological trajectories [90,116,117]. By contrast, the effects of elevation, vegetation composition, and heterogenous snow (in the early season) make it more challenging to predict flowering from satellite imagery in alpine meadows. In other montane systems, a modified approach that takes into account species level reflectance profiles might improve the flowering signal in mixed flowering species, e.g., use of vegetation color specific reflectance profiles in differentiating between flowering and green-up [89,118].

## 2.6 Conclusions

In this study, we tested whether fine-resolution imagery (<10 m) can detect flowering and whether combining multiple sources of imagery improves the detection process. By

examining alpine wildflowers at Mt. Rainier National Park (MORA), we found that combined PlanetScope (from Planet Labs) data with coarser resolution but better-quality imagery from Sentinel and Landsat satellites (10-m Sentinel and 30-m Landsat), resulted in higher accuracy in delineating the flowering season captured by ground-based phenological surveys than a finer 3-m resolution PlanetScope imagery.

Our methodology holds potential in quantifying climate-driven shifts in alpine meadows in support of both scientific and management goals. Our approach to wildflower phenology predictions using finer scale satellite imagery serve economic interests as Mt. Rainier National Park receives over 2 million yearly visitors, the majority of whom come in the summertime to enjoy the idyllic wildflower vistas. Better predictions of peak flowering could help park facilities to prepare for increased visitation and provide timely visitor updates.

## Author Contributions

Conceptualization, A.J. and E.J.T.; methodology, A.J. and J.D.O.; software, A.J., J.D.O.; validation, A.J., and J.D.O.; formal analysis, A.J.; investigation, A.J.; resources, A.J.; data curation, A.J.; writing—original draft preparation, A.J.; writing—review and editing, A.J., J.O., E.J.T., J.D.O., A.T., and J.H.; visualization, A.J.; supervision, J.H. All authors have read and agreed to the published version of the manuscript.

## Funding

AWS credits were provided by eScience Institute at University of Washington (UW) and UW-IT, and Microsoft Azure Cloud credits were provided through the Microsoft AI for Earth Grant to Aji John.

## Acknowledgments

We acknowledge Planet's Ambassadors program support in providing the 3m PlanetScope imagery.

**Conflicts of Interest**

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## 2.7 APPENDIX A

Data refinement is spread out over reprojection and cropping; however, the bulk of the data processing is done after cropping. The data is first converted into Top-of-Atmosphere reflectance (TOA) and then to Bottom-of-Atmosphere reflectance (BOA) also known as surface reflectance (SR). This is because the data gathered comes as raw Digital Numbers (DN), and the DN values are simply scaled values measured by the sensors and have no meaningful value. TOA is converted to BOA because BOA takes into account atmospheric effects, such as cloud cover, aerosol gases, etc. For Landsat, the equation for DN to TOA conversion is

$$\rho\lambda = M\rho\ Qcal + A\rho\cos(\theta SZ) \quad (A1)$$

In Equation (A1), $\rho\lambda$ = TOA planetary reflectance, $M\rho$ = band-specific multiplicative rescaling factor, $A\rho$ = band-specific additive rescaling factor, Qcal = quantized and calibrated DN pixel value, and $\theta SZ$ = solar zenith angle for solar correction.

$M\rho$ is $2 \times 10^{-5}$, and $A\rho$ is $-0.1$, which are both provided in the metadata file. The solar elevation ($\theta SE = 90 - \theta SZ$) angle is provided in the metadata file as an average of the

entire tile, but Landsat provides a tool to get the θSZ for each pixel in a specific band. This is used instead of the approximation.

Once TOA reflectance is calculated, a scatter value is subtracted in order to get BOA reflectance. TOA is only reflecting from above the atmosphere. This scatter value is calculated by using a method called Frequency 50 Minus 0.008 (F50 0.8%) developed by GIS Ag Maps [119]. This is an image-based atmospheric correction model based on the Chavez Landsat TM histogram method [120]. This model only accounts for atmospheric scattering and assumes a constant haze value throughout the entire image. This constant haze value is determined by a relative scatter lookup table provided by GIS Ag Maps. This is by no means the most accurate way to derive surface reflectance as it is an image-based correction algorithm; however, it was used here because of its accuracy and simplicity.

## 2.8 APPENDIX B

Workflows for acquiring and analyzing Landsat, Sentinel, and Planet satellite images were written and executed on SWEEP [96], a scalable workflow management platform.

#### 2.8.1.1 Appendix B.1. Landsat

The workflow begins with setting the input boundaries of the meadow sites, which are subsequently run against the USGS Earthdata Explorer API (EE API), along with the date ranges to fetch the available scenes. Scenes are available on EE API and AWS S3 (a cloud data storage product) as part of its open data program. EE API provides functionality to search and download Landsat imagery for free. The workflow was designed from the ground-up and is made available to ensure reproducibility. We chose to use both of these services to overcome an EE API limit on parallel downloads. Once a set of scenes is returned by EE API, the list of scenes was checked against AWS S3 for

availability. Once matched scenes are found, they are downloaded from S3 and then re-projected (projection WGS 84 is used for this workflow), cropped, and radio-metrically corrected using the Landsat TM histogram method by (Chavez Jr, 1988). Next, statistical measures (min, max, and mean) are calculated across all the bands in visible and NIR spectra at the meadow level. Finally, the results are written to a comma-separated values (CSV) file which includes the statistical values for each band, the name of the feature (meadow site), and the date.

#### 2.8.1.2 Appendix B.2. Sentinel

The workflow starts by setting a list of polygons referring to meadow sites. Next, the workflow gets the applicable AWS scenes using the *sentinelhub* Python package with a spatial and temporal filter. The results include the AWS S3 URL, as well as the date of each of the scenes, which are both extracted for all of the available scenes. Once the scenes are found for each meadow, retrieving and refining of the data for each scene is subsequently run in parallel. This includes fetching, re-projecting, cropping, and refining the image data. Re-projecting and cropping are similar to the Landsat implementation mentioned above.

The image data conversion for this workflow is less computationally demanding than that of the Landsat workflow because Sentinel-2 L1C data comes as Top-of-Atmosphere (TOA) reflectance. Thus, we only needed to subtract each band's scatter value to reach Bottom-of-Atmosphere (BOA) reflectance, also known as surface reflectance. The method used to find the scatter value is the same one used in the Landsat workflow, except that the Sentinel relative scatter table was used, as opposed to the Landsat relative scatter table. Both of the tables are provided by GIS Ag Maps [120] to find the scatter values for the bands.

#### 2.8.1.3 Appendix B.3. Planet

Here, all the interactions are made via the Planet API through the SWEEP tasks. The workflow starts by using the meadow polygons to search for images within the desired date range. The results are image identifiers that are passed to the next task, which issues an API call to activate each image scene. The next task sends a clipping request to the API, whereby the meadow is clipped from the larger scene. A download request is then sent for each image, which, when ready, is furnished via a time-sensitive link. Finally, the images are downloaded, and metrics calculated for each band. The output is written to a CSV file that has band metrics time-stamped for each meadow.

## 2.9 APPENDIX C

Table C1. Wavelengths (in nanometer, μm) of spectral bands by different satellite imagery providers.

| Band | Landsat 8 | Sentinel-2 | Planet |
|---|---|---|---|
| Blue | .45-.51 | .45-.52 | .45=.51 |
| Green | .53-.59 | .54-.57 | .50-.59 |
| Red | .64-.67 | .65-.68 | .59-.67 |
| NIR | .85-.88 | .78-.90 | .78-.86 |
| SWIR | 1.5-1.6 | .9-1.7 | N/A |

## 2.10 APPENDIX D

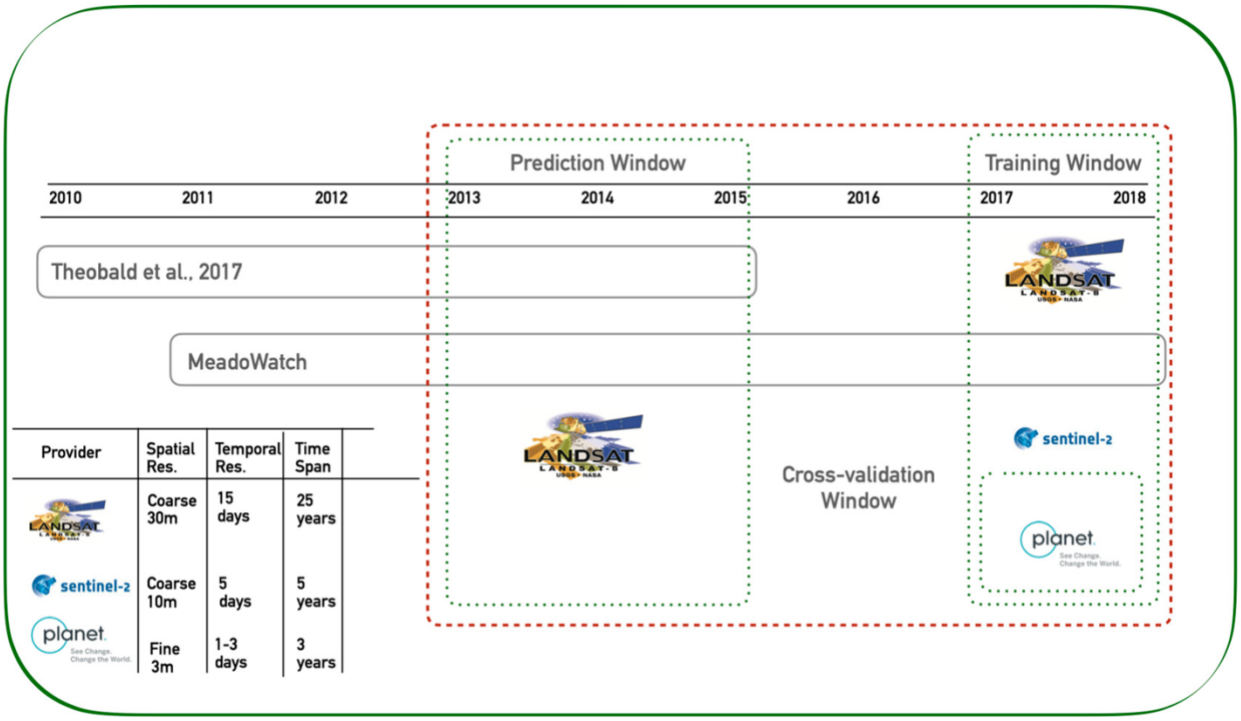


Figure A1. Time frame of fine resolution data versus coarse resolution data of the study.

## 2.11 APPENDIX E

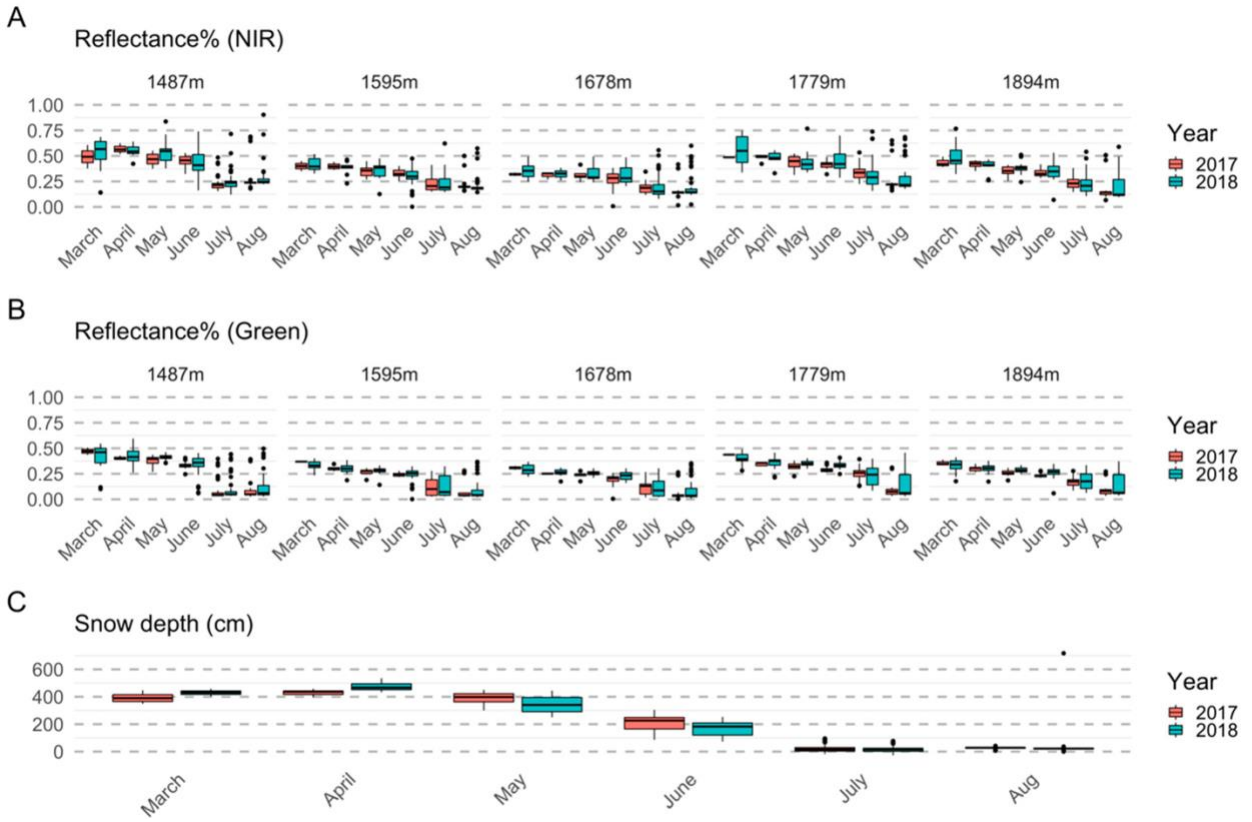


Figure A2. Planet reflectance comparison of NIR and green band showing snowmelt followed by green-up and flowering. (A) NIR reflectance is higher in snow months than in vegetative and flowering months, signified by positive correlation of reflectance's with snow-depth. (B) Green reflectance shows lower reflectance (i.e., higher absorption) after the snow-melt, signifying green-up. (C) Snow-depth plot from the metrological records highlighting snow-free months after June.

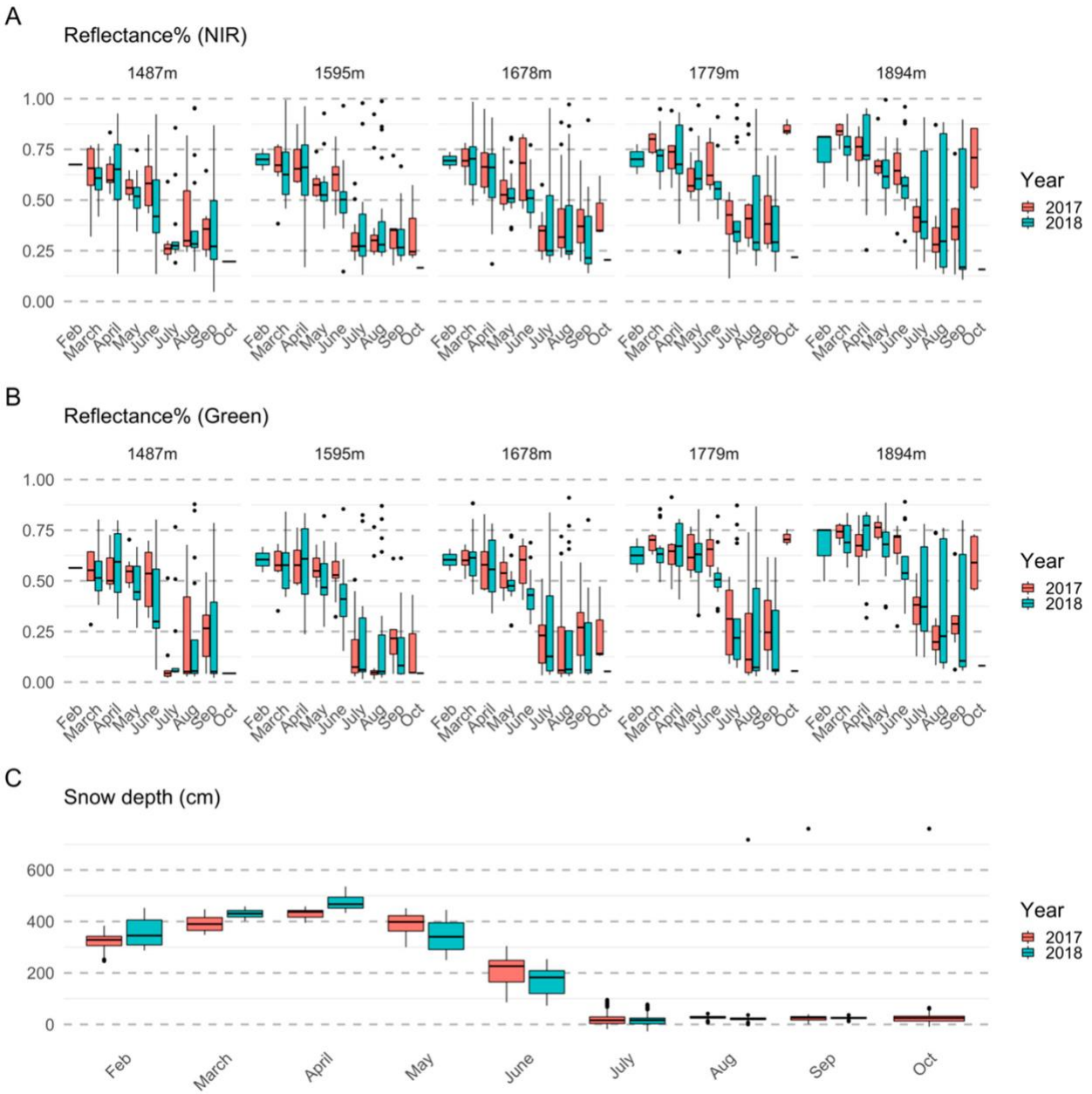


Figure A3. Combined Sentinel and Landsat reflectance comparison of NIR and green band showing snowmelt followed by green-up and flowering. (A) NIR reflectance is higher in snow months than in vegetative and flowering months, signified by positive correlation of reflectance's with snow-depth; however, note the upward trend of NIR in late September and October, signifying onset of snow or drying vegetation. (B) Green reflectance shows lower reflectance (i.e., higher absorption) after the snow-melt, signifying green-up. (C) Snow-depth plot from the metrological records highlighting snow-free months after June.

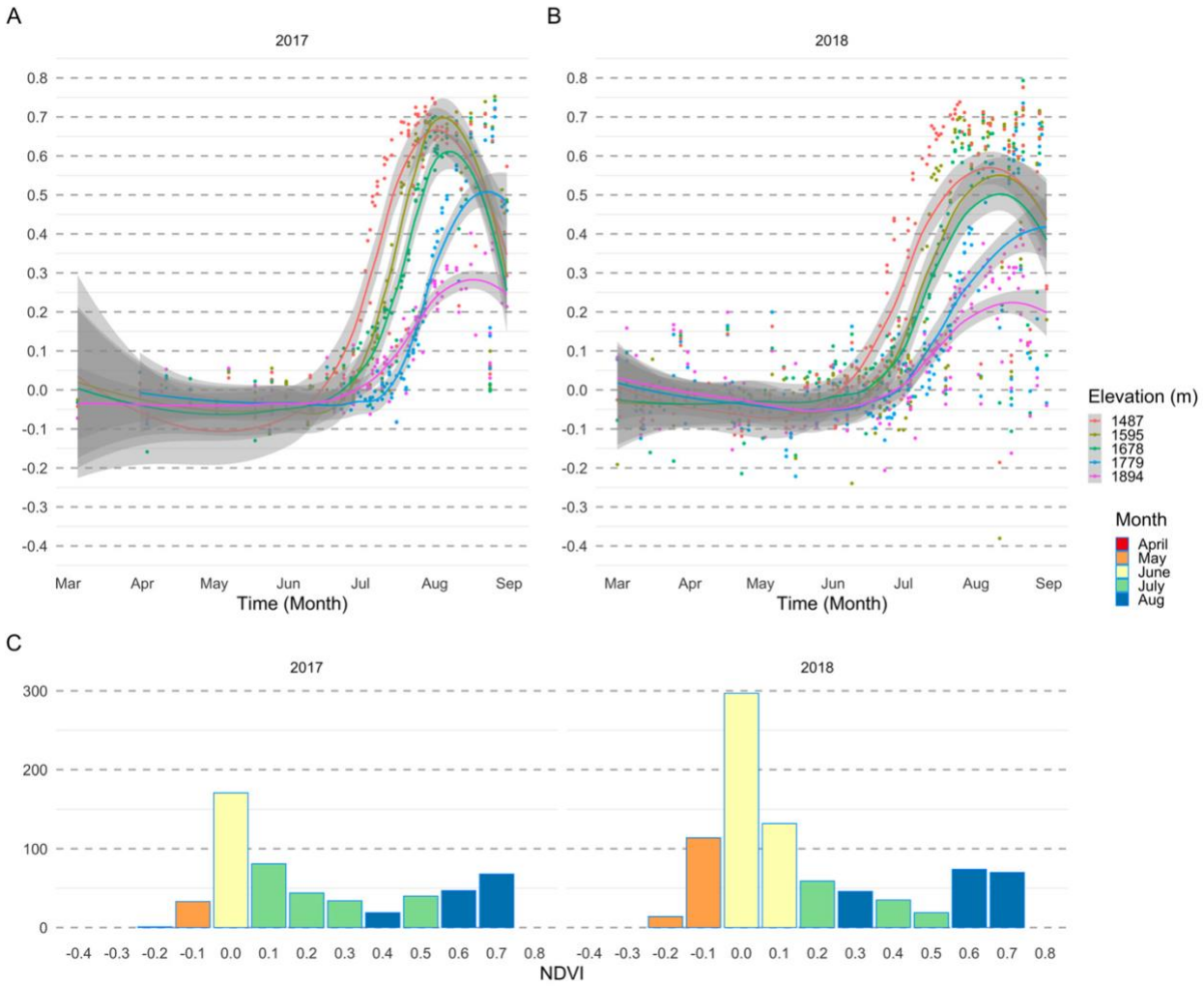


Figure A4. NDVI curve (loess smoothing) using Planet reflectance for years 2017–2018, showing the peak green-up around late July and also showing the phase lag by elevation. (A) 2017 NDVI values going from −0.2 to 0.8 capturing snow on the ground to flowering; also, notice the lag in peak NDVI confirming the snow-melt lag that is driven by elevation. (B) Similar patterns as 2017, 2018 NDVI values going from −0.2 to 0.8, capturing snow on the ground to flowering; also, notice the lag in peak NDVI confirming the snow-melt lag that is driven by elevation. (C) Histogram of NDVI values binned by year with the average month that it was observed in shaded.

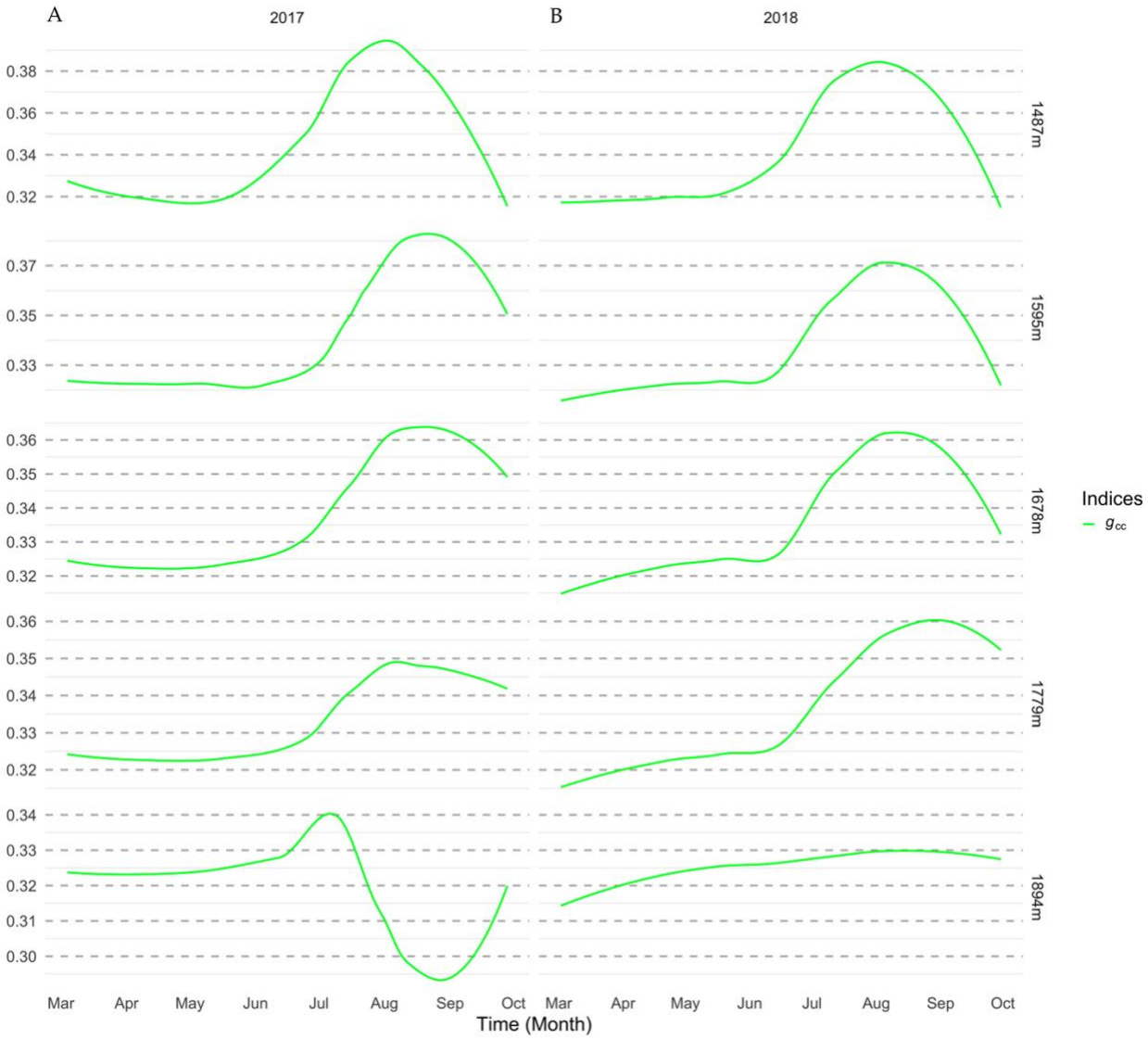


Figure A5. Green Chromatic Coordinate ($g_{cc}$) that is a proxy of greenness calculated using Sentinel reflectance as data existed after August for years 2017–2018, showing the decline in green-up around late August but moderated by elevation. (A) 2017, capturing the ramp-up after late June right after snow-melt and decline after August; highest elevation sites have very short green-up. (B) Similar patterns as 2017, 2018 shows ramp-up in greenness around late June and drop in greenness after August.

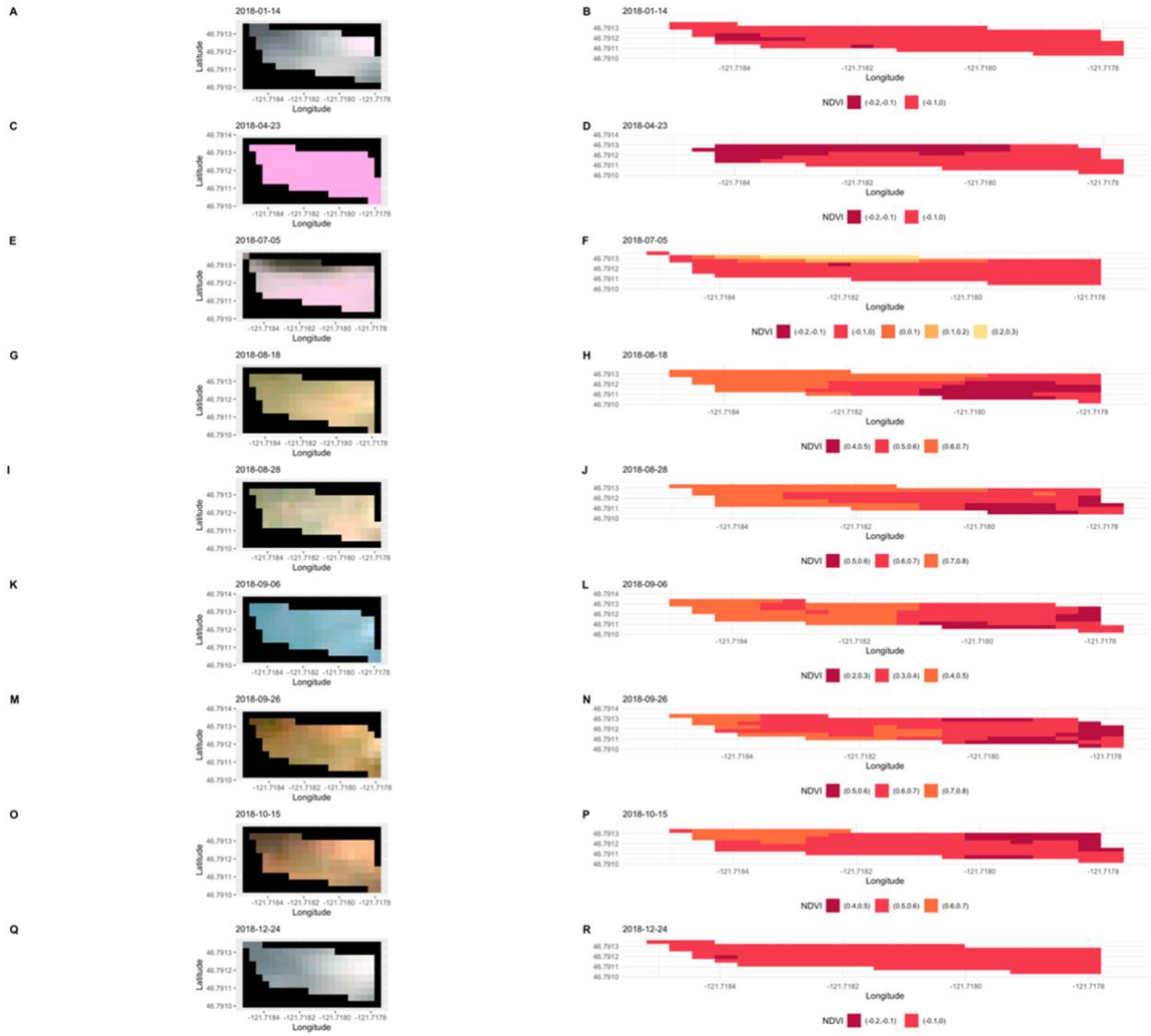


Figure A6. True color composite (red, green, and blue band combined) and corresponding NDVI of Meadow 4 (1779 m) for the year 2018. Column 1 for the panel gives a sense of how the meadow is changing during the whole year; starting from covered with snow (A,C,E), then meadow green-up and flowering (G,I,K), to senescence (M,O), and back to covered with snow (Q). Column 2 show the change in seasonality that is captured by NDVI; it shows in months till July NDVI remains below 0.2–0.3 (B,D,F), then a sharp increase in NDVI (H,J,L), and finally back to low levels of NDVI confirming snow (N,P,R).

# Chapter 3. USING PHOTOGRAPHS AND DEEP NEURAL NETWORKS TO UNDERSTAND FLOWERING PHENOLOGY AND DIVERSITY IN MOUNTAIN MEADOWS

Aji John [1,*], Elli J. Theobald [1], Nicoleta Cristea [2], Amanda Tan [3] and Janneke Hille Ris Lambers [1,4]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA; [2]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA; [3]eScience Institute, University of Washington, Seattle, WA 98195, USA; [4]Plant Ecology, Institute of Integrative Biology, D-USYS, ETH Zürich, 8092 Zürich, Switzerland

## 3.1 ABSTRACT

Mountain meadows are an essential part of the alpine-subalpine ecosystem; they provide ecosystem services like pollination and nutrient recycling and are home to diverse plant communities. Changes in climate affect the meadow ecology on multiple levels, altering the growing season dynamics and water availability. Tracking the effects of climate change on meadow diversity is critical to conservation efforts. Here, we explore how to combine crowd sourced camera images with machine learning to quantify flowering species richness across a range of elevations in alpine meadows located in Mt Rainier National Park, Washington, USA. We employed three machine learning techniques (Mask R-CNN, RetinaNet and YOLO) to detect wildflower species in images taken during two flowering seasons. We demonstrate that deep learning techniques can detect multiple species, providing information on flowering richness in photographed meadows. The results indicate higher richness just above the tree line for most of the species, which is comparable with patterns found using field studies. We found that the two-stage detector

Mask R-CNN was more accurate than single-stage detectors like RetinaNet and YOLO, with the Mask R-CNN network performing best overall with mean average precision (mAP) of 0.67 followed by RetinaNet (0.5) and YOLO (0.4). We also show that detection is possible even when pictures are interspersed with complex backgrounds and are not in focus. We expect this approach can be used to address many other ecological questions that benefit from automated flower detection, like studies of flowering phenology or floral resources, and that this approach can therefore complement a wide range of ecological approaches (e.g., community science, experiments, etc.). In all, our study suggests that ecological metrics like floral richness can be efficiently monitored by combining machine learning with easily accessible publicly curated datasets (e.g., Flickr, iNaturalist).

Keywords: Alpine wildflowers; convolutional neural net, climate-change, phenology.

## 3.2 INTRODUCTION

Alpine environments harbor a wide diversity of plant and animal species, but are also extremely susceptible to climate change [1,2]. Species in alpine meadows are limited by a short growing window that is driven by late snowmelt [3] which can make them prone to a warming climate [4]. Floral richness is one measure that provides an important yardstick on the fitness of Alpine wildflowers, as meadow species are generally perennial [19]. Furthermore, flowers are good indicators for measuring the effects of climate change [20–23].

Phenological shifts in flowering could have drastic impacts to flora and fauna[9,10]. For example, pollination might be at risk if pollinators emerge after peak flowering or earlier before the onset of flowering [4,8] as well as the experience of human visitors who

visit to enjoy the wildflower displays [121,122]. An extremely warm year might hasten flowering and make flowers susceptible to frost damage [123]. On other hand, in an extremely wet year (or late snowmelt), flowering might get delayed resulting in pollinator mismatch [124]. Hence, evaluating patterns of floral richness spatiotemporally can be important for understanding the factors driving meadow species reproductive life cycle [11].

Detecting flowers in natural settings (such as alpine meadows) is a complex task, but recent developments in computer vision algorithms are providing new opportunities. Isolating flowers in a picture of a typical mountain meadow is challenging because of the diversity of meadow species, and the wide range of floral structures and backgrounds that exist (like leaves, trees, rocks etc.). This is because flowers are inherently complex because of the extraordinary diversity of floral morphologies resulting from eco-evolutionary dynamics [125]. However, machine learning (ML) vision-based algorithms show promise in detecting life stages or phenologies (e.g., identifying blooming), and in-depth evaluation of a single phenological stage for particular species has been successful (e.g. fruit maturity in passion fruit [24], strawberries [25] and tomatoes [26]; ripeness in papayas [27], and flowering in cotton plants [28]). ML approaches have also been used for classifying the progression of flowering in apples [126,127]. These studies showed that it is possible to evaluate stages of a plant life event (for e.g., fruiting) from images with ML methods. However, these approaches have not been attempted in more species rich and complex systems like Alpine wildflower meadows, where issues related to occlusion, complex backgrounds and a high floral richness may complicate these approaches [128].

An evolving class of machine learning algorithms called deep neural networks (dNNs) have been particularly useful for detecting and segmenting multiple objects in an image. dNN methods can perform object detection (for e.g., a flower) and instance segmentation (for e.g., detect all the instances of a flower). Applications of such algorithms are wide and important, ranging from tree crown detection to identifying cancerous cells [129,130]. Our use of dNNs for mapping flowers is also motivated by their ability to extract subtle features and learn from them efficiently [126]. This means flowers can be individually tagged and segmented in a meadow image with dedicated dNNs.

Broadly, dNN approaches can be divided into one-stage and two-stage techniques. A one-stage algorithm called YOLO [131] is a well-known object instance detection method, but has issues with objects with multiple sizes (pictures where objects are taken at different distances), and background noise. RetinaNet is a comparable one-stage algorithm that overcomes these limitations [132]. However, neither of these methods are able to use some features for classification, so a two-stage algorithm like Mask R-CNN [133], where feature-extraction is built-in with segmentation and pixel-level classification is known to perform more accurately [130,134,135]. All three of these methods i.e., YOLO, RetinaNet and Mask R-CNN apply bounding box detection (i.e., localize the object), but Mask R-CNN, in addition, gives pixel level mask of the class. In addition, Mask R-CNN supports pluggable feature extraction network backbone like U-Net [136] that could a cover wide range of domains.

In this paper, we explore these three state-of-the-art dNN methods (YOLO, RetinaNet and Mask R-CNN) for the detection and localization of Alpine wildflowers in complex settings – i.e., montane meadows. We define identification of a species by its

'flowering' phenophase only i.e., the stage where the species flower(s) are visible. Specifically, we focus on two objectives; i) Evaluating the three dNN based techniques in their ability to detect flowering species in alpine meadows to derive diversity and abundance and ii) using the data to explore an ecological hypothesis on the relationship between elevation and floral richness. Our expectation is that diversity declines with increasing elevation but in a non-linear way.

## 3.3 Materials and Methods

### 3.3.1 Study Area

We conducted our studies in subalpine regions (upward of 1400 m) at MORA, which have an impressive wildflower display that is one of the main summertime visitor attractions of the area. Meadows are snow covered most of the year, even late into July, with peak-flowering that lasts between 2 and 3 weeks between late July and early September – depending on snowmelt dates [5]. We collected data along an existing trail called Reflections Lakes Trail that has been a site for community and focused meadow phenology observations [122,137]. The trail is on the south side of Mount Rainier National Park (MORA) and covers an elevational gradient of 1400 m to 1950 m (Fig. 1).

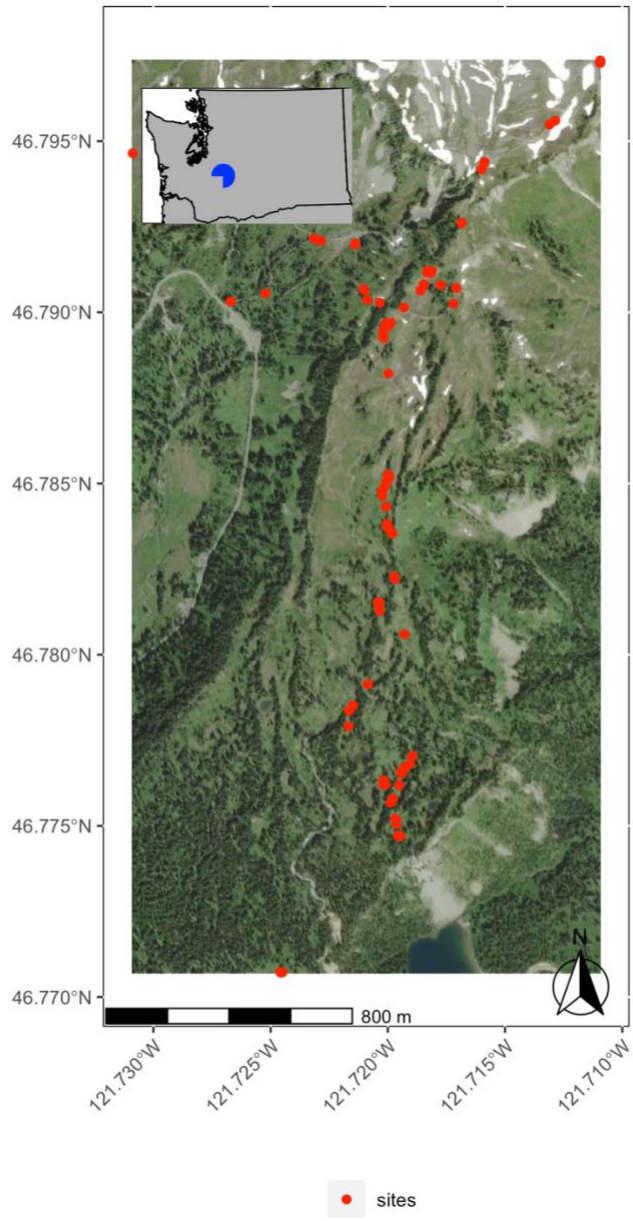


Figure 3-1. Study area at Mt Rainier National Park, red dots indicate the sites along the Reflections Lakes trail at which the pictures were taken for the study. The inset shows the location (in blue) of the study area in the context of Washington State, USA.

### 3.3.2 Data

Volunteers participating in this study collected camera images using an iPhone camera (version 11 Pro) during the flowering season of 2020 and 2021. Sixteen of the 42 species

that commonly occur at the study site were used in this study (see Table 1). The curated dataset contained 1545 high resolution photos: 768 x 1024-pixel (landscape) and 1024 x 768-pixel (portrait). From this dataset, we used 1221 images for training and 324 images for validation and testing. Photos were taken from the trail overlooking the meadows. The portrait and landscape images were then tiled for easier annotation after converting them to a common dimension of 1024 pixel x 1024 pixel. Padding was done so that it was easier to tile the images in equal sizes. For e.g., tiling into 256 x 256 of a 1024 x 1024 format image results in 8 images.

For our training dataset, we annotated each individual image, identifying the number and identity of flowers in each image. We used the LabelMe software to perform the annotation [138], which was then outputted in a JSON (JavaScript Object Notation) file format. These annotations included multiple labels (flowers) with their annotation mask (see Fig. 2 for sample annotations). For application of the Mask R-CNN and RetinaNet dNN algorithm, the annotation outputs were in JSON format, which was converted to a COCO format by the training framework. Because the YOLO algorithms require a specific format (the YOLO format), we converted the LabelMe annotations to this format prior to analysis.

Table 3-1. The training dataset comprised of a total of 1221 images having 4349 instances of flowers annotated. The validation dataset had a total of 166 images having 597 instances of flowers annotated, and the test dataset consisted of a total of 158 images having 407 instances of flowers annotated. All the images were 256 by 256 pixels and collectively represented 16 wildflower species.

| Common Name | Scientific Name | Training Instances | Validation Instances | Test Instances |
|---|---|---|---|---|
| Western Anemone | Anemone occidentalis | 508 | 75 | 24 |
| Bracted Lousewort | Pedicularis bracteosa | 119 | 20 | 21 |
| Mountain Daisy | Erigenon peregrinus | 265 | 39 | 27 |
| Subalpine Lupine | Lupinus arcticus | 902 | 108 | 31 |
| Glacier Lily | Erythronium grandiflorum | 103 | 23 | 21 |
| Cascade Aster | Aster ledophyllus | 177 | 55 | 20 |
| American Bistort | Polygonum bistortoides | 322 | 69 | 26 |
| Grays Lovage | Ligusticum grayi | 123 | 23 | 28 |
| North Microseris | Microseris alpestris | 213 | 20 | 28 |
| Scarlet Paintbrush | Castilleja miniata | 364 | 21 | 20 |
| Sharptooth Angelica | Angelica arguta | 96 | 22 | 26 |
| Avalanche Lily | Erythronium montanum | 129 | 21 | 29 |
| Magenta Paintbrush | Castilleja parviflora | 638 | 34 | 33 |
| Sitka Valerian | Valeriana sitchensis | 95 | 19 | 27 |
| Tall Bluebell | Mertensia paniculata | 145 | 26 | 26 |
| Broadleaf Arnica | Arnica latifolia | 150 | 22 | 20 |

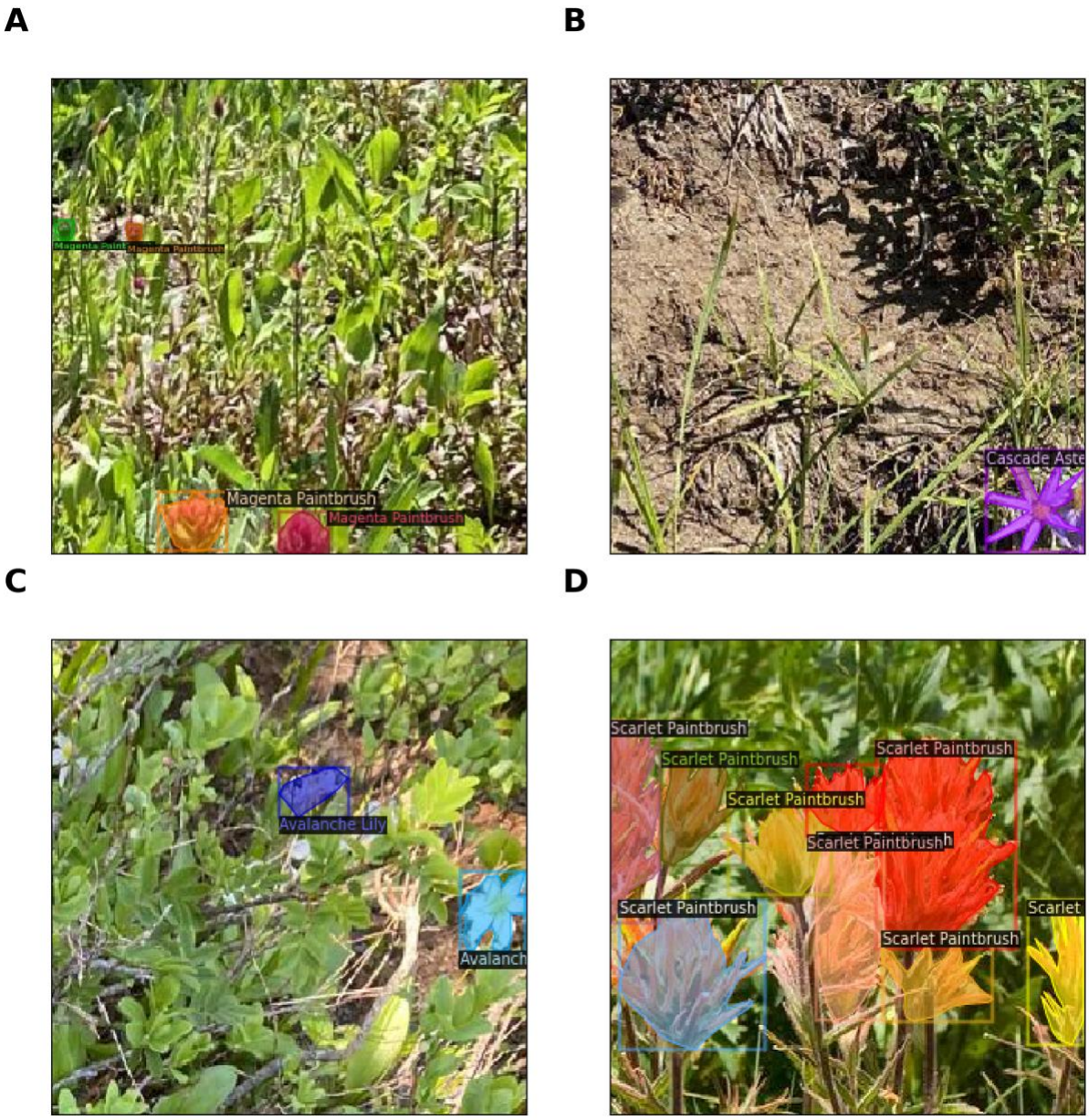


Figure 3-2. Example annotated images from the training dataset. (A) Magenta Paintbrush (Castilleja parviflora) flowers (B) A flowering Cascade Aster (Aster ledophyllus) (C) Avalanche lilies (Erythronium montanum); particularly found earlier in the season (D) Cluster of Scarlet Paintbrush (Castilleja miniata) in full bloom.

### 3.3.3 Machine Learning Methods

We explored how three dNN algorithms performed in detection and segmentation: YOLO, RetinaNet and Mask R-CNN. These three methods are inherently able to extract features and learn from them efficiently [126]. YOLO and RetinaNet support bounding box detection only whereas Mask R-CNN gives the mask of the detected object in addition to bounding-box. We briefly describe the three methods in Appendix A. We used the Detectron2 framework (an implementation of the PyTorch deep learning library) from Facebook Artificial Intelligence Research (FAIR) to train deep neural network models for flower detection for RetinaNet and Mask R-CNN. The Detectron2 has an out-of-the-box implementation of many of the latest CNN algorithms that perform detection and segmentation [139]. Detectron2 allows writing reproducible code with a standardized way of dealing with input data and annotations. Detectron2 also provides boiler plate code for dataset handling, augmentations, training, visualization, and performance metrics. The development, training and inference of the models were performed on Azure Cloud on a server having the configuration Standard NC12_Promo (12 vCPUs, 112 GiB memory with 2 NVIDIA Tesla K80 GPUs). Hyperparameter fine-tuning of all three models was achieved with training for 1500 iterations on Google Colab's K80 GPU.

### 3.3.4 Hyperparameter tuning

We performed hyperparameter tuning using the validation data composed of 500 instances (compared with 4300 instances used in training). Anchors are predefined with aspect ratio and scale to establish the location and size of the bounding box for an object. We used different anchor sizes [140] for the three algorithms. We kept the learning rate

at the default value of 0.00025 with a batch size of 128. We also kept the aspect ratio of 1:1 and scale to double with a total of 5 increments.

### 3.3.5 Evaluation Metrics

We used detection evaluation metrics as defined by the COCO dataset [141]. Intersection over Union (IoU) is a metric that evaluates how the ground truth and predicted detections overlap. IoU of 1 means an ideal detection whereby there is an exact overlap whereas IoU of zero means that there is no overlap. Average Precision (AP) indicates how a detection method performed at a particular IoU. Here, we report AP at 0.5, 0.75 and 1. AP is calculated by taking average of AP of ten IoU thresholds (t= [0.5,0.55,0.60,0.65,0.70,0.75,0.80,.0.85,90,.95] and when averaged across all categories the metric provides us mean average precision (mAP). True positive (TP) would mean IoU > = Threshold, and false positive (FP) is IoU < Threshold. We further inspect the Precision (TP/TP+FP) and Recall (TP/TP+FN) to evaluate frequency of false detections or missed ones; note that higher mAP indicates better detection performance. Mask R-CNN identifies the anchor boxes based on the IoU, which is a critical step in the first part of the network (RPN).

## 3.4 RESULTS

### 3.4.1 Hyperparameter tuning results

We found that maintaining anchor box variations that are multiple of 16 i.e., [16,32,64,128,256] yielded better accuracy (Table 2). The results were consistent with YOLO and RetinaNet.

Table 3-2. Results of anchor box variation with Mask R-CNN.

| Anchor Box Size | AP | AP50 | AP75 | AP | AP50 | AP75 |
|---|---|---|---|---|---|---|
| | Bounding Box | | | Segmentation | | |
| [4, 8, 16, 32, 64] | 11.708 | 23.180 | 10.338 | 12.230 | 22.296 | 12.217 |
| [8, 16, 32, 64, 128] | 13.515 | 24.992 | 12.023 | 13.340 | 23.783 | 13.358 |
| [16, 32, 64, 128, 256] | 18.128 | 31.767 | 17.148 | 17.871 | 30.838 | 20.211 |
| [32, 64, 128, 256,512] | 17.485 | 31.555 | 16.395 | 17.031 | 30.935 | 16.596 |

### 3.4.2 Performance of flower detection methods

We tested the three dNNs algorithms using over 316 images with an average number of 40 labels per flower category. We used two augmentations in training: Resize Shortest Edge and Random Flip. Each of the augmentation methods were trained for 10000 epochs. The time needed for training differed between the three algorithms, ranging from 3 hours for YOLO to 12 hours in Mask R-CNN.

Table 3 lists the mAPs and different IoU thresholds for all the methods. In addition, Mask R-CNN provides a classification metric by identified object size as well. In our case, for small, medium, and large flowers, the mAP was 29.27, 42.76 and 48.80, respectively. To reiterate, bounding box detection applies to all the methods, but segmentation (drawing of masks) only applies to Mask R-CNN, so segmentation scores are only reported for Mask R-CNN.

Table 3-3. Evaluation results for the BBox and Segmentation across the three algorithms. Mean Average precision (mAP) for the three methods. Mask R-CNN has better overall performance.

| Method | mAP BBox | mAP50 BBox | mAP 75 BBox | mAP Seg | mAP50 Seg | mAP75 Seg | Training time (hours) | Iterations (Epochs) |
|---|---|---|---|---|---|---|---|---|
| Mask R-CNN | 45.80 | 67.24 | 52.10 | 31.32 | 59.05 | 30.78 | 12 | 10000 |
| RetinaNet | 23.59 | 51.28 | 18.63 | N/A | N/A | N/A | 8 | 10000 |
| YOLO | 32.66 | 60.68 | 31.91 | N/A | N/A | N/A | 3 | 10000 |

Table 3-4. Per category BBox evaluation (using mAP) results across all three methods. Generally, Mask R-CNN performs better, but YOLO and RetinaNet are better for some of the flowers.

| Flower | Mask R-CNN AP50 | RetinaNet AP50 | YOLO AP50 |
|---|---|---|---|
| Western Anemone | 58.45 | 35.22 | 27.90 |
| American Bistort | 42.79 | 23.55 | 56.80 |
| Avalanche Lily | 40.30 | 18.74 | 52.2 |
| Bracted Lousewort | 58.05 | 18.28 | 4.42 |
| Gray's Lovage | 36.02 | 20.07 | 43.90 |
| Magenta Paintbrush | 42.01 | 29.98 | 27.80 |
| Mountain Daisy | 51.42 | 23.31 | 67.70 |
| North Microseris | 68.22 | 39.10 | 77.80 |
| Sitka Valerian | 45.43 | 21.31 | 10.40 |
| Subalpine Lupine | 38.36 | 24.67 | 39.20 |
| Scarlet Paintbrush | 59.91 | 33.00 | 74.60 |
| Tall Bluebell | 46.04 | 14.45 | 83.00 |
| Glacier Lily | 29.93 | 11.56 | 14.70 |
| Sharptooth Angelica | 24.34 | 08.11 | 48.30 |
| Broadleaf Arnica | 61.09 | 21.80 | 41.50 |
| Cascade Aster | 26.83 | 34.34 | 07.62 |

All three methods were able to isolate the flowers within the visual complexity of a meadow. Table 4 gives the breakdown of mAP by species for all the three methods. On average, Mask R-CNN performs better for all species, but we note differences in

identifying individual species across the methods. For example, Scarlet Paintbrush, Tall Bluebell and North Microseris are detected more accurately by YOLO than by Mask R-CNN or RetinaNet. Regardless of the species, they are detected with higher accuracy if they are towards the foreground of the image and if the image is in focus - see for example Bracted Lousewort being isolated by Mask R-CNN while two species in the horizon are missed (Fig. 3A, B). Similarly, occluded flowers are being missed for example some of the Mountain Daisy flowers that are behind the stalks of other species (Fig. 3C, D).


Figure 3-3. Predictions from the Mask R-CNN method showing the fluency of the model; the percent next to the predicted class represents model confidence. (A and B) A meadow in peak flowering with predictions (B) corresponding original tile. (C and D) Another meadow in peak flowering with predictions (D) corresponding to the original tile.

### 3.4.3 Richness and abundance of flowering species

We found that floral species richness and abundance per photo varies along the elevational gradient. Specifically, we see lower floral richness at lower and higher elevations, and increased richness along mid elevations i.e., 1700-1800 m (Fig. 4A). We also see increased density of flowers at the mid-elevation compared to lower and higher elevations (Fig. 4B). Floral abundance in pictures was also found to vary over time, even over very short periods (Fig. 5).

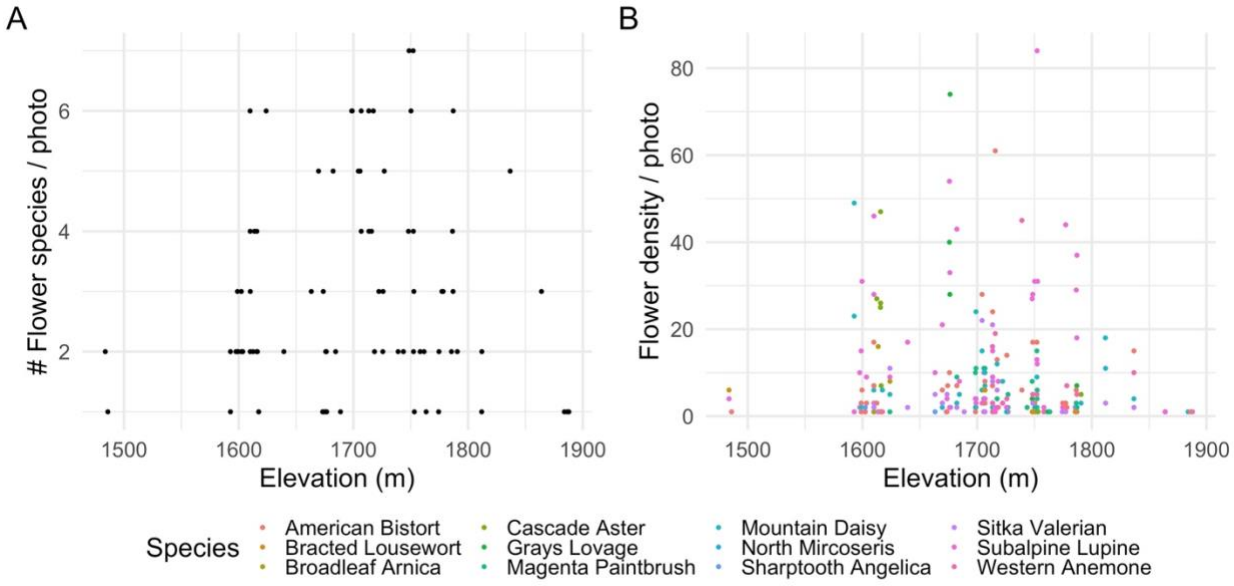


Figure 3-4. Detected species from all the collections using images collected in August 2020; predictions are from YOLO method. (A) Number of unique flower species detected per photo, (B) number of occurrences of each flower species per picture.

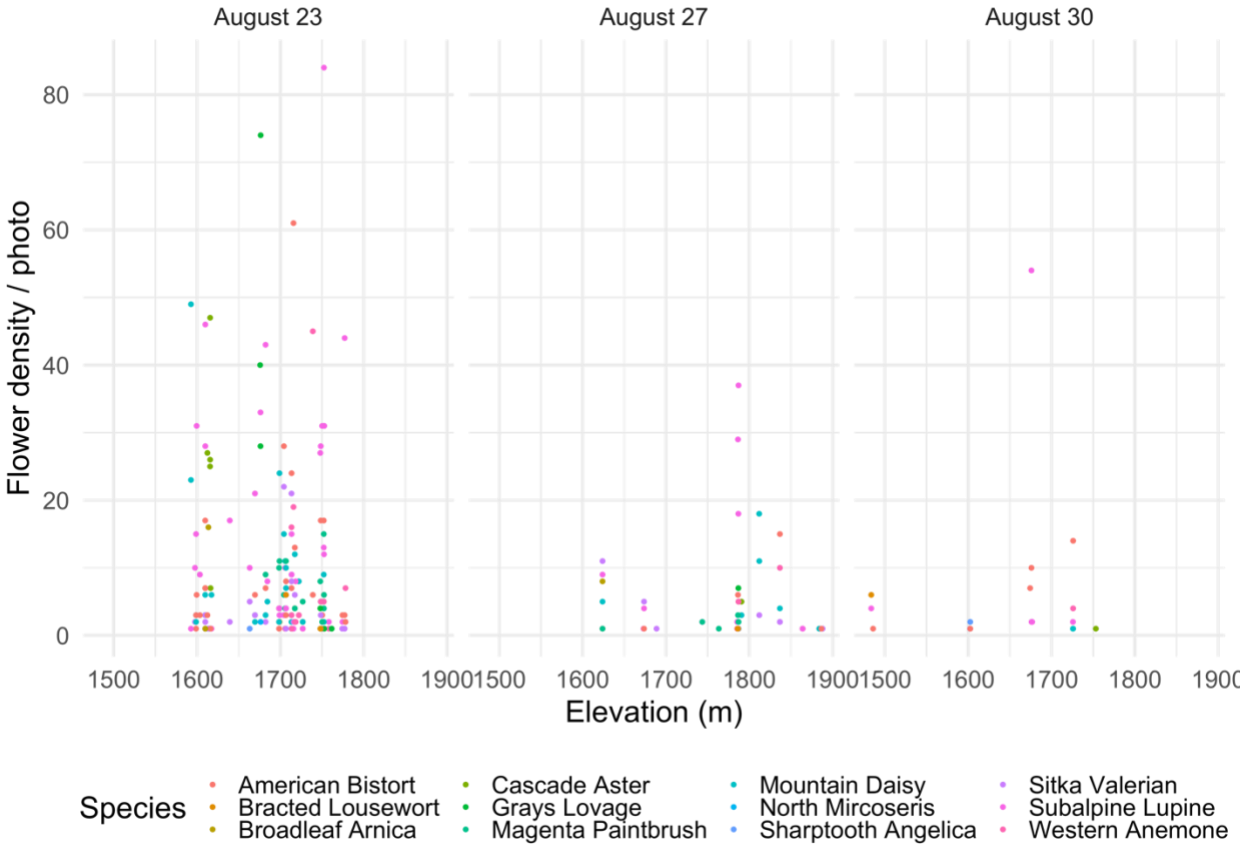


Figure 3-5. Floral density of species by picture for all the collection dates in August 2020.

## 3.5 Discussion

We demonstrate that deep learning techniques can detect, identify, and count wildflowers in photographs, and thereby provide detailed information on flowering occurrences in complex mountain meadow systems. Furthermore, we show that these approaches verify known ecological patterns – that species flowering varies temporally and by elevation at our study site [142,143]. Our results show that with timestamped, geo-coded photographs, deep learning techniques can identify and count flowers, potentially facilitating detailed analyses on flowering richness as well as floral abundance, both key ecological metrics relevant to Alpine wildflower fitness and pollinator performance.

### 3.5.1 Differences in performance and methodologies for identifying flowers in camera images

In our application, the two-stage detector (Mask-R-CNN) was more accurate but slower in its predictions than the one stage detectors (YOLO, RetinaNet). This is consistent with other studies using this approach to count and identify individual flowers of a single species [130,135,144] . Single stage detectors perform classification and regression together on the predicted anchor boxes. The trade-off is that additional stage incurs computation time (but see [145]). The YOLO algorithm which is a single state detector is able to accurately draw boxes around the flowers as well as classify them correctly more speedily than Mask R-CNN (similar finding in [146,147]). However, many flowers are missed in the labelling of the test images, but this is expected since flowers are also missed in the training images.

The imagery obtained for our study came in two dimensions (1024 x 768 and 768 x 1024), but we found that slicing large format images into 256 x 256-pixel images had the lowest classification error. The 256x256 images appear to be of adequate size for object detection for flowers of similar size in the dataset. We found that larger format image samples were computationally demanding. When interpreting the metrics, it is important to note that larger images have greater boxes and objectness errors because sub-images are taken from the same image, and if the sub-image is larger, it will have more boxes in them. And, these errors are summed for each candidate box, the box and objectness error will tend to be larger for the larger images.

### 3.5.2 Flower characteristics, species abundance and implications for detection

Several factors led to variation in our ability to classify and detect flowers. Most wildflower meadows (and thus pictures) include large variation in the abundance of species, like Subalpine Lupine, which tends to be more abundant than other flowers. This results in Subalpine Lupine being detected more so than other flowers like Mountain Daisy which appear more sparsely – which reflects real differences in floral abundance. However, the algorithm also made mistakes, and it could be that the characteristics of some flowers are similar when detection is done. Some meadow wildflowers have an elongated floral structure (e.g., Bracted Lousewort and Subalpine Lupine), while others are more circular in nature (Mountain Daisy and Cascade Aster). Misclassification tended to occur when structures were similar (e.g., Sharptooth Angelica and Gray's Lovage). Unlike in other studies, in which data enhancement helped prevent overfitting and improved accuracy [148], we did not see improvement in accuracy with image augmentations. We also see the shortcoming where occluded flowers were not being isolated (Fig. 3A, C).

Despite these challenges, we were encouraged by the accuracy we achieved in this study as these pictures are inherently complex and therefore challenging for classifiers. Some challenges include inconsistent light levels (e.g., because of slope and aspect) and heterogenous backgrounds (e.g., rocks, lots of plants, etc.). Additionally, labelling errors, namely flowers that are missed while annotating could contribute to decreased accuracy. For e.g., YOLO method penalizes incorrectly drawn boxes i.e., if a flower is not labelled in

the training set, YOLO would penalize the model for drawing a box around the flower despite it being correct.

There also is likely to be sampling bias in this study. Undoubtedly, the chances of picking a location and photographing it is positively correlated with the probability of a flower existing in that place. That means that our pictures are disproportionately likely to contain flowers, potentially predisposing our algorithm to bias toward flowers. One other limitation is that our study does not capture the length of species flowering and does not quantify the flowers by location (elevation) but this could be done if a planned study is performed for e.g. images are captured from a PhenoCams along an elevational gradient.

### 3.5.3 Applicability to ecological research

We believe that combining image analysis with automated floral detection can provide important information on metrics of ecological interest, which can either complement or in some cases even replace more traditional ecological sampling. For example, our analyses show that floral richness and abundance are linked to elevation in a non-linear way (Fig. 4A, B), similar to patterns shown by other studies. Floral richness is driven both by plant diversity (which is influenced by many biotic and abiotic factors), as well as by the probability that species are flowering [5,149,150] – an important indicator of plant fitness resources for pollinators. Thus, this approach could be used to provide information on diversity patterns across gradients, phenological patterns of flowering relative to climate, and floral resources for pollinators.

For abundant species with recognizable flowers, these approaches may also allow for the detection of distributional shifts, if conducted annually. For example, climate

change might alter the distribution of specific wildflowers found in montane environments; possibly as the result of elevational shifts [2,10]. Similarly, a rapid invasion of a non-native flowering species can be detected using this approach. The combination of the floral detection approaches introduced here with drone imagery could allow for assessment of such shifts over large spatial areas, or in remote areas.

However, there are some caveats to using flowers in pictures to quantify richness and/or abundance. For one, floral richness is not total species richness, as many perennial plants do not flower every year or in every location. A second important caveat in our measures of floral abundance is that fields of view vary from picture to picture, which means the total area available to flowers (and thus flowering abundance) will vary even if underlying patterns detected by more traditional ecological approaches (e.g., $m^2$ quadrats) do not. However, we believe the promise of these approaches for generating large volumes of data with less manpower (as described in [151]) outweigh these issues. Moreover, these challenges can be overcome with large volumes of pictures / data or by employing fixed cameras (e.g., webcams, PhenoCams) with common fields of view. For example, if pictures are taken automatically via PhenoCams or TrailCams, these fixed cameras allow for pictures to be taken in the same place repeatedly, which saves time and gives an advantage over taking them oneself as one gets to control when and where the pictures are to be taken. Then, differences in floral richness would not be due to differences in field of view (the way the picture was taken). Regardless, we feel that floral richness derived from photos is still of interest as it gives a quick proxy of site-specific flowering patterns. We also believe our approach can also be used to address questions about phenology that otherwise would not be possible with traditional ecological

sampling. As photos are engrained with date/time and location information, we could also have linked these measures to time to provide information on phenology (as in Breckheimer et al 2019). In this case, crowd sourced pictures (e.g., Flickr, iNaturalist, Instagram) increased the data volume (# of observations) and allowed for quantification of phenology in previous years (i.e., going back in time). This and similar approaches, however, require specific statistical approaches. For example, crowd sourced pictures (e.g., from Flickr) are not taken according to a specific experimental design, which means using these pictures could result in an unbalanced representation – e.g., spatial or temporal gaps in distributions (e.g., Mt. Rainier meadow pictures predominantly may be from around visitor centers, weather permitting times and might have weekend effects). However, one can deal with skewness introduced by source images by statistically controlling for any spatiotemporal bias in pictures taken [121].

Finally, we believe that the floral detection combined with ML approaches could allow for quick assessment of ecological patterns, which could be important in a changing world. For example, it might be valuable to know about the distribution of a rare species threatened by climate change and / or the co-occurrence of wildflowers that compete from pollinators whose phenology shifts at different rates or due to climate driven changes in species composition [152]. Documenting species presence-absence using any of our methods as applied to existing (crowd-sourced) pictures could address questions such as these, and provide a preliminary assessment of patterns before designing new (and costly) experiments [153].

### 3.5.4 Future directions

Although our flower detection algorithm performed well, there are likely improvements that could be made to detect a greater number of wildflowers and better distinguish between similar looking flowers. For example, we have not looked at the fluorescence signal but hypothesize that shape and color are being learnt during feature extraction, and as future directions, incorporation of additional information (for e.g., thermal, or infrared bands) might help in the presented algorithms.

Embedding the trained models in community science apps (Mobile Application) could be one of the end uses of these algorithms. In other domains, apps have increased community involvement, e.g., the bird identification app, Merlin [154] and the plant identification functionality of iNaturalist. Use in apps both provides the public with a way to identify flowers and a way to gather data.

Combining photo data with drone and remotely sensed products could potentially offer opportunities to measure things in remote areas and on larger spatial scales. Improvement of algorithms that can use additional bands (produced in hyperspectral or multispectral imagery) could be promising and might detect more kinds of flowers and distinguish between flowers that are more similar.

## 3.6 CONCLUSIONS

We explored three deep learning techniques to detect and count flowering plants in alpine meadows. The two-stage detector we used (Mask R-CNN) performed better than the single-stage detectors (YOLO and RetinaNet), with the first stage cropping flowers and filtering out the background; and the second stage classifying and detecting the flowers.

However, all techniques gave us detailed information on flowering occurrences in meadows, allowing us to calculate floral richness as well as flower abundance. These approaches were also able to capture known ecological patterns; namely that flowering varies temporally (i.e., phenology) and by elevation. Our trained models could be applied in areas where our focal species (from Mt. Rainier National Park) are also dominant, including Alpine sites in other high mountain sites in the Pacific Northwest (e.g., Washington, Oregon, British Columbia). In the future, including a larger catalog of species for training could allow for coverage from other areas, e.g., Rocky Mountain meadows or Swiss Alpine meadows. More generally, these approaches could aid in climate change studies, because they can be used in a non-invasive, low-investment way to rapidly get information on spatiotemporal flowering patterns.

**Acknowledgements**

We want to acknowledge support of eScience institute and Hille Ris Lambers lab in developing the ideas and supporting discussions. Authors want to acknowledge field help from Tommy Li and Richard Lee. We also acknowledge support from Microsoft AI for Earth and Microsoft Azure collegial support. This study was carried out on the ancestral lands of the people of the Nisqually, Puyallup, Squaxin Island, Muckleshoot, Yakama, and Cowlitz Nations.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Authors' contributions**

A.J. and J.H.R.L conceived the ideas and designed the methodology; A.J. and MeadoWatch volunteers collected the data; A.J. analyzed the data; A.J., E.J.T. and J.H.R.L. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## 3.7 APPENDIX A

**Mask R-CNN**

We used Mask R-CNN (Mask Region based CNN), a deep learning algorithm that generates bounding boxes and segmentation for each instance of an object (here a flower) in an image (Fig. A.1). It is classified as a two-pass network; where the first part is dedicated to feature learning, while the second part is for detection and segmentation. In the end, the network identifies individual objects in an image and localizes each of them with a bounding box [135] along with predicting pixel level segmentation masks. Instance segmentation in Mask R-CNN comprises of object detection and semantic segmentation. Mask R-CNN is an improvement upon Faster R-CNN [155] which used region proposal and classification networks to detect region of interest (RoI) and bounding boxes. Mask R-CNN supports various backbones as feature pyramid network (FPN) in order to produce feature map of the input image. Region proposal network (RPN) is then applied atop feature maps that returns object proposals with corresponding objectness scores. Objectness tells how good the model did to predict an object regardless of class. A RoIAlign layer is then applied to filter the proposals; the predecessor (Faster R-CNN [155]) used RoI pooling which made it slower. The key feature is the use of RoI Align instead of RoiPool (quantization free layer). The proposals are then passed to a fully

connected layer that returns class and bounding boxes for the objects. Simultaneously, the mask branch that is a fully convolutional network (FCN) works on each RoI, and outputs a segmented mask at a pixel level. Mask R-CNN uses class loss, box loss and mask loss for scoring the model. In our study Mask R-CNN was configured with ResNext 101 backbone.

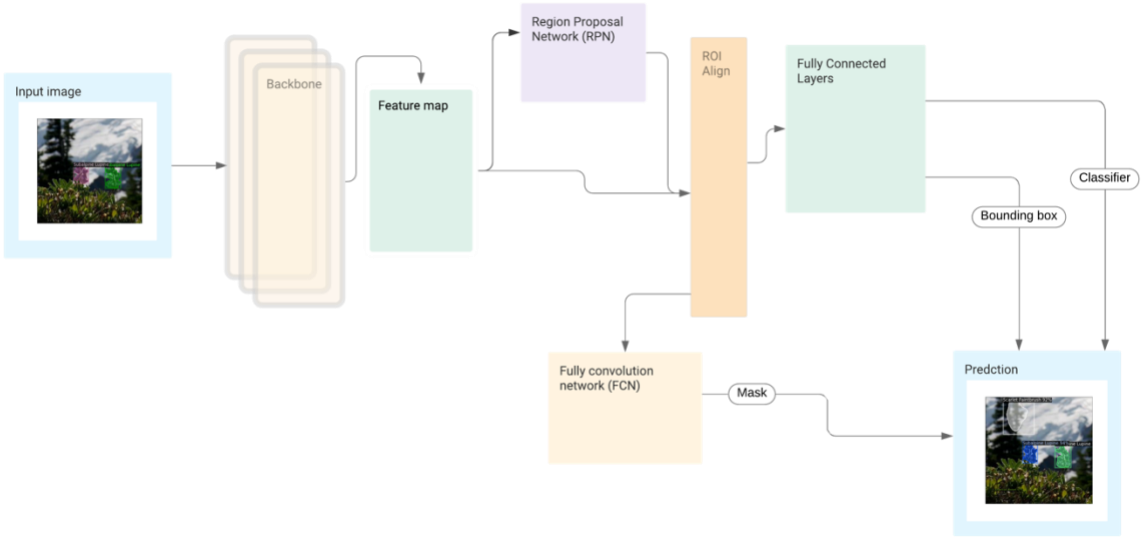


Figure A.1: High level architecture of the Mask R-CNN algorithm.

**YOLO**

YOLO stands for “You Only Look Once” that uses regression to achieve object detection and localization [131,156]. YOLO is a one stage method that works by regressing the bounding boxes and class probabilities and is known for its accuracy and high speed (Fig. A.2). YOLO has constantly evolved by releasing significant iterations with changes to its network structure resulting in better speed, consistent classification accuracy and recognizing smaller targets [131,156,157]. The YOLO method divides each input into

regions or grids (S by S where S =8 for e.g.). For each grid it predicts B bounding boxes with confidence scores along with class conditional probabilities C. Confidence score translates to accuracy of the predicted object(s) in a grid. YOLO employs non-maximum suppression (NMS) similarly to Mask R-CNN if there are multiple candidate bounding boxes for the same target object. The loss function has three parts: objectness, box and classification. Objectness like in Mask R-CNN is a measure of how well the model did to predict an object irrespective of the class, Box tells how close the predicted box aligns with the actual box regardless of class and Classification loss tells how well each grid is classified with respect to its class probabilities. The YOLO algorithm penalizes the model's inability to draw the bounding box correctly and where it missed estimating it completely. YOLO v3 was the last publicly available version by the original authors, and the recent release (unofficially known as Yolo v5) is now maintained by Ultralytics (GitHub handle - https://github.com/ ultralytics/yolov3)

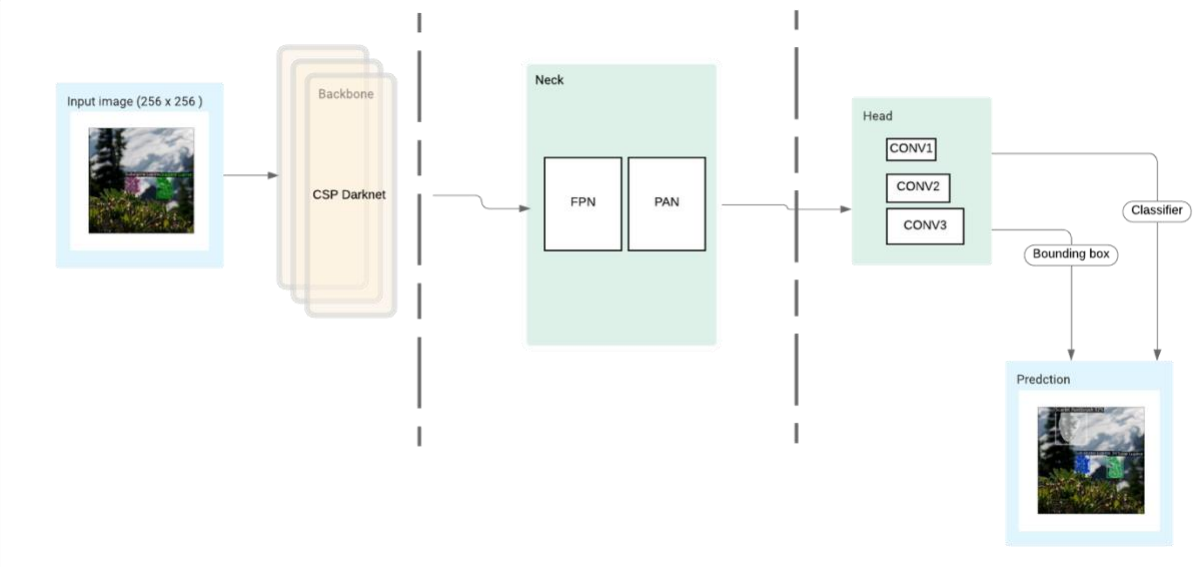


Figure A.2: High level architecture of the YOLO algorithm.

**Retina-Net**

RetinaNet is a two-part learner where the convolution layer learns from the architecture of the images. The architecture uses a feature pyramid network (FPN) backbone on the top of ResNet architecture to generate a rich, multi-scale convolutional feature pyramid (Fig. A.3). Two subnetworks are attached to it next – classification subnet and regression subnet. The first subnetwork is for classifying anchor boxes, while the second subnetwork is for regressing anchor boxes to the ground-truth boxes. The complex head and body of the network involves convolutions which reduces the image to a deep low-resolution image. The network is tuned by using the regression loss and the classification loss. Both metrics look at overall objectness and classification ability of the model. Regression loss is calculated by using smooth L1 whereas classification is evaluated by Focal loss.

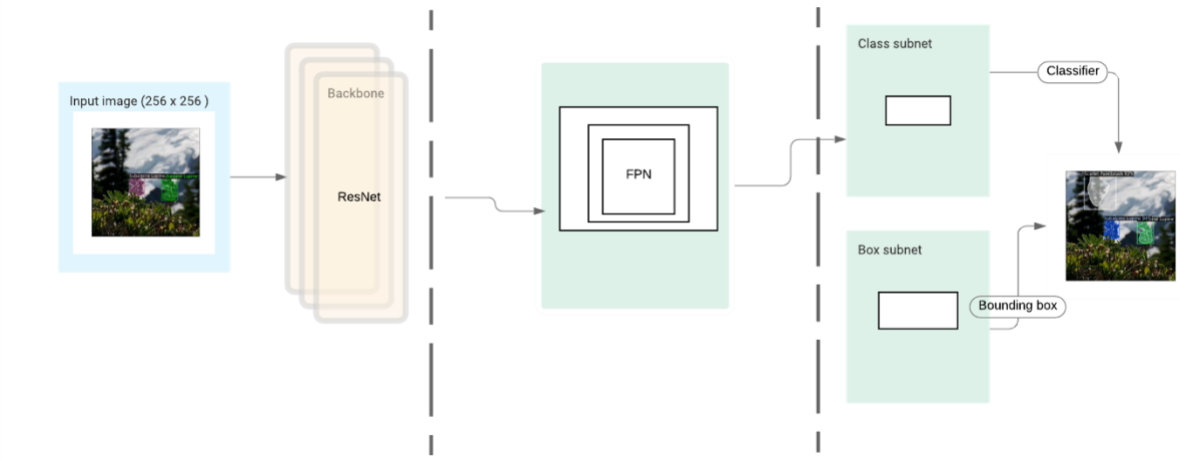


Figure A.3: High level architecture of the RetinaNet algorithm.

# CONCLUSION

In this dissertation, I set out to build tools which could monitor climatic factors influencing Alpine wildflower habitats as well as climatic impacts on Alpine Wildflowers. Specifically, I wanted to take advantage of new technological and statistical developments to improve our ability to capture important ecological drivers in Alpine systems (snow) and important ecological processes (peak wildflower season, and flower density). I therefore built and used tools in chapters 1-3 to map the heterogeneity of snow cover, monitor the location and timing of Alpine flowering at broad spatiotemporal scales, and evaluate patterns of floral richness at a finer spatial resolution. These approaches can support efforts in the study of biological impacts of climate change in montane ecosystems by enabling a better understanding of current climate impacts and future climate change implications.

The work in chapter 1 demonstrates that we can improve the identification of snow-covered areas in forested areas by combining CubeSat imagery, Lidar, and machine learning approach (convolution neural net), which would help ecologists and hydrologists estimate late-season snowpack distribution. Specifically, we explored the use of vegetation structure (canopy height and NDVI), elevation, and elevation-derived attributes (from Digital Elevation Model), in combination with satellite imagery, to map snow in forested areas. Our results show that combining NDVI (proxy of canopy greenness calculated by taking the normalized difference of NIR and Red band) with PlanetScope imagery results in the best snow-identification performance in both forested and open areas compared to other models. We also show, using data from Colorado (USA) and Engadin (CH), that models parameterized in one location (California, USA) can

accurately predict snow cover in another region - highlighting the transferability of this approach. We show that multispectral bands alone are sufficient in mapping snow-covered areas across many landscape types. The approach we developed is scalable, as the availability of additional spectral bands from future satellites or additional parameters of interest can be added for training and prediction. For instance, Planet Labs has introduced additional spectral bands (Red-Edge and Green II) that give a wider coverage of visible spectrum. More broadly, the high-resolution nature of snow cover area mapping shown here could dovetail into development of automated downstream projections (like snow water equivalent (SWE)) at a high spatial resolution.

The work in chapter 2 showed that peak flowering in alpine meadows can be detected using fine resolution CubeSat imagery, and that the inclusion of coarser resolution imagery substantially improves accuracy of these predictions. Specifically, we used moderate resolution imagery (30-m Landsat and 10-m Sentinel) with fine resolution PlanetScope (3-m) imagery to delineate flowering in Alpine meadows. In particular, NDVI was found to be a significant predictor of peak flowering phenology in alpine/subalpine meadow. Because NDVI can be subject to atmospheric effects, alternative vegetative indices like enhanced vegetation index (EVI) could be used as a replacement in future explorations. Similarly, although we found that using visible bands (red, green, and blue) and NIR band was sufficient in narrowing down the flowering window, the use of finer bands at the red-edge region might help further augment the flowering signal in future work. For example, use of short-wave infrared bands have been found to be useful in accentuating signals of senescent vegetation. Thus, the recent availability of additional bands in the visible part of the spectrum (Coastal Blue, Green II, Yellow, Red-Edge) via PlanetScope could improve these phenological assessments at finer

resolutions. Models were more accurate in prediction of flowering when all the sources of imagery were combined (as opposed to models using PlanetScope data). We were also able to identify flowering windows in an unusually warm year (2015) using this approach, despite the model being calibrated from data collected in average climate years. This implies that this approach can be used to predict and detect peak flowering windows, even in changing climatic conditions. Furthermore, improved predictions of peak flowering could be incorporated into park logistical planning to increase visitation and provide better visitor experience.

The work in chapter 3 demonstrates that we can get detailed information on flowering occurrences in alpine meadows using deep learning techniques (Mask R-CNN, RetinaNet and YOLO) and widely available images, allowing us to calculate floral richness as well as flower abundance. We found that the two-stage detector Mask R-CNN was more accurate than single-stage detectors like RetinaNet and YOLO and allowed us to capture known ecological patterns – that species flowering varies temporally and by elevation. Thus, we believe this approach is useful for monitoring flowering in any diverse ecological system, not just Alpine wildflower meadows. The ML models we developed in this study could potentially be used as is, in locations with substantial overlap in focal species (e.g., Alpine sites in Mt. Rainier National Park and Rocky Mountain National Park), which would allow us to replicate this study in other alpine sites in the Pacific Northwest (e.g., Washington, Oregon and British Columbia) harboring similar alpine meadows. We also anticipate this approach can be applied to a wide variety of data sources; including crowd-sourced images, images collected from stationary cameras (e.g., PhenoCams), and high-resolution drone imagery. In total, any of these approaches will improve our ability to monitor fine-scale flowering patterns at high spatiotemporal accuracy, even in relatively

inaccessible areas – potentially revolutionizing our ability to detect climate-change driven phenological shifts and changes to the spatial and temporal provisioning of floral resources for pollinators.

In all, this body of work shows that recent advances in remotely sensed imagery and ML approaches can be combined to build tools that can help ecologists obtain basic and applied insights in the current and future effects of climate change. The resolution of current imagery products is only bound to improve, spurring the development of increasingly improved (in ease and accuracy) technological tools. The approaches presented also demonstrate that cross-discipline investigations are pivotal in climate change impact studies because they allow scientists to integrate disparate data sources and statistical tools to address new questions. As an ecologist with a background in data science, I also believe and hope that these explorations can contribute toward actions to better protect our natural world.

# BIBLIOGRAPHY


1. Ganjurjav, H.; Gao, Q.; Gornish, E.S.; Schwartz, M.W.; Liang, Y.; Cao, X.; Zhang, W.; Zhang, Y.; Li, W.; Wan, Y.; et al. Differential Response of Alpine Steppe and Alpine Meadow to Climate Warming in the Central Qinghai–Tibetan Plateau. *Agricultural and Forest Meteorology* **2016**, *223*, 233–240, doi:10.1016/j.agrformet.2016.03.017.
2. Steinbauer, M.J.; Grytnes, J.-A.; Jurasinski, G.; Kulonen, A.; Lenoir, J.; Pauli, H.; Rixen, C.; Winkler, M.; Bardy-Durchhalter, M.; Barni, E.; et al. Accelerated Increase in Plant Species Richness on Mountain Summits Is Linked to Warming. *Nature* **2018**, *556*, 231–234, doi:10.1038/s41586-018-0005-6.
3. Cardinale, B.J.; Wright, J.P.; Cadotte, M.W.; Carroll, I.T.; Hector, A.; Srivastava, D.S.; Loreau, M.; Weis, J.J. Impacts of Plant Diversity on Biomass Production Increase through Time Because of Species Complementarity. *Proceedings of the National Academy of Sciences* **2007**, *104*, 18123–18128, doi:10.1073/pnas.0709069104.
4. Inouye, D.W. EFFECTS OF CLIMATE CHANGE ON PHENOLOGY, FROST DAMAGE, AND FLORAL ABUNDANCE OF MONTANE WILDFLOWERS. *Ecology* **2008**, *89*, 353–362, doi:10.1890/06-2128.1.
5. Theobald, E.J.; Breckheimer, I.; HilleRisLambers, J. Climate Drives Phenological Reassembly of a Mountain Wildflower Meadow Community. *Ecology* **2017**, *98*, 2799–2812, doi:10.1002/ecy.1996.
6. Burrows, M.T.; Schoeman, D.S.; Buckley, L.B.; Moore, P.; Poloczanska, E.S.; Brander, K.M.; Brown, C.; Bruno, J.F.; Duarte, C.M.; Halpern, B.S.; et al. The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science* **2011**, *334*, 652–655, doi:10.1126/science.1210288.
7. Parmesan, C. Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics* **2006**, *37*, 637–669, doi:10.1146/annurev.ecolsys.37.091305.110100.
8. Ogilvie, J.E.; Griffin, S.R.; Gezon, Z.J.; Inouye, B.D.; Underwood, N.; Inouye, D.W.; Irwin, R.E. Interannual Bumble Bee Abundance Is Driven by Indirect Climate Effects on Floral Resource Phenology. *Ecol Lett* **2017**, *20*, 1507–1515, doi:10.1111/ele.12854.
9. Anthelme, F.; Dangles, O. Plant–Plant Interactions in Tropical Alpine Environments. *Perspectives in Plant Ecology, Evolution and Systematics* **2012**, *14*, 363–372, doi:10.1016/j.ppees.2012.05.002.
10. Gottfried, M.; Pauli, H.; Futschik, A.; Akhalkatsi, M.; Barančok, P.; Benito Alonso, J.L.; Coldea, G.; Dick, J.; Erschbamer, B.; Fernández Calzado, M.R.; et al. Continent-Wide Response of Mountain Vegetation to Climate Change. *Nature Clim Change* **2012**, *2*, 111–115, doi:10.1038/nclimate1329.
11. Inouye, D.W.; Saavedra, F.; Lee-Yang, W. Environmental Influences on the Phenology and Abundance of Flowering by Androsace Septentrionalis (Primulaceae). *American Journal of Botany* **2003**, doi:10.3732/ajb.90.6.905.
12. Planet, T. *Planet Application Program Interface: In Space for Life on Earth*; Planet, 2018;
13. Houborg, R.; McCabe, M. Daily Retrieval of NDVI and LAI at 3 m Resolution via the Fusion of CubeSat, Landsat, and MODIS Data. *Remote Sensing* **2018**, *10*, 890, doi:10.3390/rs10060890.
14. Mazzotti, G.; Currier, W.R.; Deems, J.S.; Pflug, J.M.; Lundquist, J.D.; Jonas, T. Revisiting Snow Cover Variability and Canopy Structure Within Forest Stands: Insights From

Airborne Lidar Data. *Water Resour. Res.* **2019**, *55*, 6198–6216, doi:10.1029/2019WR024898.
15. Douglas, T.A.; Zhang, C. Machine Learning Analyses of Remote Sensing Measurements Establish Strong Relationships between Vegetation and Snow Depth in the Boreal Forest of Interior Alaska. *Environ. Res. Lett.* **2021**, *16*, 065014, doi:10.1088/1748-9326/ac04d8.
16. Raleigh, M.S.; Small, E.E. Snowpack Density Modeling Is the Primary Source of Uncertainty When Mapping Basin-wide SWE with Lidar. *Geophys. Res. Lett.* **2017**, *44*, 3700–3709, doi:10.1002/2016GL071999.
17. Cannistra, A.F.; Shean, D.E.; Cristea, N.C. High-Resolution CubeSat Imagery and Machine Learning for Detailed Snow-Covered Area. *Remote Sensing of Environment* **2021**, *258*, 112399, doi:10.1016/j.rse.2021.112399.
18. McGowan, H.; Borthwick, K.; Schwartz, A.; Callow, J.N.; Bilish, S.; Browning, S. Atmospheric Rivers: An Overlooked Threat to the Marginal Snowpack of the Australian Alps. *Journal of Hydrometeorology* **2021**, *22*, 2521–2532.
19. Roslin, T.; Antão, L.; Hällfors, M.; Meyke, E.; Lo, C.; Tikhonov, G.; Delgado, M. del M.; Gurarie, E.; Abadonova, M.; Abduraimov, O.; et al. Phenological Shifts of Abiotic Events, Producers and Consumers across a Continent. *Nat. Clim. Chang.* **2021**, *11*, 241–248, doi:10.1038/s41558-020-00967-7.
20. Dreyer, L.L.; Esler, K.J.; Zietsman, J. Flowering Phenology of South African Oxalis—Possible Indicator of Climate Change? *South African Journal of Botany* **2006**, *72*, 150–156, doi:10.1016/j.sajb.2005.06.009.
21. Miller-Rushing, A.J.; Inouye, D.W.; Primack, R.B. How Well Do First Flowering Dates Measure Plant Responses to Climate Change? The Effects of Population Size and Sampling Frequency. *Journal of Ecology* **2008**, *96*, 1289–1296, doi:10.1111/j.1365-2745.2008.01436.x.
22. Gallagher, R.V.; Hughes, L.; Leishman, M.R. Phenological Trends among Australian Alpine Species: Using Herbarium Records to Identify Climate-Change Indicators. *Aust. J. Bot.* **2009**, *57*, 1, doi:10.1071/BT08051.
23. Rawal, D.S.; Kasel, S.; Keatley, M.R.; Nitschke, C.R. Herbarium Records Identify Sensitivity of Flowering Phenology of Eucalypts to Climate: Implications for Species Response to Climate Change: Herbarium Records Identify Phenological Shift. *Austral Ecology* **2015**, *40*, 117–125, doi:10.1111/aec.12183.
24. Tu, S.; Xue, Y.; Zheng, C.; Qi, Y.; Wan, H.; Mao, L. Detection of Passion Fruits and Maturity Classification Using Red-Green-Blue Depth Images. *Biosystems Engineering* **2018**, *175*, 156–167, doi:10.1016/j.biosystemseng.2018.09.004.
25. Chen, Y.; Lee, W.S.; Gan, H.; Peres, N.; Fraisse, C.; Zhang, Y.; He, Y. Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages. *Remote Sensing* **2019**, *11*, 1584, doi:10.3390/rs11131584.
26. Wan, P.; Toudeshki, A.; Tan, H.; Ehsani, R. A Methodology for Fresh Tomato Maturity Detection Using Computer Vision. *Computers and Electronics in Agriculture* **2018**, *146*, 43–50, doi:10.1016/j.compag.2018.01.011.
27. Santos Pereira, L.F.; Barbon, S.; Valous, N.A.; Barbin, D.F. Predicting the Ripening of Papaya Fruit with Digital Imaging and Random Forests. *Computers and Electronics in Agriculture* **2018**, *145*, 76–82, doi:10.1016/j.compag.2017.12.029.

28. Jiang, Y.; Li, C.; Xu, R.; Sun, S.; Robertson, J.S.; Paterson, A.H. DeepFlower: A Deep Learning-Based Approach to Characterize Flowering Patterns of Cotton Plants in the Field. *Plant Methods* **2020**, *16*, 156, doi:10.1186/s13007-020-00698-y.
29. Mountain Research Initiative EDW Working Group Elevation-Dependent Warming in Mountain Regions of the World. *Nature Climate Change* **2015**, *5*, 424–430, doi:10.1038/nclimate2563.
30. Li, D.; Wrzesien, M.L.; Durand, M.; Adam, J.; Lettenmaier, D.P. How Much Runoff Originates as Snow in the Western United States, and How Will That Change in the Future?: Western U.S. Snowmelt-Derived Runoff. *Geophys. Res. Lett.* **2017**, *44*, 6163–6172, doi:10.1002/2017GL073551.
31. Thackeray, C.W.; Derksen, C.; Fletcher, C.G.; Hall, A. Snow and Climate: Feedbacks, Drivers, and Indices of Change. *Curr Clim Change Rep* **2019**, *5*, 322–333, doi:10.1007/s40641-019-00143-w.
32. Mote, P.W.; Hamlet, A.F.; Clark, M.P.; Lettenmaier, D.P. DECLINING MOUNTAIN SNOWPACK IN WESTERN NORTH AMERICA*. *Bull. Amer. Meteor. Soc.* **2005**, *86*, 39–50, doi:10.1175/BAMS-86-1-39.
33. Sethi, M.L.; Theobald, E.J.; Breckheimer, I.; Hille Ris Lambers, J. Early Snowmelt and Warmer, Drier Summers Shrink Postflowering Transition Times in Subalpine Wildflowers. *Ecology* **2020**, *101*, doi:10.1002/ecy.3171.
34. CaraDonna, P.J.; Iler, A.M.; Inouye, D.W. Shifts in Flowering Phenology Reshape a Subalpine Plant Community. *Proc Natl Acad Sci USA* **2014**, *111*, 4916–4921, doi:10.1073/pnas.1323073111.
35. Hessburg, P.; Reynolds, K.; Salter, R.; Dickinson, J.; Gaines, W.; Harrod, R. Landscape Evaluation for Restoration Planning on the Okanogan-Wenatchee National Forest, USA. *Sustainability* **2013**, *5*, 805–840, doi:10.3390/su5030805.
36. Churchill, D.J.; Larson, A.J.; Dahlgreen, M.C.; Franklin, J.F.; Hessburg, P.F.; Lutz, J.A. Restoring Forest Resilience: From Reference Spatial Patterns to Silvicultural Prescriptions and Monitoring. *Forest Ecology and Management* **2013**, *291*, 442–457, doi:10.1016/j.foreco.2012.11.007.
37. Kobori, H.; Dickinson, J.L.; Washitani, I.; Sakurai, R.; Amano, T.; Komatsu, N.; Kitamura, W.; Takagawa, S.; Koyama, K.; Ogawara, T.; et al. Citizen Science: A New Approach to Advance Ecology, Education, and Conservation. *Ecol Res* **2016**, *31*, 1–19, doi:10.1007/s11284-015-1314-y.
38. Inouye, D.W.; Morales, M.A.; Dodge, G.J. Variation in Timing and Abundance of Flowering by Delphinium Barbeyi Huth (Ranunculaceae): The Roles of Snowpack, Frost, and La Niña, in the Context of Climate Change. *Oecologia* **2002**, *130*, 543–550, doi:10.1007/s00442-001-0835-y.
39. Meshyk, A.; Barushka, M.; Marozava, V. Snow as a Contributor to Spring Flooding in Belarus. *Environ Sci Pollut Res* **2021**, *28*, 18826–18836, doi:10.1007/s11356-020-09638-8.
40. Jonas, T.; Marty, C.; Magnusson, J. Estimating the Snow Water Equivalent from Snow Depth Measurements in the Swiss Alps. *Journal of Hydrology* **2009**, *378*, 161–167, doi:10.1016/j.jhydrol.2009.09.021.
41. Currier, W.R.; Lundquist, J.D. Snow Depth Variability at the Forest Edge in Multiple Climates in the Western United States. *Water Resour. Res.* **2018**, *54*, 8756–8773, doi:10.1029/2018WR022553.

42. Currier, W.R.; Pflug, J.; Mazzotti, G.; Jonas, T.; Deems, J.S.; Bormann, K.J.; Painter, T.H.; Hiemstra, C.A.; Gelvin, A.; Uhlmann, Z.; et al. Comparing Aerial Lidar Observations With Terrestrial Lidar and Snow-Probe Transects From NASA's 2017 SnowEx Campaign. *Water Resour. Res.* **2019**, *55*, 6285–6294, doi:10.1029/2018WR024533.
43. 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.; et al. 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* **2016**, *184*, 139–152, doi:10.1016/j.rse.2016.06.018.
44. Frazier, A.E.; Hemingway, B.L. A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery. *Remote Sensing* **2021**, *13*, 3930, doi:10.3390/rs13193930.
45. Kuter, S. Completing the Machine Learning Saga in Fractional Snow Cover Estimation from MODIS Terra Reflectance Data: Random Forests versus Support Vector Regression. *Remote Sensing of Environment* **2021**, *255*, 112294, doi:10.1016/j.rse.2021.112294.
46. Wang, X.; Gao, X.; Zhang, X.; Wang, W.; Yang, F. An Automated Method for Surface Ice/Snow Mapping Based on Objects and Pixels from Landsat Imagery in a Mountainous Region. *Remote Sensing* **2020**, *12*, 485, doi:10.3390/rs12030485.
47. Czyzowska-Wisniewski, E.H.; van Leeuwen, W.J.D.; Hirschboeck, K.K.; Marsh, S.E.; Wisniewski, W.T. Fractional Snow Cover Estimation in Complex Alpine-Forested Environments Using an Artificial Neural Network. *Remote Sensing of Environment* **2015**, *156*, 403–417, doi:10.1016/j.rse.2014.09.026.
48. Liu, C.; Huang, X.; Li, X.; Liang, T. MODIS Fractional Snow Cover Mapping Using Machine Learning Technology in a Mountainous Area. *Remote Sensing* **2020**, *12*, 962, doi:10.3390/rs12060962.
49. Belmonte, A.; Sankey, T.; Biederman, J.; Bradford, J.; Goetz, S.; Kolb, T. UAV-Based Estimate of Snow Cover Dynamics: Optimizing Semi-Arid Forest Structure for Snow Persistence. *Remote Sensing* **2021**, *13*, 1036, doi:10.3390/rs13051036.
50. Nijhawan, R.; Das, J.; Raman, B. A Hybrid of Deep Learning and Hand-Crafted Features Based Approach for Snow Cover Mapping. *International Journal of Remote Sensing* **2019**, *40*, 759–773, doi:10.1080/01431161.2018.1519277.
51. Tsai; Dietz; Oppelt; Kuenzer Wet and Dry Snow Detection Using Sentinel-1 SAR Data for Mountainous Areas with a Machine Learning Technique. *Remote Sensing* **2019**, *11*, 895, doi:10.3390/rs11080895.
52. Musselman, K.N.; Molotch, N.P.; Brooks, P.D. Effects of Vegetation on Snow Accumulation and Ablation in a Mid-Latitude Sub-Alpine Forest. *Hydrol. Process.* **2008**, *22*, 2767–2776, doi:10.1002/hyp.7050.
53. Kirchner, P.B.; Bales, R.C.; Molotch, N.P.; Flanagan, J.; Guo, Q. LiDAR Measurement of Seasonal Snow Accumulation along an Elevation Gradient in the Southern Sierra Nevada, California. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4261–4275, doi:10.5194/hess-18-4261-2014.
54. Tennant, C.J.; Harpold, A.A.; Lohse, K.A.; Godsey, S.E.; Crosby, B.T.; Larsen, L.G.; Brooks, P.D.; Van Kirk, R.W.; Glenn, N.F. Regional Sensitivities of Seasonal Snowpack to Elevation, Aspect, and Vegetation Cover in Western North America: CLIMATE-PHYSIOGRAPHY EFFECTS ON SNOWPACK. *Water Resour. Res.* **2017**, *53*, 6908–6926, doi:10.1002/2016WR019374.

55. Lundquist, J.D.; Dickerson-Lange, S.E.; Lutz, J.A.; Cristea, N.C. Lower Forest Density Enhances Snow Retention in Regions with Warmer Winters: A Global Framework Developed from Plot-Scale Observations and Modeling: Forests and Snow Retention. *Water Resour. Res.* **2013**, *49*, 6356–6370, doi:10.1002/wrcr.20504.
56. Kostadinov, T.S.; Schumer, R.; Hausner, M.; Bormann, K.J.; Gaffney, R.; McGwire, K.; Painter, T.H.; Tyler, S.; Harpold, A.A. Watershed-Scale Mapping of Fractional Snow Cover under Conifer Forest Canopy Using Lidar. *Remote Sensing of Environment* **2019**, *222*, 34–49, doi:10.1016/j.rse.2018.11.037.
57. Houborg, R.; McCabe, M. High-Resolution NDVI from Planet's Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. *Remote Sensing* **2016**, *8*, 768, doi:10.3390/rs8090768.
58. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853, doi:10.1126/science.1244693.
59. McCabe, G.J. Relationships between Atmospheric Circulation and Snowpack in the Gunnison River Basin, Colorado. *Journal of Hydrology* **1994**, *157*, 157–175, doi:10.1016/0022-1694(94)90103-1.
60. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic Differentiation in PyTorch. 4.
61. Iglovikov, V.; Shvets, A. TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation. *arXiv:1801.05746 [cs]* **2018**.
62. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.-M.; Tucker, C.J.; Stenseth, N.Chr. Using the Satellite-Derived NDVI to Assess Ecological Responses to Environmental Change. *Trends in Ecology & Evolution* **2005**, *20*, 503–510, doi:10.1016/j.tree.2005.05.011.
63. Carlson, T.N.; Ripley, D.A. On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index. *Remote sensing of Environment* **1997**, *62*, 241–252.
64. Ye, S.; Pontius, R.G.; Rakshit, R. A Review of Accuracy Assessment for Object-Based Image Analysis: From per-Pixel to per-Polygon Approaches. *ISPRS Journal of Photogrammetry and Remote Sensing* **2018**, *141*, 137–147, doi:10.1016/j.isprsjprs.2018.04.002.
65. Whiteside, T.G.; Maier, S.W.; Boggs, G.S. Area-Based and Location-Based Validation of Classified Image Objects. *International Journal of Applied Earth Observation and Geoinformation* **2014**, *28*, 117–130, doi:10.1016/j.jag.2013.11.009.
66. Dobreva, I.D.; Klein, A.G. Fractional Snow Cover Mapping through Artificial Neural Network Analysis of MODIS Surface Reflectance. *Remote Sensing of Environment* **2011**, *115*, 3355–3366, doi:10.1016/j.rse.2011.07.018.
67. Kostadinov, T.S.; Lookingbill, T.R. Snow Cover Variability in a Forest Ecotone of the Oregon Cascades via MODIS Terra Products. *Remote Sensing of Environment* **2015**, *164*, 155–169, doi:10.1016/j.rse.2015.04.002.
68. Safa, H.; Krogh, S.; Greenberg, J.; Kostadinov, T.S.; Harpold, A.A. *Unraveling the Controls on Snow Disappearance in Montane Forests Using Multi-Site Lidar Observations*; Hydrology, 2020;
69. Guth, P.L.; Geoffroy, T.M. LiDAR Point Cloud and *ICESat-2* Evaluation of 1 Second Global Digital Elevation Models: Copernicus Wins. *Transactions in GIS* **2021**, *25*, 2245–2261, doi:10.1111/tgis.12825.

70. Klein, A.G.; Hall, D.K.; Riggs, G.A. Improving Snow Cover Mapping in Forests through the Use of a Canopy Reflectance Model. *Hydrol. Process.* **1998**, *12*, 1723–1744, doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1723::AID-HYP691>3.0.CO;2-2.
71. Helfricht, K.; Schöber, J.; Seiser, B.; Fischer, A.; Stötter, J.; Kuhn, M. Snow Accumulation of a High Alpine Catchment Derived from LiDAR Measurements. *Adv. Geosci.* **2012**, *32*, 31–39, doi:10.5194/adgeo-32-31-2012.
72. Grünewald, T.; Stötter, J.; Pomeroy, J.W.; Dadic, R.; Moreno Baños, I.; Marturià, J.; Spross, M.; Hopkinson, C.; Burlando, P.; Lehning, M. Statistical Modelling of the Snow Depth Distribution in Open Alpine Terrain. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 3005–3021, doi:10.5194/hess-17-3005-2013.
73. Huang, X.; Deng, J.; Wang, W.; Feng, Q.; Liang, T. Impact of Climate and Elevation on Snow Cover Using Integrated Remote Sensing Snow Products in Tibetan Plateau. *Remote Sensing of Environment* **2017**, *190*, 274–288, doi:10.1016/j.rse.2016.12.028.
74. Zhao, W.; Singh, R.; Joshi, T.; Sudjianto, A.; Nair, V.N. Self-Interpretable Convolutional Neural Networks for Text Classification. *arXiv preprint arXiv:2105.08589* **2021**.
75. Sun, Z.; Sandoval, L.; Crystal-Ornelas, R.; Mousavi, S.M.; Wang, J.; Lin, C.; Cristea, N.; Tong, D.; Carande, W.H.; Ma, X.; et al. A Review of Earth Artificial Intelligence. *Computers & Geosciences* **2022**, *159*, 105034, doi:10.1016/j.cageo.2022.105034.
76. John, A.; Ong, J.; Theobald, E.J.; Olden, J.D.; Tan, A.; HilleRisLambers, J. Detecting Montane Flowering Phenology with CubeSat Imagery. *Remote Sensing* **2020**, *12*, 2894, doi:10.3390/rs12182894.
77. Gitelson, A.; Merzlyak, M.N. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus Hippocastanum L. and Acer Platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of plant physiology* **1994**, *143*, 286–292.
78. Cooley, S.; Smith, L.; Stepan, L.; Mascaro, J. Tracking Dynamic Northern Surface Water Changes with High-Frequency Planet CubeSat Imagery. *Remote Sensing* **2017**, *9*, 1306, doi:10.3390/rs9121306.
79. Barrou Dumont, Z.; Gascoin, S.; Hagolle, O.; Ablain, M.; Jugier, R.; Salgues, G.; Marti, F.; Dupuis, A.; Dumont, M.; Morin, S. Brief Communication: Evaluation of the Snow Cover Detection in the Copernicus High Resolution Snow & Ice Monitoring Service. *The Cryosphere* **2021**, *15*, 4975–4980, doi:10.5194/tc-15-4975-2021.
80. Lv, Z.; Pomeroy, J.W. Detecting Intercepted Snow on Mountain Needleleaf Forest Canopies Using Satellite Remote Sensing. *Remote Sensing of Environment* **2019**, *231*, 111222, doi:10.1016/j.rse.2019.111222.
81. Parmesan, C.; Hanley, M.E. Plants and Climate Change: Complexities and Surprises. *Annals of Botany* **2015**, *116*, 849–864, doi:10.1093/aob/mcv169.
82. Panetta, A.M.; Stanton, M.L.; Harte, J. Climate Warming Drives Local Extinction: Evidence from Observation and Experimentation. *Sci. Adv.* **2018**, *4*, eaaq1819, doi:10.1126/sciadv.aaq1819.
83. Theobald, E.J.; Ettinger, A.K.; Burgess, H.K.; DeBey, L.B.; Schmidt, N.R.; Froehlich, H.E.; Wagner, C.; HilleRisLambers, J.; Tewksbury, J.; Harsch, M.A.; et al. Global Change and Local Solutions: Tapping the Unrealized Potential of Citizen Science for Biodiversity Research. *Biological Conservation* **2015**, *181*, 236–244, doi:10.1016/j.biocon.2014.10.021.

84. Schwartz, M.D.; Betancourt, J.L.; Weltzin, J.F. From Caprio's Lilacs to the USA National Phenology Network. *Frontiers in Ecology and the Environment* **2012**, *10*, 324–327, doi:10.1890/110281.
85. Kudo, G. Dynamics of Flowering Phenology of Alpine Plant Communities in Response to Temperature and Snowmelt Time: Analysis of a Nine-Year Phenological Record Collected by Citizen Volunteers. *Environmental and Experimental Botany* **2020**, *170*, 103843, doi:10.1016/j.envexpbot.2019.103843.
86. Dunne, J.A.; Harte, J.; Taylor, K.J. Subalpine Meadow Flowering Phenology Responses to Climate Change: Integrating Experimental and Gradient Methods. *Ecological Monographs* **2003**, *73*, 69–86, doi:10.1890/0012-9615(2003)073[0069:SMFPRT]2.0.CO;2.
87. Wolkovich, E.M.; Cook, B.I.; Allen, J.M.; Crimmins, T.M.; Betancourt, J.L.; Travers, S.E.; Pau, S.; Regetz, J.; Davies, T.J.; Kraft, N.J.B.; et al. Warming Experiments Underpredict Plant Phenological Responses to Climate Change. *Nature* **2012**, *485*, 494–497, doi:10.1038/nature11014.
88. Shores, C.R.; Mikle, N.; Graves, T.A. Mapping a Keystone Shrub Species, Huckleberry ( *Vaccinium Membranaceum* ), Using Seasonal Colour Change in the Rocky Mountains. *International Journal of Remote Sensing* **2019**, *40*, 5695–5715, doi:10.1080/01431161.2019.1580819.
89. Chen, B.; Jin, Y.; Brown, P. An Enhanced Bloom Index for Quantifying Floral Phenology Using Multi-Scale Remote Sensing Observations. *ISPRS Journal of Photogrammetry and Remote Sensing* **2019**, *156*, 108–120, doi:10.1016/j.isprsjprs.2019.08.006.
90. Fang, S.; Tang, W.; Peng, Y.; Gong, Y.; Dai, C.; Chai, R.; Liu, K. Remote Estimation of Vegetation Fraction and Flower Fraction in Oilseed Rape with Unmanned Aerial Vehicle Data. *Remote Sensing* **2016**, *8*, 416.
91. Horton, R.; Cano, E.; Bulanon, D.; Fallahi, E. Peach Flower Monitoring Using Aerial Multispectral Imaging. *J. Imaging* **2017**, *3*, 2, doi:10.3390/jimaging3010002.
92. Oliphant, A.J.; Thenkabail, P.S.; Teluguntla, P.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Yadav, K. Mapping Cropland Extent of Southeast and Northeast Asia Using Multi-Year Time-Series Landsat 30-m Data Using a Random Forest Classifier on the Google Earth Engine Cloud. *International Journal of Applied Earth Observation and Geoinformation* **2019**, *81*, 110–124, doi:10.1016/j.jag.2018.11.014.
93. Bolton, D.K.; Friedl, M.A. Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics. *Agricultural and Forest Meteorology* **2013**, *173*, 74–84.
94. Shen, M.; Chen, J.; Zhu, X.; Tang, Y. Yellow Flowers Can Decrease NDVI and EVI Values: Evidence from a Field Experiment in an Alpine Meadow. *Canadian Journal of Remote Sensing* **2009**, *35*, 99–106.
95. Herbei, M.V.; Sala, F.; others Use Landsat Image to Evaluate Vegetation Stage in Sunflower Crops. *AgroLife Scientific Journal* **2015**, *4*, 79–86.
96. John, A.; Ausmees, K.; Muenzen, K.; Kuhn, C.; Tan, A. SWEEP: Accelerating Scientific Research Through Scalable Serverless Workflows. In Proceedings of the Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing Companion - UCC '19 Companion; ACM Press: Auckland, New Zealand, 2019; pp. 43–50.

97. Huete, A.R. REMOTE SENSING FOR ENVIRONMENTAL MONITORING. In *Environmental Monitoring and Characterization*; Elsevier, 2004; pp. 183–206 ISBN 978-0-12-064477-3.
98. Jolliffe, I.T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Phil. Trans. R. Soc. A* **2016**, *374*, 20150202, doi:10.1098/rsta.2015.0202.
99. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY. *Ecology* **2007**, *88*, 2783–2792, doi:10.1890/07-0539.1.
100. Belgiu, M.; Csillik, O. Sentinel-2 Cropland Mapping Using Pixel-Based and Object-Based Time-Weighted Dynamic Time Warping Analysis. *Remote Sensing of Environment* **2018**, *204*, 509–523, doi:10.1016/j.rse.2017.10.005.
101. Feng, Q.; Liu, J.; Gong, J. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. *Remote sensing* **2015**, *7*, 1074–1094.
102. d'Andrimont, R.; Taymans, M.; Lemoine, G.; Ceglar, A.; Yordanov, M.; van der Velde, M. Detecting Flowering Phenology in Oil Seed Rape Parcels with Sentinel-1 and -2 Time Series. *Remote Sensing of Environment* **2020**, *239*, 111660, doi:10.1016/j.rse.2020.111660.
103. Liu, F.; Liao, Y.-Y.; Li, W.; Chen, J.-M.; Wang, Q.-F.; Motley, T.J. The Effect of Pollination on Resource Allocation among Sexual Reproduction, Clonal Reproduction, and Vegetative Growth in Sagittaria Potamogetifolia (Alismataceae). *Ecol Res* **2010**, *25*, 495–499, doi:10.1007/s11284-009-0679-1.
104. Liu, J.; Miller, J.R.; Haboudane, D.; Pattey, E.; Hochheim, K. Crop Fraction Estimation from Casi Hyperspectral Data Using Linear Spectral Unmixing and Vegetation Indices. *Canadian Journal of Remote Sensing* **2008**, *34*, S124–S138.
105. Ciganda, V.S.; Gitelson, A.A.; Schepers, J. How Deep Does a Remote Sensor Sense? Expression of Chlorophyll Content in a Maize Canopy. *Remote Sensing of Environment* **2012**, *126*, 240–247, doi:10.1016/j.rse.2012.08.019.
106. Pasqualotto, N.; Delegido, J.; Van Wittenberghe, S.; Rinaldi, M.; Moreno, J. Multi-Crop Green LAI Estimation with a New Simple Sentinel-2 LAI Index (SeLI). *Sensors* **2019**, *19*, 904, doi:10.3390/s19040904.
107. Zhu, X.; Liu, D. Improving Forest Aboveground Biomass Estimation Using Seasonal Landsat NDVI Time-Series. *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *102*, 222–231, doi:10.1016/j.isprsjprs.2014.08.014.
108. Gamon, J.A.; Field, C.B.; Goulden, M.L.; Griffin, K.L.; Hartley, A.E.; Joel, G.; Penuelas, J.; Valentini, R. Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecological Applications* **1995**, *5*, 28–41, doi:10.2307/1942049.
109. Tian, H.; Huang, N.; Niu, Z.; Qin, Y.; Pei, J.; Wang, J. Mapping Winter Crops in China with Multi-Source Satellite Imagery and Phenology-Based Algorithm. *Remote Sensing* **2019**, *11*, 820, doi:10.3390/rs11070820.
110. Liu, H.Q.; Huete, A. A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise. *IEEE transactions on geoscience and remote sensing* **1995**, *33*, 457–465.
111. Bentz, B.J.; Duncan, J.P.; Powell, J.A. Elevational Shifts in Thermal Suitability for Mountain Pine Beetle Population Growth in a Changing Climate. *Forestry* **2016**, *89*, 271–283, doi:10.1093/forestry/cpv054.

112. Okin, G.S.; Roberts, D.A.; Murray, B.; Okin, W.J. Practical Limits on Hyperspectral Vegetation Discrimination in Arid and Semiarid Environments. *Remote Sensing of Environment* **2001**, *77*, 212–225, doi:10.1016/S0034-4257(01)00207-3.
113. Bourgoin, C.; Blanc, L.; Bailly, J.-S.; Cornu, G.; Berenguer, E.; Oszwald, J.; Tritsch, I.; Laurent, F.; Hasan, A.; Sist, P.; et al. The Potential of Multisource Remote Sensing for Mapping the Biomass of a Degraded Amazonian Forest. *Forests* **2018**, *9*, 303, doi:10.3390/f9060303.
114. Leach, N.; Coops, N.C.; Obrknezev, N. Normalization Method for Multi-Sensor High Spatial and Temporal Resolution Satellite Imagery with Radiometric Inconsistencies. *Computers and Electronics in Agriculture* **2019**, *164*, 104893, doi:10.1016/j.compag.2019.104893.
115. Wicaksono, P.; Lazuardi, W. Assessment of PlanetScope Images for Benthic Habitat and Seagrass Species Mapping in a Complex Optically Shallow Water Environment. *International Journal of Remote Sensing* **2018**, *39*, 5739–5765, doi:10.1080/01431161.2018.1506951.
116. Guerini Filho, M.; Kuplich, T.M.; Quadros, F.L.F.D. Estimating Natural Grassland Biomass by Vegetation Indices Using Sentinel 2 Remote Sensing Data. *International Journal of Remote Sensing* **2020**, *41*, 2861–2876, doi:10.1080/01431161.2019.1697004.
117. Li, C.; Zhu, X.; Wei, Y.; Cao, S.; Guo, X.; Yu, X.; Chang, C. Estimating Apple Tree Canopy Chlorophyll Content Based on Sentinel-2A Remote Sensing Imaging. *Sci Rep* **2018**, *8*, 3756, doi:10.1038/s41598-018-21963-0.
118. van der Kooi, C.J.; Elzenga, J.T.M.; Staal, M.; Stavenga, D.G. How to Colour a Flower: On the Optical Principles of Flower Coloration. *Proc. R. Soc. B* **2016**, *283*, 20160429, doi:10.1098/rspb.2016.0429.
119. Chavez, P.S.; others Image-Based Atmospheric Corrections-Revisited and Improved. *Photogrammetric engineering and remote sensing* **1996**, *62*, 1025–1035.
120. Chavez Jr, P.S. An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data. *Remote sensing of environment* **1988**, *24*, 459–479.
121. Breckheimer, I.K.; Theobald, E.J.; Cristea, N.C.; Wilson, A.K.; Lundquist, J.D.; Rochefort, R.M.; HilleRisLambers, J. Crowd-sourced Data Reveal Social–Ecological Mismatches in Phenology Driven by Climate. *Front Ecol Environ* **2020**, *18*, 76–82, doi:10.1002/fee.2142.
122. Hille Ris Lambers, J.; Cannistra, A.F.; John, A.; Lia, E.; Manzanedo, R.D.; Sethi, M.; Sevigny, J.; Theobald, E.J.; Waugh, J.K. Climate Change Impacts on Natural Icons: Do Phenological Shifts Threaten the Relationship between Peak Wildflowers and Visitor Satisfaction? *Climate Change Ecology* **2021**, *2*, 100008, doi:10.1016/j.ecochg.2021.100008.
123. Klein, G.; Rebetez, M.; Rixen, C.; Vitasse, Y. Unchanged Risk of Frost Exposure for Subalpine and Alpine Plants after Snowmelt in Switzerland despite Climate Warming. *Int J Biometeorol* **2018**, *62*, 1755–1762, doi:10.1007/s00484-018-1578-3.
124. Vorkauf, M.; Kahmen, A.; Körner, C.; Hiltbrunner, E. Flowering Phenology in Alpine Grassland Strongly Responds to Shifts in Snowmelt but Weakly to Summer Drought. *Alp Botany* **2021**, *131*, 73–88, doi:10.1007/s00035-021-00252-z.
125. Moyroud, E.; Glover, B.J. The Evolution of Diverse Floral Morphologies. *Current Biology* **2017**, *27*, R941–R951, doi:10.1016/j.cub.2017.06.053.

126. Dias, P.A.; Tabb, A.; Medeiros, H. Apple Flower Detection Using Deep Convolutional Networks. *Computers in Industry* **2018**, *99*, 17–28, doi:10.1016/j.compind.2018.03.010.
127. Wang, X. (Annie); Tang, J.; Whitty, M. DeepPhenology: Estimation of Apple Flower Phenology Distributions Based on Deep Learning. *Computers and Electronics in Agriculture* **2021**, *185*, 106123, doi:10.1016/j.compag.2021.106123.
128. Osherov, E.; Lindenbaum, M. Increasing CNN Robustness to Occlusions by Reducing Filter Support. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV); IEEE: Venice, October 2017; pp. 550–561.
129. Abdolali, F.; Kapur, J.; Jaremko, J.L.; Noga, M.; Hareendranathan, A.R.; Punithakumar, K. Automated Thyroid Nodule Detection from Ultrasound Imaging Using Deep Convolutional Neural Networks. *Computers in Biology and Medicine* **2020**, *122*, 103871, doi:10.1016/j.compbiomed.2020.103871.
130. Hao, Z.; Lin, L.; Post, C.J.; Mikhailova, E.A.; Li, M.; Chen, Y.; Yu, K.; Liu, J. Automated Tree-Crown and Height Detection in a Young Forest Plantation Using Mask Region-Based Convolutional Neural Network (Mask R-CNN). *ISPRS Journal of Photogrammetry and Remote Sensing* **2021**, *178*, 112–123, doi:10.1016/j.isprsjprs.2021.06.003.
131. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv:1506.02640 [cs]* **2016**.
132. Iglovikov, V.; Shvets, A. TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation. *arXiv:1801.05746 [cs]* **2018**.
133. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-Cnn. In Proceedings of the Proceedings of the IEEE international conference on computer vision; 2017; pp. 2961–2969.
134. Jia, W.; Tian, Y.; Luo, R.; Zhang, Z.; Lian, J.; Zheng, Y. Detection and Segmentation of Overlapped Fruits Based on Optimized Mask R-CNN Application in Apple Harvesting Robot. *Computers and Electronics in Agriculture* **2020**, *172*, 105380, doi:10.1016/j.compag.2020.105380.
135. Machefer, M.; Lemarchand, F.; Bonnefond, V.; Hitchins, A.; Sidiropoulos, P. Mask R-CNN Refitting Strategy for Plant Counting and Sizing in UAV Imagery. *Remote Sensing* **2020**, *12*, 3015, doi:10.3390/rs12183015.
136. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical image computing and computer-assisted intervention; Springer, 2015; pp. 234–241.
137. Manzanedo, R.D.; John, A.; Sethi, M.L.; Theobald, E.J.; Brosi, B.; Jenkins, J.; Kloss-Schmidt, A.; Lia, E.; Schiffer, A.; Sevigny, J.; et al. MeadoWatch: A Long-Term Community-Science Database of Wildflower Phenology in Mount Rainier National Park. *Sci Data* **2022**, *9*, 151, doi:10.1038/s41597-022-01206-8.
138. Russell, B.C.; Torralba, A.; Murphy, K.P.; Freeman, W.T. LabelMe: A Database and Web-Based Tool for Image Annotation. *Int J Comput Vis* **2008**, *77*, 157–173, doi:10.1007/s11263-007-0090-8.
139. Wu, Y.; Kirillov, A.; Massa, F.; Lo, W.-Y.; Girshick, R. Detectron2 2019.
140. Kalantar, A.; Edan, Y.; Gur, A.; Klapp, I. A Deep Learning System for Single and Overall Weight Estimation of Melons Using Unmanned Aerial Vehicle Images. *Computers and Electronics in Agriculture* **2020**, *178*, 105748, doi:10.1016/j.compag.2020.105748.

141. Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft Coco: Common Objects in Context. In Proceedings of the European conference on computer vision; Springer, 2014; pp. 740–755.
142. KUDO, G.; Suzuki, S. Flowering Phenology of Alpine Plant Communities along Gradient of Snowmelt Timing. *Polar Bioscience* **1999**, *12*, 100–113.
143. Prevéy, J.S. Climate Change: Flowering Time May Be Shifting in Surprising Ways. *Current Biology* **2020**, *30*, R112–R114, doi:10.1016/j.cub.2019.12.009.
144. Yu, Y.; Zhang, K.; Yang, L.; Zhang, D. Fruit Detection for Strawberry Harvesting Robot in Non-Structural Environment Based on Mask-RCNN. *Computers and Electronics in Agriculture* **2019**, *163*, 104846, doi:10.1016/j.compag.2019.06.001.
145. Suh, S.; Park, Y.; Ko, K.; Yang, S.; Ahn, J.; Shin, J.-K.; Kim, S. Weighted Mask R-CNN for Improving Adjacent Boundary Segmentation. *Journal of Sensors* **2021**, *2021*, 1–8, doi:10.1155/2021/8872947.
146. Prasetyo, E.; Suciati, N.; Fatichah, C. A Comparison of YOLO and Mask R-CNN for Segmenting Head and Tail of Fish. In Proceedings of the 2020 4th International Conference on Informatics and Computational Sciences (ICICoS); 2020; pp. 1–6.
147. Sumit, S.S.; Watada, J.; Roy, A.; Rambli, D. In Object Detection Deep Learning Methods, YOLO Shows Supremum to Mask R-CNN. *J. Phys.: Conf. Ser.* **2020**, *1529*, 042086, doi:10.1088/1742-6596/1529/4/042086.
148. Tian, Y.; Yang, G.; Wang, Z.; Li, E.; Liang, Z. Instance Segmentation of Apple Flowers Using the Improved Mask R–CNN Model. *Biosystems Engineering* **2020**, *193*, 264–278, doi:10.1016/j.biosystemseng.2020.03.008.
149. Fabbro, T.; Körner, C. Altitudinal Differences in Flower Traits and Reproductive Allocation. *Flora - Morphology, Distribution, Functional Ecology of Plants* **2004**, *199*, 70–81, doi:10.1078/0367-2530-00128.
150. Sargent, R.D.; Ackerly, D.D. Plant–Pollinator Interactions and the Assembly of Plant Communities. *Trends in Ecology & Evolution* **2008**, *23*, 123–130, doi:10.1016/j.tree.2007.11.003.
151. Wilson, A.; Bacher, K.; Breckheimer, I.; Lundquist, J.; Rochefort, R.; Theobald, E.; Whiteaker, L.; HilleRisLambers, J. Monitoring Wildflower Phenology Using Traditional Science, Citizen Science, and Crowd Sourcing. *Park Science* **2017**, *33*, 17–26.
152. Taheri, S.; Naimi, B.; Rahbek, C.; Araújo, M.B. Improvements in Reports of Species Redistribution under Climate Change Are Required. *Sci. Adv.* **2021**, *7*, eabe1110, doi:10.1126/sciadv.abe1110.
153. Elliott, T.L.; Davies, T.J. A Comparison of Phylogenetic and Species Beta Diversity Measures Describing Vegetation Assemblages along an Elevation Gradient. *J Veg Sci* **2019**, *30*, 98–107, doi:10.1111/jvs.12700.
154. Nugent, J. Citizen Science: Birds, Binoculars, and Biodiversity. *Sci. Scope* **2018**, *041*, doi:10.2505/4/ss18_041_05_16.
155. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149, doi:10.1109/TPAMI.2016.2577031.
156. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv:1804.02767 [cs]* **2018**.

157. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE: Honolulu, HI, July 2017; pp. 6517–6525.