

GEOSPATIAL ANALYSIS OF THE EFFECTS OF CANOPY INSOLATION  
PARTITIONING ON BIODIVERSITY IN A TEMPERATE MONTANE FOREST

A Thesis submitted to the faculty of  
San Francisco State University  
In partial fulfillment of  
the requirements for  
the Degree

Master of Science

In

Geographic Information Science

by

Paul Christian Dunn

San Francisco, California

August 2019

Copyright by  
Paul Christian Dunn  
2019

## CERTIFICATION OF APPROVAL

I certify that I have read Geospatial Analysis of the Effects of Canopy Insolation Partitioning on Biodiversity in a Temperate Montane Forest by Paul Christian Dunn, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Science in Geographic Information Science at San Francisco State University.

---

Leonhard Blesius, Ph.D.  
Associate Professor

---

Andrew Oliphant, Ph.D.  
Professor

---

Chad Hanson, Ph.D.  
Forest Ecologist

GEOSPATIAL ANALYSIS OF THE EFFECTS OF CANOPY INSOLATION  
PARTITIONING ON BIODIVERSITY IN A TEMPERATE MONTANE FOREST

Paul Christian Dunn  
San Francisco, California  
2019

Incident solar radiation (insolation) passing through the forest canopy to the ground surface is either absorbed or scattered. This phenomenon is known as attenuation and is measured in forest ecology using the extinction coefficient ( $K$ ). The effect of  $K$  on understory photosynthetically available radiation (PAR) and microclimate may be associated with plant species diversity, as distinct species and communities have unique habitat requirements. The objective of his study is to model insolation and canopy structure to observe effects of predictors representative of  $K$  on understory plant biodiversity using remotely sensed and botanical field data.

We used two taxonomic diversity indexes (Menhinick's and Simpson) to describe the surveyed plant community in a natural temperate montane forest, modeling the index values at the plot level as response variables. Independent variables included localized area incident solar radiation estimated using a solar model, LiDAR derived canopy height model, effective leaf area index (LAI) estimates derived from multi-spectral imagery and canopy strata metrics derived from LiDAR point cloud data. Considering the impact of atmospheric components above the canopy layer and an assumption that incident short-wave solar radiation to Earth's vegetated surfaces is *primarily* absorbed in the canopy layer, we used a multiple linear model to predict canopy metrics controlling the sub-canopy surface radiant flux to develop the hypothesis that 1.) canopy structural variability is associated with the biodiversity of stand plant species through habitat partitioning and, 2.) a prediction model can be developed to validate this relationship spatially. The available data indicated many and varied correlations between predictor and response variables as well as a statistically valid linear model comprising the canopy relief, the texture, and vegetation density with understory plant diversity. When analyzed for spatial autocorrelation, the predicted biodiversity data exhibited non-random spatial continuity.

I certify that the Abstract is a correct representation of the content of this thesis.

---

Chair, Thesis Committee

Date

## PREFACE AND ACKNOWLEDGEMENTS

The author would like to express gratitude to the thesis committee, Drs. Leonhard Blesius, Andrew Oliphant and Chad Hanson for review and insights on the many and complex biogeographical, geospatial and climatological variables that entered into the conceptual framework for this study. Dr. Blesius granted me several years of his time and trained me in key areas of remote sensing theory and practice. Thanks is also due to Dr. John Gallo for mentoring me in geospatial and ecological science while acting as Senior Landscape Ecologist at the TWS-Center for Landscape Analysis, Diane Montoya, research geomorphologist at the USFS Redwood Sciences Laboratory in Arcata, CA for reviewing the original concepts, Anna Studwell and Quinten Clark at the Institute for Geographic Information Science (IGISc) at San Francisco State University's College of Science and Engineering for software support and technical assistance, and especially to Mark Schutz and Kevin Physioc for their fieldwork support, without which the project would remain entirely incomplete. Finally, I would like to dedicate this thesis to the loving memory my mother, Paula Atkinson, and to my father, Jack Opie, who both provided me with a rich and challenging learning environment as the predicate for lifelong curiosity and the motivation for scientific inquiry.

## TABLE OF CONTENTS

List of Tables .....	viii
List of Figures .....	ix
List of Appendices .....	x
Introduction.....	1
Study Area .....	3
Regional Geological and Biogeographic History .....	5
Study Area Soil Taxonomy.....	7
Regional Climate and Meteorology .....	8
Regional Biogeography and Forest Species .....	9
Area Solar Radiation, Canopy Structure and Insolation Partitioning.....	13
Canopy Structure Metrics .....	14
Measures of Biodiversity .....	20
Data.....	22
Multi-spectral Imagery Data Specification & Processing .....	22
LiDAR Data Acquisition & Specification .....	23
Data Processing Summary .....	25
Methods.....	26
Input Data Analysis.....	27
Study Area Predictor Variable Distributions .....	29
Sample Plot Selection .....	32
Biodiversity Response Variables .....	34
Multiple Linear Regression and Prediction Modeling.....	35
Results.....	37

Discussion.....	52
References.....	55
Appendices.....	61

## LIST OF TABLES

Table	Page
1. Study Area Soil Taxonomy .....	8
2. Vascular Plants, Bryophytes and Lichens of the Montane KMC Forest.....	12
3. Global Short-wave Radiation Balance .....	13
4. WorldView2 Multispectral Image Bands.....	23
5. Absolute Accuracy of LiDAR Ground Control Points .....	24
6. Vertical Relative LiDAR Accuracy .....	24
7. Plot Survey Data.....	34
8. Global Moran's I and General G Results for Predicted MDI .....	50

## LIST OF FIGURES

Figures	Page
1. Study Area Map.....	5
2. Photo of Study Area Topography.....	7
3. Extruded LiDAR Digital Image of Study Area Topography .....	7
4. Study Area Soil Map .....	8
5. Photos of Study Area Biota.....	11
6. Study Area LiDAR Point-cloud Tile.....	25
7. Analysis Extent Data Processing Steps .....	26
8. Study Area NDVI Image Map.....	29
9. Study Area Predictor Variable Distributions.....	30
10. Study Area ASR, LAI, CH and TVD Collinearity Analysis.....	31
11. Study Area Plot Map .....	33
12. Study Area Response Variable Distributions.....	35
13. Study Area ASR, R, A, MDI and SDI Collinearity Analysis .....	37
14. Study Area R Collinearity Analysis .....	38
15. Study Area A Collinearity Analysis.....	39
16. Study Area MDI Collinearity Analysis .....	40
17. Study Area SDI Collinearity Analysis .....	41
18. Model Cross-validation Analysis .....	42
19. Linear Model 0 Summary Statistics .....	43
20. Linear Model 1 Summary Statistics .....	43
21. Linear Model 2 Summary Statistics .....	44
22. Linear Model 1 and 2 ANOVA Analysis.....	44
23. Linear Model 3 Summary Statistics .....	45
24. Linear Model 3 Residuals vs. Fitted Plot .....	46
25. Linear Model 3 QQ Plot.....	46
26. Linear Model 3 Predictive Scatterplot.....	47
27. Study Area MDI Prediction Map .....	48
28. Study Area Adjusted MDI Prediction Map.....	49
29. Study Area MDI Clustering Map.....	52

## LIST OF APPENDICES

Appendix	Page
1. Appendix 1 .....	52
2. Appendix 2 .....	53
3. Appendix 3 .....	54
4. Appendix 4 .....	69

## Introduction

The disciplines of spatial ecology and biogeography maintain similar theories that attempt to explain the variation of species distribution across spatial scales and latitudinal gradients. This spatial variability and its link to ecological and biogeochemical processes are considered to be fundamental to biological inquiry and relating phenomena across scales a central research problem in biology (Levin 1992). The spatial analysis of biota often focuses on either species distribution or species biodiversity within the context of community, landscape, ecosystem and genetic systems associated with spatial scales such as species populations. For plant species, the discipline of phytogeography combines botany and geography to investigate the spatial distributions of plant species and their communities. An historical example is given in Alexander von Humboldt and Aimé Jacques Bonpland's "*Essai sur la géographie des plantes*", published in 1807. In this work they considered the variation of several ecological gradients to put forth a theory of the geographical repartition of species, visualizing these concepts in portraiture – the "*Tableau physique des Andes et pays voisins*" a prominent work among them – and contributing significantly to the formation of modern biogeography (von Humboldt & Bonpland 1807).

Biodiversity measures describe species and trait richness and evenness on different bases and scales including taxonomic or genetic (trait based) and habitat or community and ecosystem. Alpha biodiversity ( $\alpha$ -diversity), or species diversity in habitats at a local scale, is influenced both by the number of types of habitat and ecological processes (Dufour et al. 2006). It has been shown that the rates of such processes are affected by the spatial configuration of the environment, and a hypothesis will be made that  $\alpha$ -diversity is influenced by two aspects of environmental heterogeneity: the range of environmental conditions (i.e. environmental variability) influencing the number of types of habitats available and the spatial configuration of those habitats (Dufour et al. 2006). Vegetation biodiversity in forest ecosystems has been positively correlated with productivity and resiliency in forest stands due to increased spatial, temporal and biogeochemical efficiencies in site utilization (Schulze 2005, Zhang et al. 2012, Ishii et al. 2004). In addition to higher productivity, the resultant plant communities tend to be less vulnerable to pathogens, wildfire and wind related disturbances and better adapt to environmental change (Schneider et al. 2017).

The spatial variability of biodiversity on larger scales has been characterized by biogeographers with respect to latitudinal gradients. Meta-studies indicate that mechanisms such as solar energy, climate and area-specific processes are likely to

contribute to species diversity on many spatial scales (Hillenbrand 2004) and within the vertical and horizontal structural variation in forest landscapes. The asymmetric competition for light in the canopy is noted in another meta-study as a key negative influence on overall productivity (Ali 2019).

A forest's structure is a landscape scale terrestrial ecosystem including canopy flora, soil type and depth, subsurface biota and hydrology. The canopy is of considerable interest to researchers because of its functional interface with the atmosphere with respect to carbon, water and energy exchange and being the site of primary production (Pan et al. 2013). On a global scale, forest canopies support approximately 40% of extant species, 10% of which are predicted to be canopy specialists (Ozanne et al. 2003). Attributable to the complex three-dimensional structure of the canopy affording niche diversification and vertical stratification, approximately 10% of all vascular plants are epiphytic canopy species (Ozanne et al. 2003).

Landscape scale studies have observed stand structure as a biodiversity indicator from ecosystem data repositories with no adjoining fieldwork. Gao et al. (2014) observed that mature stands with a stratified canopy had the highest plant species diversity of the 26 stand structure types and across the nine soil classes in the study, in particular stands comprising mixed conifer and broadleaved species with a semi-open canopy, whereas younger single-layered stands had consistently low species diversity. Age of canopy trees was closely associated with taxonomic diversity, followed by canopy stratification, tree species composition and canopy coverage (Gao et al. 2014).

Chronosequential field studies designed to observe forest succession related  $\alpha$ -diversity in a temperate mixed-conifer forest region indicate that post-disturbance taxonomic diversity trends upward until canopy closure. It then decreases for up to several decades before increasing again as canopy structure variability increases (Schoonmaker & McKee 1988). Notably, although late seral species returned in significant percentages in as soon as five years following the disturbance, two mycotrophs noted in the study area, reliant on the subsurface mycorrhizal biota, were lost to the survey habitat due to mechanical disturbance – these species potentially required more complex dispersal mechanism to recolonize the sites after an industrial harvest (Schoonmaker & McKee 1988). In contrast to more or less uniform mechanical stand replacement disturbances, Donato et al. (2009) note that natural and compound disturbance regimes, such as mixed-severity fire, create alternate successional pathways and distinct early seral communities contributing to canopy heterogeneity.

Due to its three-dimensional distribution of leaves, branches and stems comprising their topology within the overall canopy, as well as its physical relationship to incoming solar radiation, canopy structure is a source of habitat niche partitioning for plant species (Pan et al. 2013). Properties of this structure can be described by metrics such as stand height, density, distribution and volume – each as a proxy for the more complex and unmeasurable distribution of forest canopy structure itself (Leiterer et al. 2015).

Canopy structure can be considered in terms of both vertical components, such as strata, and horizontal ones, such as age or disturbance patches, cover and fragments. Stands undisturbed by landscape level events such as fire or major windfall tend to develop an uneven-aged stand structure as natural disturbances generally provide open gaps in the canopy for younger trees (Pan et al. 2011). Depending on the scale of the canopy measurement, such as is the case with optical methods analyzing multi-spectral imagery or LiDAR point-cloud data gridded to a 30-meter raster resolution, elements of both can be captured and modeled at once.

The objective of this study is to develop an approach for the accurate geospatial analysis of the impact of canopy strata diversity on the radiant flux and extinction on biodiversity at a stand-level scale. We used area solar radiation models and canopy metrics derived from multi-spectral imagery and light detection and ranging (LiDAR) systems to estimate radiant energy and canopy structural properties that act to absorb or scatter the radiant flux as it moves downward from the canopy surface to the forest floor.

This derived data, and field data from randomly selected biodiversity plots sampled from an unmodified natural temperate montane conifer forest, was used to develop multivariate regression model to predict the spatial distribution of plant species  $\alpha$ -diversity. If the results are adequate to initiate a more comprehensive study, localized data sets similar to those used for stand volume and site quality could be developed from data acquired over time for forest management and ecosystem services prescriptions.

## **Study Area**

The study area is located in the center of the Klamath Ecoregion of Northern California and Southern Oregon, an area studied extensively for its geological antiquity and diversity of plant species. Geographically, the region is bounded on the north and south by lower elevation coast range mountains and interior valleys, on the west by the Pacific Ocean, and on the east by montane valleys and desert plateaus. Due in part to its topographic variability and an abundance of alpine water sources, the landscape contains

a conifer biodiversity ‘hot-spot’ within a region that represents a global maxima comprising 30 native and several endemic species (Olson 2016).

The analysis extent is bound by an arid east-slope canyon at the base of Russian Peak in the Upper Sugar Creek Watershed. The analysis extent and the study area are two distinct geographic areas. The first is the topographic horizon captured in the bare earth digital elevation model (DEM) produced from the LiDAR products. The study area for plot selection is defined by a Klamath Mixed Conifer (KMC) plant community, comprising plots that were within one standard deviation from the mean canopy height derived at a 30 meter resolution.

The study area is an approximately 18.5 hectare area located at 507090 E, 4572336 N (Universal Transverse Mercator (UTM) Zone 10, NAD 1983) as shown in Figure 1. The elevation was between 1500 meters and 1700 meters height above ellipsoid (HAE), with the 2019 plots samples observed primarily near 1500 meters HAE. Sample plots were therefore selected from within a relatively uniform climate, soil type and vegetation profile, although soil water content varied considerably due to varied water sources, topography and organic structures such as down woody debris.

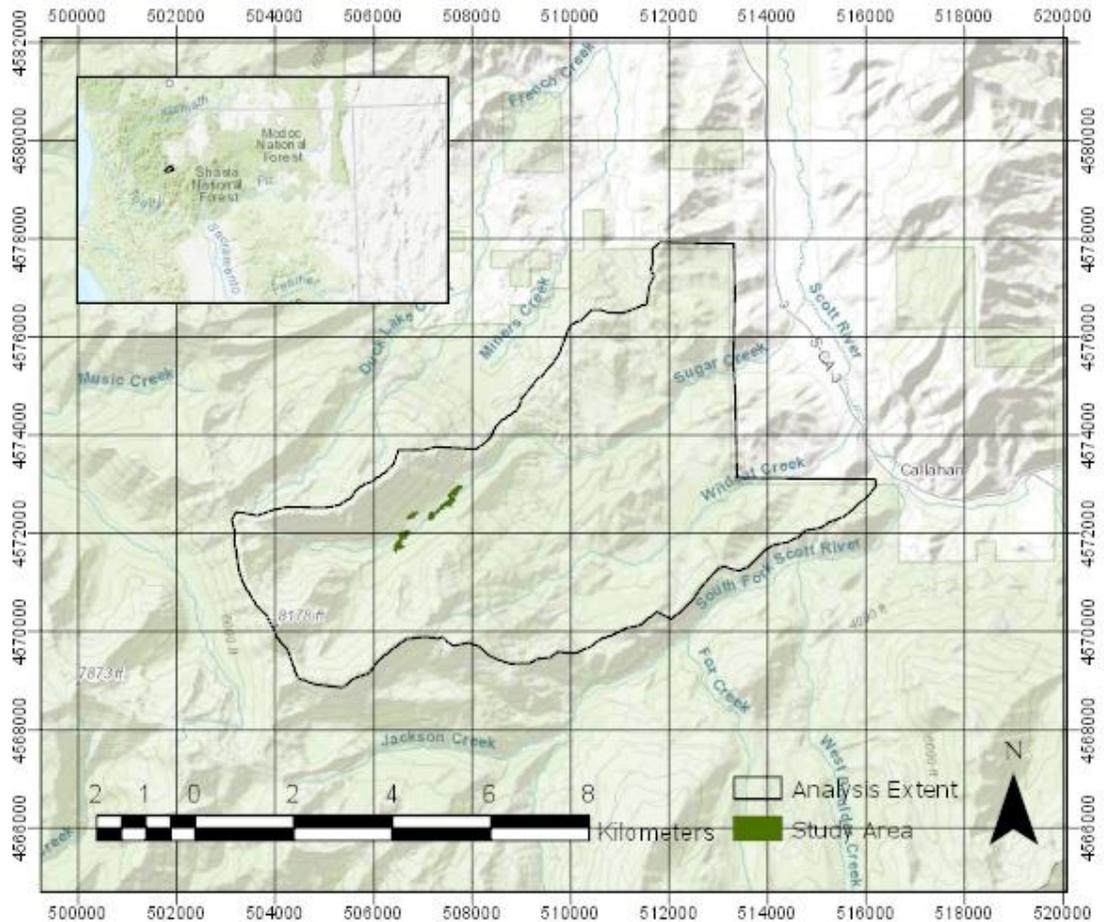


Figure 1: The study area comprises an 18.4896 hectare area of montane Klamath Mixed Conifer stands. The analysis extent includes the Upper Sugar Creek Watershed. Source: Existing Vegetation-CALVEG 2007.

**Regional Geological and Biogeographic History**

The geology of the Klamath Ecoregion is complex and plays a key role in its biogeographic formation. The overarching geological framework originated as island arcs and continental fragments in the Pacific Ocean consisting of the rifted fragments of pre-existing continents and volcanic island masses created over subduction zones. These contain rocks believed to be 500 million years old that date to the early Paleozoic Era. A succession of eight island terranes moved eastward on the ancient Farallon plate and collided with the North American plate an estimated 260 to 130 million years ago

(Skinner et al. 2006). Each accretion left a terrane of rock of a single age. During the accretion, subduction of the plate metamorphosed the overlying rock and produced magma which intruded the overlying rock as plutons. Serpentinite, produced by the metamorphism of basaltic oceanic rocks, and intrusive rocks of gabbroic to granodiorite composition are common rocks within the Klamath Ecoregion terranes (Skinner et al. 2006).

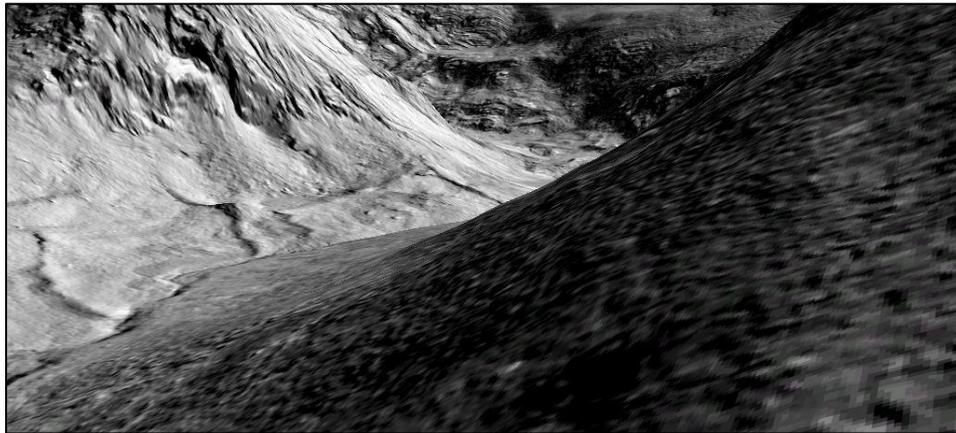
The Wisconsin glaciation that began approximately 80,000 years ago and lasted until the Holocene was the last major geological event shaping the landforms proximate to the analysis extent. Glacial ice collected during the Wisconsin in the higher elevations, especially in the Marble Mountains, Trinity Alps, and Trinity Mountains (Sawyer 2006). Specific regional influences of the Wisconsin glacial period on landform and soil development involved a series of ice-free times (interstadials), and ice advance (stadial). The last major ice advance, the Late Wisconsin, occurred about 22,000 years ago. Since then, the climate has warmed in spurts with several reversals, during which the ice advanced and then retreated. During the interstadials, clay, silt, sand, gravel, and boulders (collectively called till), eroded from higher elevations, accumulated in mounds and ridges (moraines), as the ice melted. The last glacial in the region ended approximately 170 years ago after the end of the Little Ice Age (Sawyer 2006).

The region's geological position and the topographic variability provided by its landforms have provided refugia that have allowed several endemic and relic species to persist which were once dominant in forests of previous geological epochs and during periods of alternating climate conditions in the Northern Hemisphere. Due to its proximity to the Cascade and Sierra Nevada ranges to the north and south, and to the Oregon and California coastal ranges to the west, the region's conifer biodiversity can be traced to its biogeographic locale. This is believed to be related to the convergence of these regions in the Klamath, relic, endemic and regionally dominant species coexist in a diverse series of xeric, hydric and mesic habitats among abrupt alternates in elevation and climate (Keeler-Wolf 1990). Although regional vegetation diversity is higher along the inland-continental gradient, especially in temperate woodlands with a highly developed herb strata (Skinner et al. 2006), the taxonomic biodiversity present in the study area itself is thought to be the unique result of its topographic-climate variability and biogeographic history (Keeler-Wolf 1990).

The ruggedness and topographic variability of the study areas landscape is evident in a photograph taken from the steep ridge above the study area in Figures 2. Figure 3 depicts the study area landforms taken from a 3D rendering of the LiDAR data (last returns).



**Figure 2: Ridgeline below Russian Peak looking west above study area in the Upper Sugar Creek Watershed  
Source: Kauffmann 2014.**



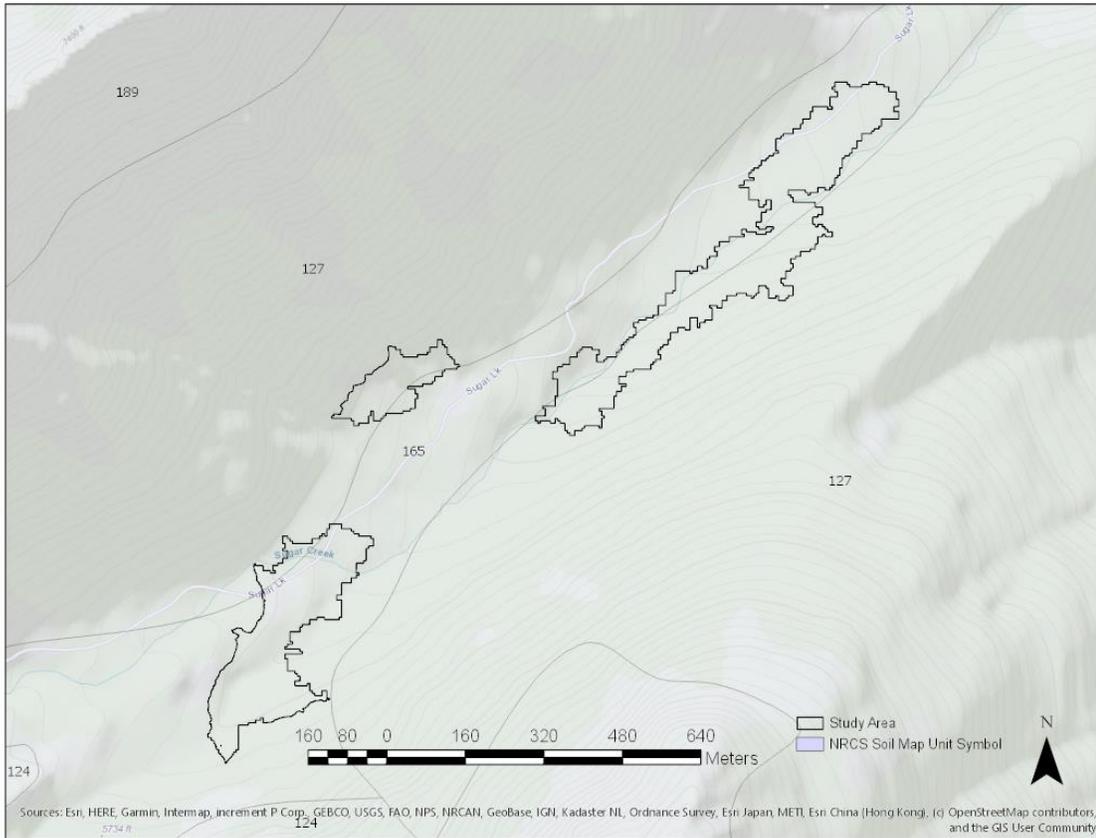
**Figure 3: Study area viewed from 506223E 4572098N using an extruded LiDAR based Digital Elevation Model (DEM). The 'bare earth' model with hill-shade rendering provides a detailed view of the site's landforms.**

### **Study Area Soil Taxonomy**

Soil profiles in the study area include those composed of the Nanny and Gerle families as shown in Table 1 (USDA 1982) and labeled with the map unit symbol to indicate overlap with the study area (Figure 4).

**Table 1: Soil Taxonomy of the Upper Sugar Creek Watershed (Source: USDA)**

Soil Name & Class	Subgroup	Order (Taxon)	Map Unit Symbol
Nanny, loamy-skeletal,mixed, frigid	Typic Xerumbrepts	Inceptisols	165
Gerle, coarse-loamy,mixed, frigid	Typic Xerumbrepts	Inceptisols	127



**Figure 4: Study area soil map by map unit (Source: NRCS).**

Loamy inceptisols are the dominant soil taxon in the study extent. These soils commonly occur on landscapes that are relatively active, such as mountain slopes, where erosional processes expose unweathered materials formed in late-Pleistocene glacial drift.

**Regional Climate and Meteorology**

Significant shifts in climate over time make the Klamath Ecoregion both an intersection and a transition zone for several major regional scale biotas, specifically the Great Basin, the Oregon Coast Range, the Cascades Range, the Sierra Nevada, the California Central

Valley, and Coastal [Mountain] Province of Northern California (Olson 2016). Evidence indicates that the regional climates of the Quaternary period, including Pleistocene and the Holocene, were generally colder than were those of the Tertiary.

The ecoregion's modern climate is characterized by warm dry summers and wet winters with a strong precipitation gradient decreasing as it tracks from the Pacific coast to the eastern mountain interior. The temperature gradient from the coast to the interior basins increases in range as it tracks eastward from marine to continental. From west to east along the climate gradient away from the moderating marine layer, conifer species tend to occupy increasingly higher elevation sites, from sea level in the west, to montane to subalpine zones in the east.

The region's weather patterns are a key source of canopy disturbance in its forest ecosystems. These include the Pacific High–Post-Frontal, Pacific High–Pre-Frontal and Subtropical High Aloft patterns. Post-Frontal conditions are defined by high pressure following the passage of a cold front causing strong winds from the north and/or northeast. Pre-Frontal conditions are defined by strong, southwesterly or westerly winds created by dry southern tail-end of a rapidly moving cold front. In this case the strong winds are the relevant physical process. Subtropical high conditions are defined by descending air from high pressure, increasing temperatures, decreasing humidity and are related to strong temperature inversions (Skinner et al. 2006). These conditions are related to high-velocity winds and low-humidity and represent a key physical component of natural forest dynamics and stand succession including fire and windfall (Skinner et al. 2006).

The Klamath ecoregion is described as having had a mixed-severity fire regime, with a fire return interval of about 15 years in lower montane conifer forests (a term used when observing fire frequency at the scale of a stand or relatively small landscape area). The fire rotation interval presents a more nuanced view of the local study area fire regime, as it describes the fire cycle over the larger scale landscape with variable spatio-temporal frequency and intensity. Notably, the study area has missed some intervals in its natural fire regime due to suppression (DeSiervo et al. 2016).

### **Regional Biogeography and Forest Species**

The Klamath Ecoregion is characterized by “complex biogeographic patterns, high endemism, and unusual community assemblages” (Olson 2016). It is one of seven IUCN (International Union for Conservation of Nature) Areas of Global Botanical Significance in North America and is a proposed World Heritage Site and United Nations Educational,

Scientific and Cultural Organization (UNESCO) Biosphere Reserve. It contains one of the four richest temperate coniferous forests in the world along with the Southeastern conifer forests of North America and those of the Primorye region of the Russian Far East and Sichuan region of China (Olson 2016).

The region was relatively unglaciated during the late-Pleistocene, when it served as a refugia for northern plant species (Thorson et al. 2003). Sawyer & Thornburgh (1974) note paleobotanical studies that indicate that the Klamath Ecoregion maintains “forests most nearly equivalent to the western North American Arcto-Tertiary forests”. The region has not undergone any significant geological alterations since the Miocene, which is significant with respect to the magnitude of ecosystem disturbance from glaciation relative to the Wisconsin period. Late-Pleistocene glaciation was localized where it did occur and had a varied impact on the region’s conifer species, destroying many populations but allowing others to persist. The diverse topography played a role in preserving conifer species in some cases, such as sub-alpine fir which survived as krummholz (stunted, windblown trees) on cliffs above ice margins in the Russian Peak area, and subsequently repopulated hydric habitats as glaciers retreated. It is believed that Engelmann spruce survived along streams below the glaciers, and relatively low intensity fire regimes in the eastern montane allowed weeping spruce to persist (Sawyer & Thornburgh 1974).

Larger forest reserves are generally found in the highest elevations of the region, with few significant areas of lower elevation habitat remaining undisturbed. The Dillon Creek watershed on the middle Klamath River reach is one of the last remaining unfragmented lowland forests of old growth Klamath Mixed Conifer in the region (Olson 2016). The montane Klamath Mixed Conifer (KMC) forest type that defines the study extent in the Upper Sugar Creek Watershed is a regionally unique assemblage composed of tall, dense to moderately open, conifer forests with patches of broad-leaved evergreen and deciduous low trees and shrubs typical of the assemblage (Küchler 1977). It is dominated by evergreen conifers up to 60 meters in height and a rich shrub and herbaceous layer on undisturbed mesic sites. The overstory layer is characterized by a mixture of conifer species dominated by white fir (*Abies concolor*), Douglas-fir (*Pseudotsuga menziesii*), ponderosa pine (*Pinus ponderosa*), incense cedar (*Calocedrus decurrens*) and sugar pine (*Pinus lambertiana*). On xeric sites the forest canopy is less continuous, but the shrub layer is still abundant, as shown in Figure 5 (Sawyer & Thornburg 1977).



**Figure 5: Montane Klamath Mixed Conifer stand in the Sugar Creek study area and the lower strata herbaceous community at a study area plot.**

Plant species in temperate forest stands are influenced by insolation and microclimate, which impacts the presence (or the potential for the presence) of each under different canopy conditions. Table 2 lists common vascular plants, bryophytes and canopy lichens found in the analysis extent and the respective canopy position of each.

**Table 2: Vascular plants, bryophytes and lichens (Sawyer & Thornburg 1974, USDA Forest Service 2019)**

Klamath Mixed Conifer Species (Montane)	Canopy Position and Strata Level Dominance	Habitat & Ecological Factors
<i>Pseudotsuga menziesii</i>	Upper strata, dominant/codominant.	Shade-intolerant; ectomycorrhizal and ectendomycorrhizal relationships increase productivity.
<i>Picea breweriana</i>	Upper strata, dominant/codominant.	Shade tolerant.
<i>Abies concolor</i>	Upper strata, dominant/codominant.	Shade-intolerant.
<i>Picea engelmannii</i>	Upper strata, dominant/codominant.	Shade-intolerant.
<i>Pinus ponderosa</i>	Upper strata, dominant in xeric sites, codominant in mesic sites.	Shade-intolerant.
<i>Pinus contorta</i>	Upper strata, dominant/codominant.	Moderately shade tolerant.
<i>Calocedrus decurrens</i>	Upper strata, codominant on xeric sites, subdominant on mesic sites.	Shade-tolerant.
<i>Pinus lambertiana</i>	Upper strata, dominant/codominant.	Semi-tolerant to shade.
<i>Prunus emarginata</i>	Lower strata, dominant in disturbed stands with low canopy closure.	Shade-intolerant.
<i>Quercus sadleriana</i>	Lower strata.	
<i>Rosa gymnocarpa</i>	Lower strata, sub-dominant.	Shade-tolerant; persists most successfully in shaded xeric sites.
<i>Prosartes hookeri</i>	Lower strata.	
<i>Amelanchier pumila</i>	Lower strata.	
<i>Peridium aquilinum</i>	Lower strata, dominance varies.	Shade-intolerant; indicator of solar radiation intensity in understory.
<i>Arctostaphylos patula</i>	Lower strata, codominant.	Shade-intolerant.
<i>Quercus vacciniifolia</i>	Lower strata, dominant/codominant.	Xeric sites without canopy closure.
<i>Festuca idahoensis</i>	Codominant/subdominant.	Moderately shade tolerant.
<i>Amelanchier alnifolia</i>	Codominant/subdominant.	Shade-intolerant
<i>Berberis aquifolium</i>	Lower strata.	
<i>Arctostaphylos nevadensis</i>	Lower strata, dominant.	Does not tolerate canopy closure.
<i>Chimaphila umbellata</i>	Lower strata.	
<i>Clintonia uniflora</i>	Lower strata, codominant.	Surveyed in a range of 1.5% to over 60% full sunlight; higher abundance in climax conifer forest.
<i>Chrysolepi sempervirens</i>	Lower strata.	
<i>Leucothoe davisiae</i>	Lower strata.	
<i>Ceanothus velutinus</i>	Lower strata, codominant.	Moderately shade-tolerant.
<i>Symphoricarpos albus</i>	Lower strata, codominant.	Moderately shade-tolerant.
<i>Apocynum androsaemifolium</i>	Lower strata, codominant.	Shade-tolerant.
<i>Ceanothus prostrates</i>	Lower strata.	
<i>Galium triflorum</i>	Lower strata.	Favors diffuse light over full sun or full shade conditions.
<i>Taxus brevifolia</i>	Mid-strata, subdominant.	Shade tolerant.
<i>Adenocaulon bicolor</i>	Lower strata.	
<i>Calamagrostis koelerioides</i>	Lower strata.	
<i>Linnaea borealis</i> var. <i>longiflora</i>	Lower strata, dominant.	Surveyed in a range of 2% to full sunlight.
<i>Goodyera oblongifolia</i>	Lower strata, dominant.	
<i>Amelanchier semiintegrifolia</i>		
<i>Rubus parviflorus</i>	Lower strata, subdominant	Preference for mesic sites; shade tolerant in closed canopy.
<i>Pyrola picta</i>	Lower strata.	
<i>Sarcodes sanguinea</i>	Lower strata, subdominant.	Non-photosynthetic mycoheterotroph.
<i>Lobaria hallii</i>	All strata	
<i>Letharia vulpine</i>	All strata	

### Canopy Structure and Insolation Partitioning

The majority of incident solar radiation to Earth's vegetated surfaces is absorbed in the canopy layer, independent of the vegetation height above the surface (Schulze 2005). In the context of the surface radiation balance, insolation is best understood as a shortwave radiation flux influenced by atmospheric water vapor, aerosols and ozone (Table 3). Short-wave radiation is both absorbed and reflected by molecules in the atmosphere and the surface depending on the type and distribution of vegetative cover (Schulze 2005).

**Table 3: Radiation balance based on average global incident solar radiation of 342 Wm<sup>2</sup>. Source: Schulz 2005**

Location	Incoming Solar Radiation (Absorption)	Outgoing SW Radiation (Reflectance)
Space	100%	31%
Atmosphere	Absorbed by water vapor, aerosols and O <sub>3</sub> : 16%  Absorbed by clouds: 4%	Back scattered: 6%  Reflected by clouds: 16%
Surface	49%	Reflected by land surface: 9%

The spatial distribution of insolation on the surface is governed by solar elevation, surface orientation and albedo as well as the screening or reflection effects from the surrounding terrain and the diffuse fraction of radiant flux (Olseth & Skartveit 1997, Oliphant et al 2006). The spatial variability of annual photosynthetically active radiation (PAR) irradiation on the terrain surface is substantial in complex topography (Olseth & Skartveit 1997). Seasonal variation due to atmospheric opaqueness distributes radiation receipt spatially due to topographic effects, with aspects receiving direct radiant fluxes contrasting with aspects that receive primarily diffuse solar radiation (Nunez 1980).

Canopy cover is the proportion of the forest floor covered by the vertical projection of the crown area, as contrasted with canopy closure, which is the proportion of sky hemisphere obscured by vegetation from a single point (Smith et al. 2008). The upper strata of the canopy determine, therefore, the type of vegetation on the forest floor (Schulze 2005). In addition, understory microclimate is related to production in the overstory canopy as well as to the distribution of understory species and the maintenance of subsurface processes (Geiger 1965); understory microclimate is additionally determined by the topographic variability and availability of water at or near the surface.

Structural variation in the canopy should therefore create habitat variability and niche partitioning allowing greater potential for, all things being equal, measurable variation in

the diversity and richness of species. Since the stands in the study area are essentially unmanaged (except for historical fire suppression), it is important to understand that natural succession and regeneration in managed stands can differ considerably. Natural stands retain biogeochemical processes related to meteorological phenomena (lightening, wind, etc.) and temporal physical biomass processes (gravity-mass disturbances and overall nutrient load) over longer rotation periods, whereas silvicultural methods such as clear-cutting and selective harvest tend to maintain conifer stands in a perpetual state of early to pre-maturity seral classification by design and can result in a significant change in dominant conifer species composition at the landscape level. Management methods like these have resulted in biodiversity decline and generated interest in managing for increased structural complexity to enhance stand productivity, as noted before, by promoting complimentary resource utilization through spatial, temporal and physiological differentiation (Schulze 2005, Zhang et al 2012, Ishii et al. 2004). Other research has indicated that diversity related productivity may be more pronounced in boreal than in temperate forests generally, but that site-specific conditions are a determinant of this general observation (Paquette & Messier 2014).

### **Area Solar Radiation and Canopy Structure Metrics**

The leaf area index (LAI) quantitatively characterizes plant canopy layers and is often used to model forest canopy structure. It is defined as the one-sided leaf area per unit ground surface area for broadleaf or half of the total needle surface area per unit ground surface area for conifer forests.

It is notated as,

$$LAI = \frac{\text{leaf area}(m^2)}{\text{ground area}(m^2)} \quad (1)$$

(Chen & Black 1992).

LAI is also a key variable for regional and global models of biosphere-atmosphere exchanges of energy, carbon dioxide, water vapor, and other materials (Asner et al. 2003). Canopy photosynthesis and its equivalent, gross primary productivity (GPP), should theoretically reach a maximum as leaf area (LAI) increases to a value where PAR, or spectral emittance in the .4 to .7  $\mu\text{m}$  visible wavelength, is totally intercepted (Perry et al. 2008).

Photosynthetically available radiation (PAR) and its relationship to LAI is well established in the literature (Smith et al. 2008). The theoretical basis for the absorption of light as it passes through a forest canopy is described by the Beer-Lambert law:

$$t(x) = \exp[-K(LAI(x))] \quad (2)$$

where,

- $t$  is the proportion of PAR incident at the top of the canopy that is transmitted to a given point  $x$  within the canopy.
- $LAI(x)$  is the total leaf area above point  $x$ .
- $K$  is the extinction coefficient.

The coefficient ( $K$ ) indicates that light intensity decreases exponentially as it passes through each canopy layer (Perry et al. 2008, Schulze 2005). Also known as the attenuation coefficient, it describes the extent to which the radiant flux of a beam is reduced as it passes through a specific material, in this case the vegetation canopy. When a narrow (collimated) beam passes through canopy strata, the beam will lose intensity due to two processes: absorption and scattering. A detector can be used to measure a beam's directional path, or conversely using a non-narrow beam, one can measure how much of the lost radiant flux was scattered, and how much was absorbed. The extinction coefficient is therefore the sum of the absorption coefficient and the scattering coefficient (Wikipedia 2019).

The determination of the extinction coefficient requires direct measurement of LAI over consecutive seasons, as there is significant atmospheric and canopy structural variability in the determination of its value (Saitoh et al 2012). A meta-study of canopy light extinction showed significant intra-annual negative correlations between  $K$  and seasonal changes in LAI in natural ecosystems (Zhang et al. 2014). In another study, a  $K$  value of .48 and LAI of 6.2 are thresholds at which 95% of incident solar radiation is intercepted by the forest canopy in a stand of *Pseudotsuga menziesii* (Perry et al. 2008), a dominant upper strata canopy species in the Klamath Mixed Conifer forest type.

The LAI variable defines the number of equivalent layers of leaves relative to a unit of ground area, but the fraction of PAR that is absorbed (fPAR) measures the proportion of available radiation in the photosynthetically active wavelengths that are absorbed by a canopy (NASA 2017).

The basis for calculating fPAR using an exponential function is based on Beer's law:

$$fPAR = P_{\infty}[1 - \exp(-LAI)] \quad (3)$$

where  $P_{\infty}$  is the asymptotically limiting value of PAR absorption for an infinitely thick canopy and was set to 0.94, with the assumption made that the canopy leaves are randomly distributed (Oliphant et al. 2006). This assumption was made in the LAI calculation for this study as well. Both variables are used as parameters for calculating surface photosynthesis, evapotranspiration, and primary productivity, and those products are in turn used to calculate ecosystem functions such as terrestrial energy, carbon, water cycle processes, and biogeochemistry of vegetation (NASA 2017). Given its usefulness in this study, fPAR was nevertheless not directly calculated due to the need for complex adjustments required to 'effective LAI', i.e., the unadjusted product derived from optical methods. It can be inferred, however, through the interpretation of and relationship to other measures such as the normalized differentiation vegetation index NDVI (Perry et al. 2008).

In addition to LAI, metrics were derived from LiDAR data to measure area insolation and the canopy strata variables effecting solar radiation in the canopy understory.

- 1.) Canopy Height (CHM) in meters was derived from the LiDAR digital surface model (DSM) and the first returns data at 1-meter resolution. It was used to derive area incident solar radiation adjusted for topography as well as to provide data required for sample plot selection by using the average values at 30 meters resolution matched to the gridded canopy metrics raster (BCAL, 2017, Evans et al. 2009, QSI 2015).
- 2.) Area Solar Radiation (ASR) in  $WH/m^2$  was derived using the sum of 12 monthly values (2015 calendar year) for the analysis extent and resampled to 30 meters resolution. ASR was modeled in ArcGIS using the sum of the LiDAR bare earth surface product (DEM) and the canopy height model (CHM) product as the topographic parameter, including topographic elevation data for the entire watershed used as the horizon parameters for the ASR model equations (Fu & Rich 2000, 2002).

Radiation parameters included diffuse model type (radiation flux varied with zenith angle in a non-uniform overcast sky condition), diffuse proportion (proportion of global normal radiation flux that is diffuse by month), and

atmospheric transmissivity (fraction of radiation that passes through the atmosphere by month) input parameters were derived from meteorological data acquired from a field station within close proximity to the analysis extent in 2015.

ASR is a term used to describe irradiation, or the sum of downward area irradiance per unit area over a stated time interval expressed in  $\text{WH}/\text{m}^2$ . Irradiance is the instantaneous density of solar radiation on a unit area expressed in  $\text{W}/\text{m}^2$ . It comprises Global radiation ( $\text{Global}_{\text{tot}}$ ), or the sum of direct ( $\text{Dir}_{\text{tot}}$ ) and diffuse ( $\text{Dif}_{\text{tot}}$ ) radiation of all sun map and sky map sectors as shown in the following equation:

$$\mathbf{Global}_{\text{tot}} = \mathbf{Dir}_{\text{tot}} + \mathbf{Dif}_{\text{tot}} \quad (4)$$

(Fu & Rich 2000, 2002)

$\text{Dir}_{\text{tot}}$  for a given location is the sum of the direct insolation ( $\text{Dir}_{\theta,\alpha}$ ) from all sun map sectors. Direct insolation from the sun map sector ( $\text{Dir}_{\theta,\alpha}$ ) with a centroid at zenith angle ( $\theta$ ) and azimuth angle ( $\alpha$ ) is calculated using the following equation:

$$\mathbf{Dir}_{\theta,\alpha} = \mathbf{S}_{\text{Const}} \times \beta^{m(\theta)} \times \mathbf{SunDur}_{\theta,\alpha} \times \mathbf{SunGap}_{\theta,\alpha} \times \cos(\mathbf{AngIn}_{\theta,\alpha}) \quad (5)$$

where:

- $\mathbf{S}_{\text{Const}}$  is the solar flux outside the atmosphere at the mean earth-sun distance, known as solar constant. The solar constant used in the analysis is  $1367 \text{ W}/\text{m}^2$ . This is consistent with the World Radiation Center (WRC) solar constant.
- $\beta$  is the transmissivity of the atmosphere (averaged over all wavelengths) for the shortest path (in the direction of the zenith).
- $m(\theta)$  is the relative optical path length, measured as a proportion relative to the zenith path length.
- $\text{SunDur}_{\theta,\alpha}$  is the time duration represented by the sky sector. For most sectors, it is equal to the day interval (for example, a month) multiplied by the hour interval (for example, a half hour). For partial sectors (near the horizon), the duration is calculated using spherical geometry.
- $\text{SunGap}_{\theta,\alpha}$  is the gap fraction for the sun map sector.
- $\text{AngIn}_{\theta,\alpha}$  the angle of incidence between the centroid of the sky sector and the axis normal to the surface. (Fu & Rich 2000, 2002)

Total diffuse solar radiation for the location ( $Dif_{tot}$ ) is calculated as the sum of the diffuse solar radiation ( $Dif$ ) from all the sky map sectors.

The diffuse radiation at its centroid ( $Dif$ ) is calculated, integrated over the input time interval and corrected by the gap fraction and angle of incidence using the following equation:

$$Dif_{\theta,\alpha} = R_{glb} \times P_{dif} \times Dur \times SkyGap_{\theta,\alpha} \times Weight_{\theta,\alpha} \times \cos (AngIn_{\theta,\alpha}) \quad (6)$$

where:

- $R_{glb}$  is the global normal radiation.
- $P_{dif}$  is the proportion of global normal radiation flux that is diffused.
- $Dur$  is the time interval for analysis.
- $SkyGap_{\theta,\alpha}$  is the gap fraction (proportion of visible sky) for the sky sector.
- $Weight_{\theta,\alpha}$  is the proportion of diffuse radiation originating in a given sky sector relative to all sectors.
- $AngIn_{\theta,\alpha}$  is the angle of incidence between the centroid of the sky sector and the intercepting surface. (Fu & Rich 2002)

- 3.) Intensity of return (IR) is an amplitude describing the peak power ratio of the laser return to the emitted laser, calculated as a function of surface reflectivity. Values are corrected for variability between flight lines and pre-processed at a 0.5-meter pixel resolution before being processed using the BCAL vegetation intensity tools and output to a 30 meter resolution for the vegetation excluding bare earth data (BCAL 2017, QSI 2015).

Forest remote sensing research indicates that the returns of high-intensity and the low intensity peak count of the intensity distribution were predictive of live and dead tree biomass, respectively (Kim et al. 2009).

- 4.) Total Vegetation Density is a derived percent ratio of vegetation to ground returns within each pixel (per  $m^2$ ) or

$$TVD = nV/nG * 100 \quad (7)$$

where  $nV$  is the number of vegetation returns and  $nG$  that of ground returns (BCAL 2017, QSI 2015).

- 5.) Canopy Relief Ratio is a derived mean height less the minimum height divided by the maximum height less the minimum height within each pixel (per  $m^2$ ) or

$$CRR = \frac{\mu(\text{height}) - \min(\text{height})}{\max(\text{height}) - \min(\text{height})} \quad (8)$$

It represents a quantitative descriptor of the relative shape of the canopy from altimetry observation which describes the degree to which canopy surfaces are in the upper ( $CRR > 0.5$ ) or in the lower ( $CRR < 0.5$ ) portions of the height range (Pike & Wilson 1971, Parker & Russ 2004, BCAL 2017, Evans et al. 2009, QSI 2015).

- 6.) Texture of Heights (TH) is the variance of height of points per pixel equal to the SD of height above the ground threshold and height below crown threshold per  $m^2$  or

$$TH = \sigma(\text{height} > GT \text{ and } \text{height} < CT) \quad (9)$$

where GT is the ground threshold and CT is the crown threshold (BCAL 2017, QSI 2015). Thresholds were based on canopy height data for the study area.

- 7.) Foliage Height Diversity is a derived Shannon diversity index statistic calculated as percentage cover at different heights per  $m^2$ , or

$$FHD = - \sum p_i \ln p_i \quad (10)$$

where  $p_i$  is the proportion of the number of returns in the  $i$ th layer to the sum of points of all the layers (using all points) (BCAL 2017, Evans et al. 2009, QSI 2015).

### Measures of Biodiversity

Biodiversity as a biogeographic concept cannot be described without reasonable consideration of spatial and temporal scales (Schultze 2005). Alpha diversity ( $\alpha$ -diversity) was developed by R. H. Whittaker during his study of plant communities in the

Klamath Ecoregion and introduced with the complementary statistics, beta diversity ( $\beta$ -diversity) and gamma diversity ( $\gamma$ -diversity). Beta diversity is dimensionless, as a comparative statistic, i.e., the ratio between  $\gamma$ -diversity (regional) and  $\alpha$ -diversity (local) diversities (Whittaker, 1960);  $\alpha$ -diversity and  $\gamma$ -diversity are limited to discrete units of space – roughly communities and ecosystems respectively (Schultze 2005).

Whittaker postulated that total species diversity in a landscape, described by  $\gamma$ -diversity, is determined by the mean species diversity at a community scale ( $\alpha$ -diversity) and also by the differentiation among those communities ( $\beta$ -diversity) (Whittaker 1960, 1972). Whittaker's subsequent usage of  $\alpha$ -diversity implies the application of the statistic across multiple sites in a landscape, strongly influencing its primary use at the assemblage and community scale (Whittaker 1960, 1972).

Currently there are three varied types of biodiversity measures used in most plant ecology research, each providing evidence of an advancing understanding of diversity and ecological processes. These are taxa based (taxonomic diversity), trait based (functional, based on phenotypes expressions of traits) and phylogenetic diversity, defined as the minimum total length of phylogenetic branches required to span a given set of taxa on the phylogenetic tree (Faith 1992).

Taxonomic index values that provide a quantitative measure of richness and abundance in a nominal scale dataset. Menhinick's Index is a widely used measure of  $\alpha$ -diversity that is based on the ratio of number of species ( $S$ ) and the square root of the total number of individuals ( $N$ ). It is notated:

$$MDI = \frac{s}{\sqrt{N}} \quad (11)$$

where

- $s$  is the number of different species in your sample and
- $N$  is the total number of individual organisms in the sample.

To assess the impact of species abundances, the Simpson Index relates the contribution made by each species to the total number of individuals present.

$$SDI = 1 - \sum_{i=1}^s p_i^2 \quad (12)$$

where  $p_i$  is the proportion of individuals found in species  $i$ . As SDI's summation increases, evenness decreases. The difference between that value and 1 produces a range of 0-1, 1 being a monoculture. SDI is a non-parametric index value less sensitive to species richness in that species specific abundances are considered.

There has been a significant recognition among researchers that classifying terrestrial plant species into functional types rather than their higher taxonomic identity improves our understanding of processes at the ecosystem scale, including vegetation responses and effects on climate, atmospheric chemistry, land use and disturbances (Cornelissen et al 2003). Although there is no consensus on a uniform definition for functional diversity in the literature, there are clear contrasts in areas of its measurement: those that use trait values directly and those that use distance-based and dendrogram-based constructs (Petchey et al. 2009). The number of traits included in the analysis must be adequate to capture the specific function of interest, continuous traits being more effective at capturing interspecific variability in trait values than categorical traits (Laureto et al. 2015).

It has also been proposed that functional diversity can be divided into functional richness, evenness and divergence as well, and that functional diversity is affected by the range of trait values (phenotype expressions) present as well as the distinct species in that range, and notably, that the trait measured is more important than the specific measure used (Petchey et al. 2009). Low functional richness, for instance, indicates that some alpha niches (i.e., resources) potentially available to the community remain unused, reducing productivity (Mason et al. 2005). The community weighted mean (CWM), as a metric of functional composition, reflects the functional mean of a single (focal) trait (Pla et al. 2012).

$$CWM = \sum_{i=1}^s w_i x_i \quad (13)$$

where

- $W_i$  is the niche space filled by the species within the community, i.e., meters of cover by species in a survey plot
- $X_i$  is the trait value (see Cornelissen et al 2003) for species  $i$ , (Mason et al. 2005).

Given the many leaf traits available to research that impact productivity, such as specific leaf area (SLA) in  $\text{mm}^2\text{g}^{-1}$ , they represent a practical area of focus in any canopy insolation study concerned with the productivity of forest biota. Forest productivity research indicates that functional and phylogenetic diversity significantly influence biomass productivity on small scales, while taxonomic diversity evidenced “only indirect effects” at that scale (Hao et al 2017).

## Data

### **Multispectral Imagery Data Specification**

WorldView-2 is a high-resolution 8-band multispectral satellite operating at an altitude of 770 km. A single early season ‘leaf-on’, corrected satellite image was used for the study, minimizing radiometric variability (manifested as differences in the coloration of image features - trees and openings) and radial displacement (or the apparent elongation or displacement of objects having height - such as trees - in satellite imagery) (Hamilton et al. 2013). The imagery provides 1.85 m multispectral resolution. Its sensor resolution for multispectral is 1.85 m GSD at nadir, 2.07 m GSD at 20° off-nadir, and the swath width is 16.4 km at nadir with a demonstrated geolocation accuracy of <3.5 m without ground control.

Table 4 describes the multispectral bands in the Sugar Creek image and highlights the bands used in canopy analysis (NDVI and LAI).

**Table 4: Worldview2 multispectral band wavelength and utilization (DigitalGlobe).**

Coastal Blue	0.40 - 0.45 $\mu\text{m}$	Chlorophyll absorption sensing
Blue	0.45 - 0.51 $\mu\text{m}$	Soil & vegetation discrimination sensing, forest type mapping
Green	0.51 - 0.58 $\mu\text{m}$	Green reflectance of vegetation sensing
Yellow	0.585 - 0.625 $\mu\text{m}$	Vegetation feature identification
<b>Red</b>	<b>0.63 - 0.69 <math>\mu\text{m}</math></b>	<b>Chlorophyll absorption sensing; vegetation analysis and type differentiation</b>
Red Edge	0.705 - 0.745 $\mu\text{m}$	Vegetative condition sensing; band directly related to plant health revealed through chlorophyll status
<b>Near Infrared (NIR1)</b>	<b>0.77 - 0.895 <math>\mu\text{m}</math></b>	<b>Vegetation type sensing, vigor and biomass survey</b>
NIR2	0.86 – 1.04 $\mu\text{m}$	Overlaps the NIR1 band; supports vegetation analysis and biomass studies; water vapor influence

### LiDAR Data Acquisition and Specification

The Sugar Creek LiDAR survey used the NAD83 (CORS96) datum as the basis for the mission coordinate system (epoch 2002.00). It was accomplished using a Leica ALS50 system in late season ‘leaf-on’ conditions. The Leica ALS50 laser system records up to four range measurements (returns) per pulse. Discrepancies between first return and overall delivered density vary with terrain, land cover, and the presence of water bodies (QSI 2015).

The LiDAR survey was accomplished with an opposing flight line side-lap of  $\geq 50\%$  and  $\geq 100\%$  overlap in order to reduce laser shadowing and increase surface coverage. Ground control surveys were conducted to support the acquisition. These ground control data were used to geospatially correct the aircraft positional coordinates as well as perform quality assurance checks on final LiDAR products. Collection procedures for ground control surveys used real time kinematic (RTK) and post processed kinematic (PPK) survey techniques (QSI 2015).

Ground survey data were collected with a Position Dilution of Precision (PDOP) of  $\leq 3.0$  with at least six satellites in view of the stationary and roving receivers (QSI 2015). Post data acquisition processing included the use of a suite of automated and manual techniques to process the data into the product specifications requested including a first-return density of no less than 8 points/m<sup>2</sup>. Actual first return average point density for the survey averaged 16.11 points/m<sup>2</sup> for the analysis extent (while ground classified point density, used for accuracy assessments with ground control survey data, was 2.36

points/m<sup>2</sup>). First return density describes the density of pulses emitted from the laser that return at least one echo to the system. First returns reflect off the highest feature on the landscape within the footprint of the laser pulse. In forested areas the highest feature could be a tree or other natural feature not including the bare earth surface, while in areas of unobstructed ground the first return will be the only return echo and will represent the bare earth surface (QSI 2015).

The accuracy of the LiDAR data collection can be described in terms of absolute and relative accuracy. Absolute accuracy is the consistency of the data with external data sources. Relative accuracy is the consistency of the dataset with itself. The acquisition engineers assessed absolute accuracy, shown in Table 5, using Fundamental Vertical Accuracy (FVA) methods outlined in the FGDC National Standard for Spatial Data Accuracy. FVA compares known RTK ground control data collected on open, bare earth surfaces with level slope (<20°) to the triangulated surface generated by the LiDAR points. FVA is a measure of the accuracy of LiDAR point data in areas where the system has a high probability of measuring the ground surface at the 95% confidence interval (1.96 \* RMSE) (QSI 2015).

**Table 5: Absolute Accuracy of Ground Control Points Source: Quantum Spatial**

Sample Size	1.96*RMSE	Mean	Median	RMSE	Standard Deviation
109	0.079m	-0.004m	-0.016m	0.041m	0.041m

Vertical relative accuracy refers to the internal consistency of the data set as a whole. When the system is well calibrated, the swath-to-swath vertical divergence is low (<0.10 meters). The vertical relative accuracy was computed by comparing the ground surface model of each individual flight line with adjacent points in overlapping survey areas (QSI 2015). Vertical relative accuracy for the analysis extent is shown in Table 6.

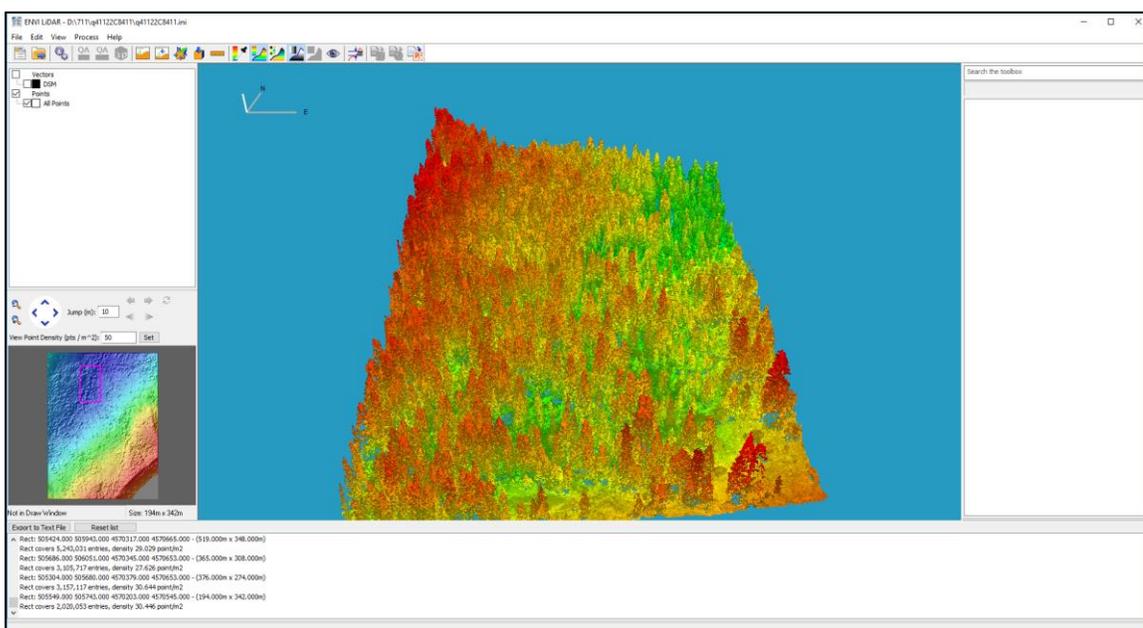
**Table 6: Vertical Relative Accuracy Source: Quantum Spatial**

Surfaces	Mean	Median	RMSE	Standard Deviation
144	0.050m	0.050m	0.053m	0.012m

Quantum Spatial used ENVI software and other specialized tools to perform post data acquisition processing to derive the geospatial products used in the study, including the DEM, CHM and point-cloud dataset.

## Data Processing Summary

Post data acquisition processing for the LiDAR and imagery products provided the inputs required to develop a georeferenced canopy model. Boise Center Aerospace Laboratory (BCAL) Vegetation and Intensity tools were used to develop canopy model metrics from the LiDAR point-cloud data in ENVI, while ERDAS Spatial Modeler was used to derive LAI data from the satellite imagery. Figure 6 shows a LiDAR tile that was mosaiced with others before gridding to raster for analysis at a 30-meter resolution (intensity and vegetation products).



**Figure 6:** A Sugar Creek LiDAR point-cloud tile rendered in the ENVI LiDAR analysis interface using default display settings. The tile is pre-processing artifact, the scale related to data acquisition and storage factors and not mapping products. The screen is useful for viewing the canopy laser light return and structure visually.

The Klamath Mixed Conifer (KMC) forest type is a California Wildlife Habitat Relationship System (CWHRS) attribute “crosswalked” in the CALVEG (Classification and Assessment with LANDSAT of Visible Ecological Groupings) data layer. All LiDAR and imagery data comprising the predictor variables were subsequently resampled to 30-meter resolution and clipped to the KMC vegetation type of the analysis extent in ArcGIS. This is shown in a diagram (Figure 7) that details the complete pre-processing dataflow required to select plots for biodiversity sampling.

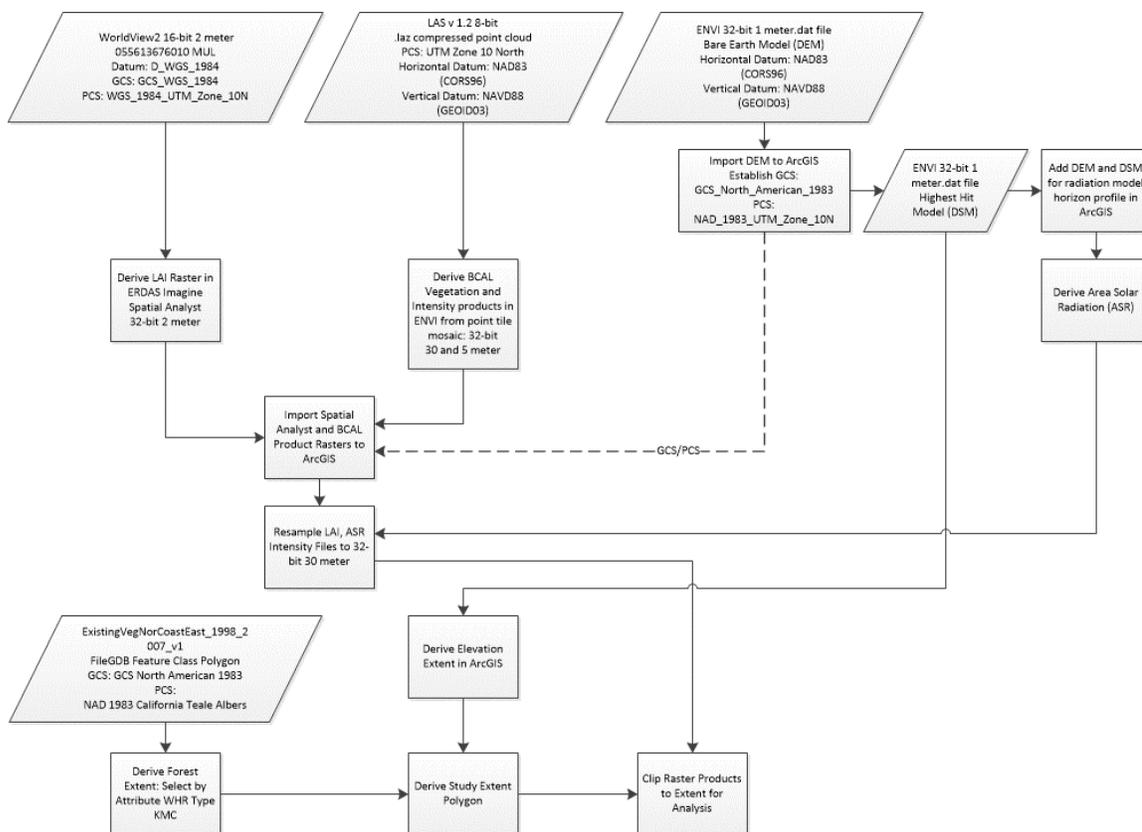


Figure 7: Data processing steps for the analysis extent and study area.

The geographic coordinate system and projected coordinates systems were based on the survey information provided in the post-acquisition LiDAR products (bare earth DEM). The WorldView 2 imagery was resampled and transformed to these coordinate systems as well. Once raster images containing the variable information were processed, a series of related tables were created in ArcGIS to analyze the processed data.

## Methods

Modeling methods consisted of developing a predictive multiple linear regression model from the remotely sensed and field data and subsequently analyzing the spatial autocorrelation and clustering of this product with spatial statistics. Imagery and LiDAR data acquisition dates were both within an acceptable if not ideal ‘leaf-on’ seasonal timeframe (deciduous species were found primarily in the lower strata as opposed to the dominant upper strata species).

Forest dynamics, notably disturbance type and succession generally, are not the focus of the modeling. Each 30-meter field plot references a pixel of gridded LiDAR data of the same dimension. All diversity values resulting from plot data are numerically ordinal and assumed to be spatially continuous. There is a noted time-lag of five years between the LiDAR and multispectral data acquisition and the final field sampling of biodiversity plots. The assumption that a temporal lag for early seral species recruitment is likely after a disturbance would address this issue on average. For this reason, sites evidencing recent canopy disturbance were to be excluded from the study to maintain data continuity.

All fine-scale data was resampled to 30-meter resolution to match plot size and location (using bilinear or nearest neighbor methods for discrete and continuous data) in order to create distinct metrics with unique centroid coordinates for the analysis. Forest ecosystems with less fine scale variability, such as closed canopy forests characterized by large, contiguous stands of trees, require coarse-scale analyses. High-resolution imagery can be used but coarse resolution imagery is the better option because the larger pixels reduce fine scale variability that is not of interest such as shadows within the canopy (Hamilton et al 2013).

### **Input Data Analysis**

LiDAR point-cloud, elevation model data and WorldView2 8-band multispectral imagery were processed to derive incident radiation and canopy mensuration variables that describe the variation in energy and canopy structure for stand level modeling within the study extent.

It can be challenging to quantify LAI accurately due to significant spatio-temporal variability and to measurement limitations inherent in current methodologies (Smith et al. 2008). These methods utilize ‘direct’ approaches employing field sampling or ‘indirect’ approaches involving optical instruments combined with modeling (Smith et al. 2008). Optical instruments used to estimate LAI rely on the measurement of light transmittance and, because they are based on Beer’s law, the resultant measures are termed ‘effective LAI’ (Smith et al. 2008).

Further, these methods assume a random spatial distribution of and represent insolation interception without the distinction between photosynthetically active materials and woody components of the canopy (Smith et al. 2008). Non-randomness in the spatial distribution of canopy components is a more realistic perspective of the true structure of forest canopies. When corrections for non-randomness are not made to the optical model,

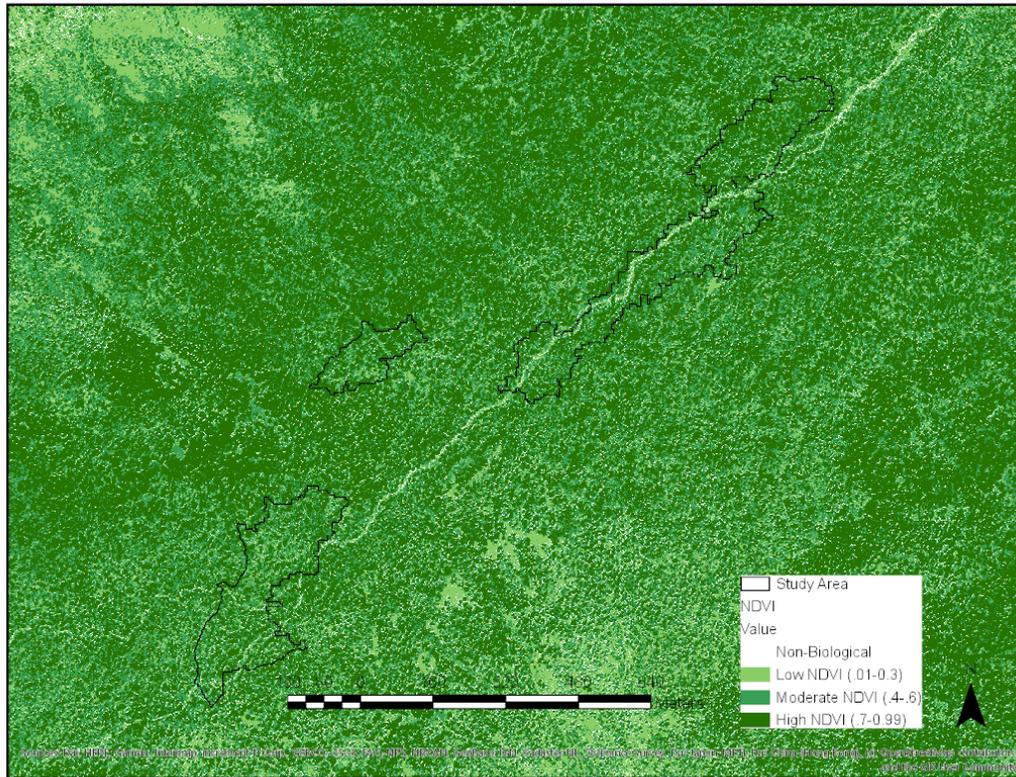
the terms plant area index (PAI) or surface area index (SAI) are often used to characterize the results (Smith et al. 2008).

LAI was derived in ERDAS Imagine by first deriving a NDVI from the multispectral image and then using a linear model to determine LAI (Wulder et al. 1998). NDVI is calculated from the visible red and near-infrared light (NIR) reflected by vegetative cover, where cells scatter (i.e., reflect and transmit) solar radiation in NIR spectral region (Hearne 2008).

It is notated,

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (15)$$

NDVI uses normalization to minimize effects of variable irradiance and is commonly used to indicate the amount and vigor of vegetation and to differentiate vegetated and non-vegetated areas in an image. Plants appear relatively dark in the PAR and relatively bright in the NIR (Hearne 2008). The biophysical interpretation of NDVI for Figure 8 is the fraction of absorbed photosynthetically active radiation (fPAR). Although derived index values can range from -1.0 to +1.0, values less than zero do not have any biological meaning and the range was adjusted for the model to exclude them.



**Figure 8: Classified NDVI map of study area in 2-meter resolution. Source: DigitalGlobe.**

The empirical conversion from the NDVI described by Wulder et al. (1998) was then used to derive the canopy leaf area in the ERDAS Spatial Modeler tool.

$$LAI = (17.35 * NDVI) - 9.01 \quad (16)$$

(Oliphant et al. 2006)

### **Study Area Predictor Variable Distributions**

The histograms in Figure 9 depict distributions of the 144 predictor variable data points.

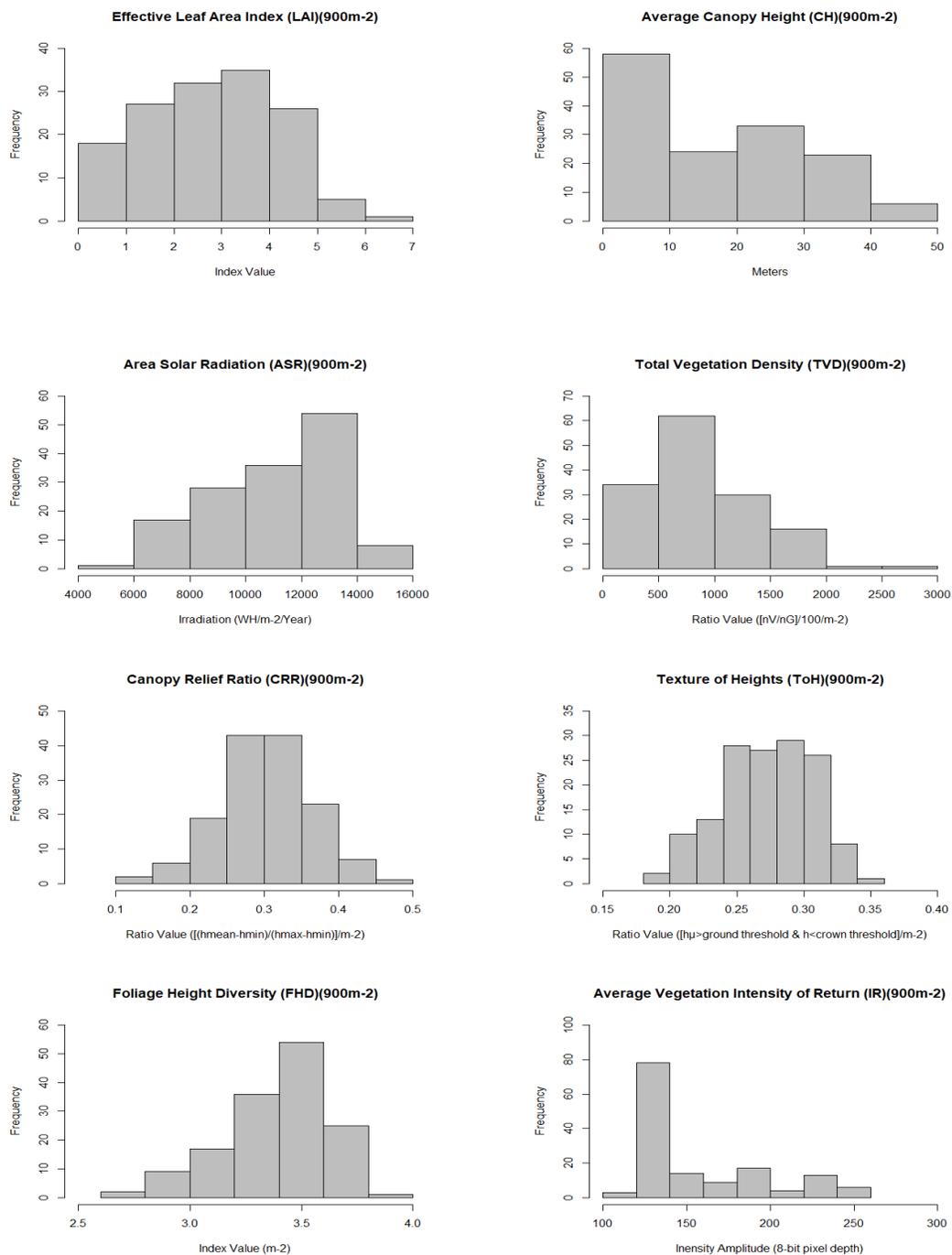


Figure 9: Study area predictor variable distributions.

A collinearity analysis of predictor variables (Figure 10) displays coefficient values in the upper half of the matrix and nonparametric trend lines (loess smoothed fits) and correlation ellipses below the diagonal. Modelled irradiation (ASR) within the study area

data set exhibited a correlation with canopy height. A significant positive correlation coefficient ( $r$ ) of .61 provided a credibility test of the irradiation data since canopy height should reasonably have a relationship to  $WH/m^2$ . This is because of the direct relationship between photosynthetic processes and productivity to the quantity of radiant energy, in this sense, in one area of space relative to another area of space.

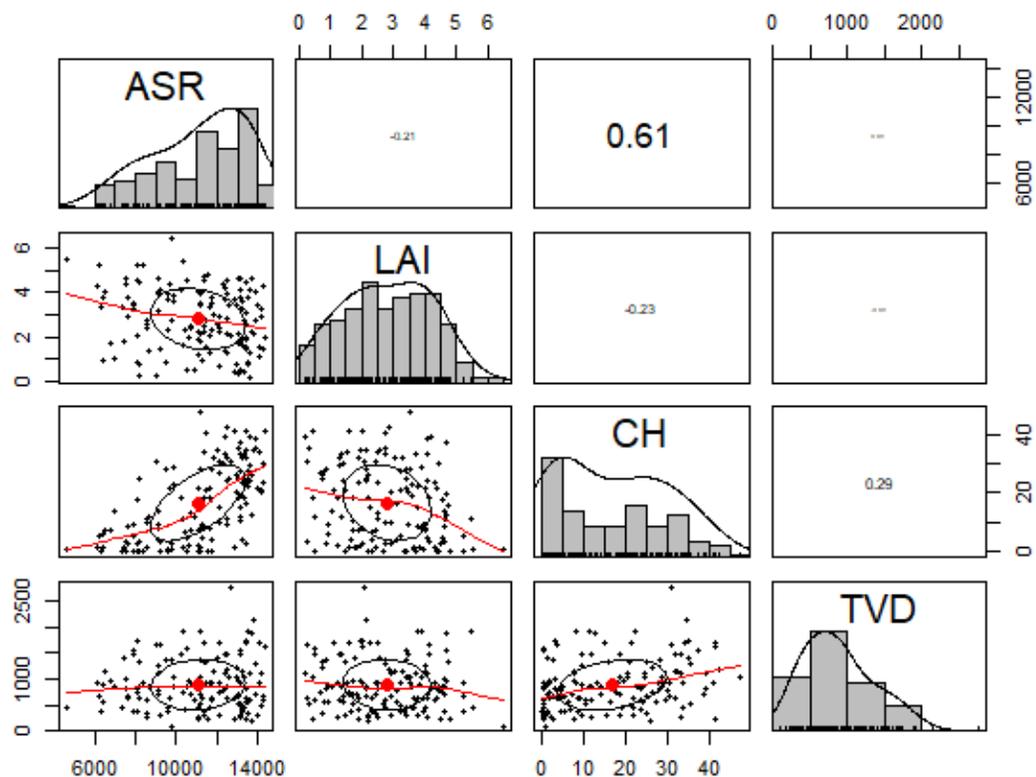


Figure 10: Density matrix analysis of ASR, LAI, CH and TVD

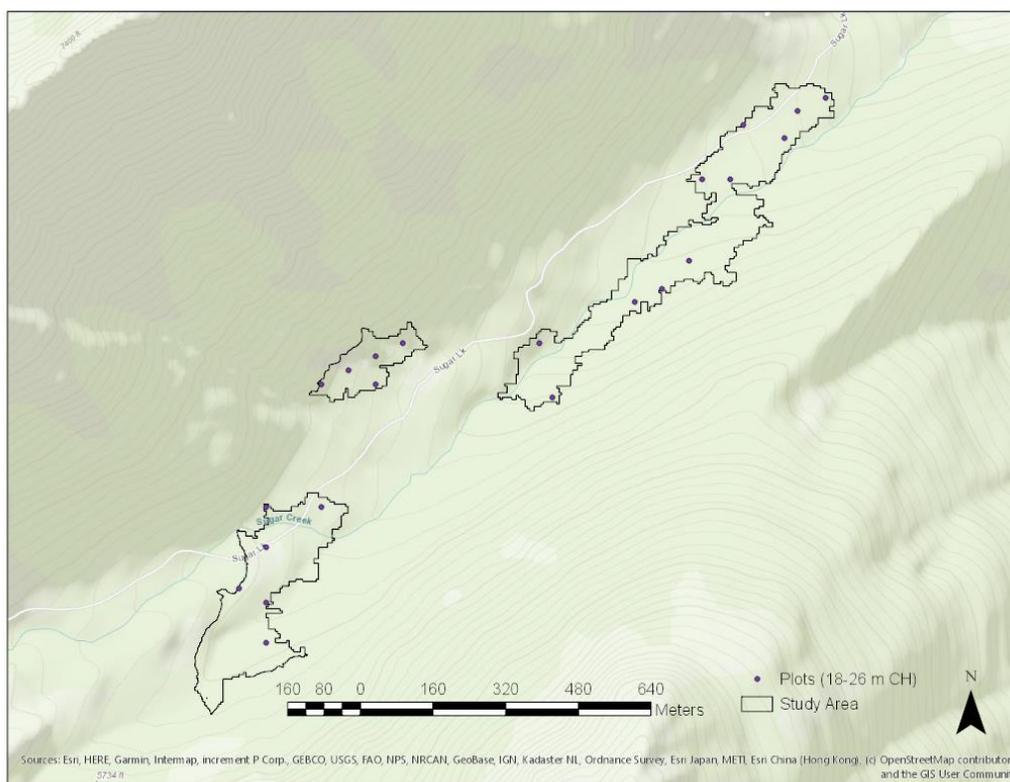
Effective leaf area (LAI) & vegetation density relationships with radiance were not significant, and leaf area and radiance were negatively correlated. This may be partially attributable to the acquisition date for the multi-spectral data being early in the growth season. A seasonal analysis of LAI would be useful to validate data to use for further canopy studies. However, other controlling factors such as soil and water limitations would also require review.

### **Sample Plot Selection**

Foresters often delineate stands using photogrammetry or even LiDAR based methods which aide in defining contiguous communities of conifer forest sufficiently uniform in structure, age and size class, assemblage, and site quality to distinguish it from adjacent communities. For purposes of this study, stands or patches of contiguous forest of similar age and tree species were defined by canopy height and species assemblage (KMC).

The study area extent was defined by identifying KMC stands within a contiguous 18.4896 hectare area and sampling plots that were in a range of one standard deviation of the mean canopy height. Selected stand strata were used to represent the best sample variability and increase the area for predictive modeling. Thirty-two data points met this criterion after buffering major stream channels from the analysis.

Random points were then selected, yielding possible plots ( $n/N =$  stratified sampling fraction  $= 22/32 = 0.69$ ) as shown in Figure 11. Six plots were sampled ( $6/32 = 0.19$ ). This provided an expansion factor of 5.3 for the stratified data and a population sampling fraction of  $6/144 = 0.04$  (an expansion factor of 24). Because the canopy height was not uniform across the population data, models with canopy height as a predictor would be limited by the range of the sample, while others would have greater range and applicability.



**Figure 11: Study Area Plots. Random plots were selected from canopy heights in the range of one standard deviation above the mean canopy height within the study area.**

There was no discrimination of sampling method based on slope, aspect or relative soil hydrology of the site (mesic, xeric, hydric).

The plot map was published to a cloud server that could be subsequently downloaded to a field data collection tablet paired with a GNSS receiver using a satellite based real-time kinematic (RTK) network for real-time plot position corrections (or SBAS if the RTK was not available). This method achieved an estimated horizontal plot accuracy of approximately 2 meters or less with respect to the LiDAR survey data used to produce the map service. Manual reconciliation of projected map coordinates and GNSS receiver data indicated that plots with 2 to 4 meters of estimated field data logger accuracy were matched to map coordinates within that level of estimated accuracy.

Each sample plot provided 900 square meters of taxonomic field data; the survey was conducted by estimating species abundance and noting canopy position within the plot

perimeter. Quantitative precision was not the focus but was estimated with respect to abundance. Surveyed taxa included vascular plants, bryophytes and canopy lichens.

### **Biodiversity Response Variables**

Six random stratified plots were sampled totaling an area of 5,400 m<sup>2</sup> (Table 7). Accuracy of species identification was high, with only a few visible species remaining unidentified or uncounted. The data are representative on a comparative basis from one plot to another, but as an exploratory survey they likely understate actual richness and abundance.

**Table 7: Plot Survey Data**

Plot	Richness (R)	Abundance (A)	MDI	SDI
32	14	361	0.737	0.829
30	15	222	1.007	0.629
26	18	655	0.703	0.866
29	18	601	0.734	0.723
25	11	171	0.841	0.852
23	22	662	0.855	0.771

Richness (number of specific taxa), abundance (quantity of each specific taxa) and two measures of biodiversity were computed for the model (Figure 12).

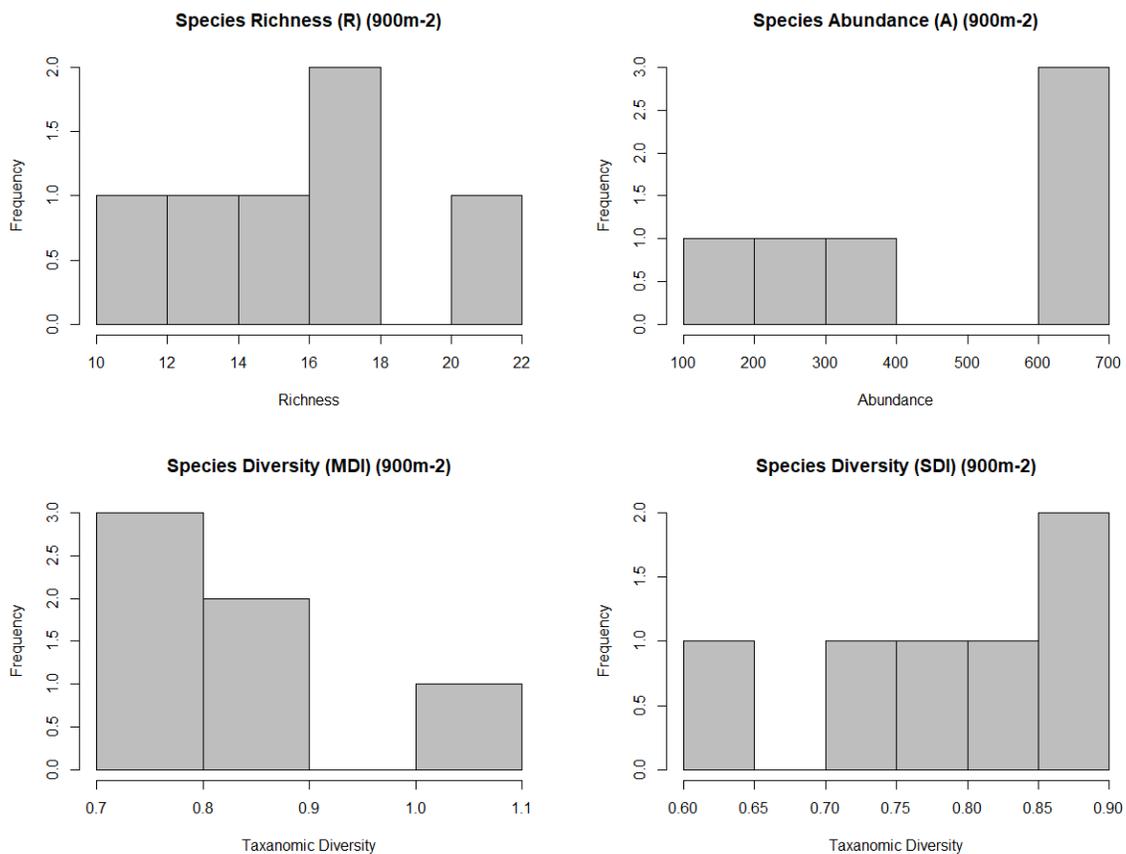


Figure 12: Distribution of plot values for species richness, abundance and diversity indexes.

## Multiple Linear Regression and Prediction Modeling

Multiple linear regression, used often by ecologists to model species distributions, was selected to model canopy metrics as predictor variables and a biodiversity response variable. Future work would allow testing the model over greater spatio-temporal conditions that would better define the model's range of applications and prediction suitability (Guisan & Zimmermann 2000).

The formula is notated as

$$Y_i = b_0 + b_1x_{1,i} + b_2x_{2,i} + \dots + b_px_{p,i} + \varepsilon_i \quad (14)$$

where

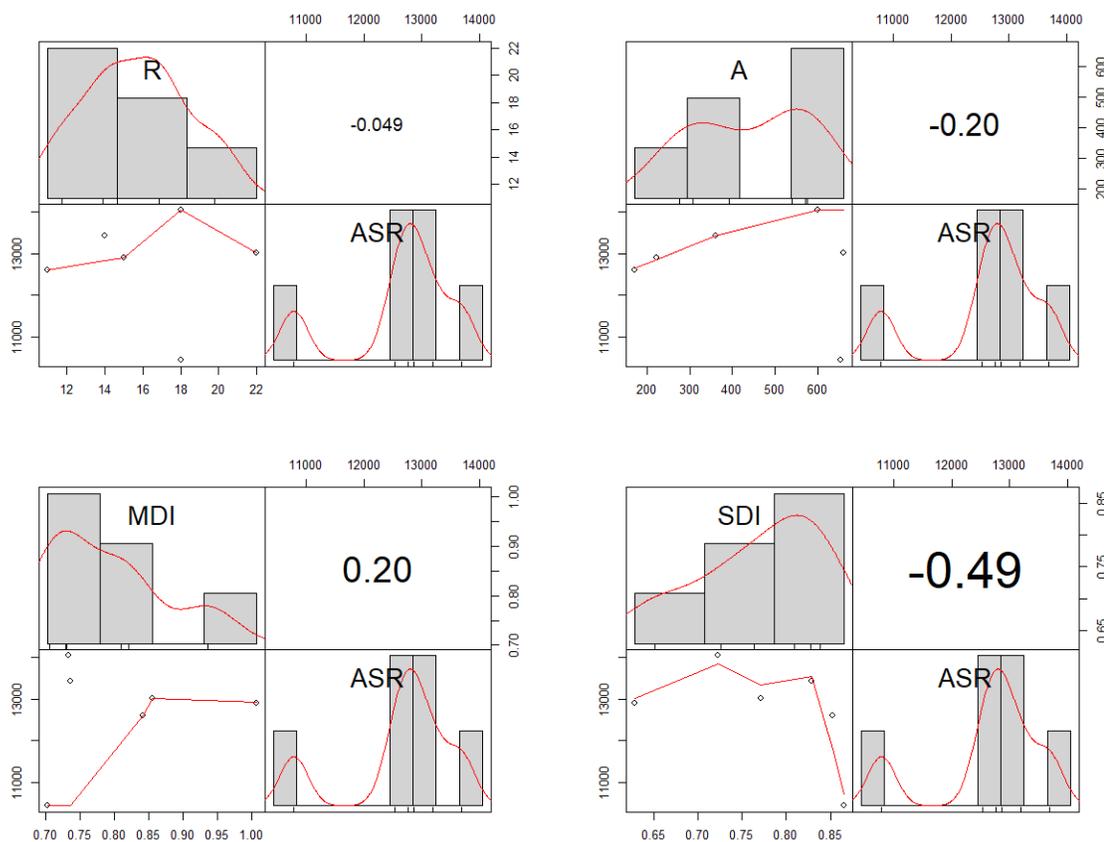
- $Y_i$  is the response variable,
- $x_{1,i}, \dots, x_{p,i}$  are predictor variables (fixed, nonrandom),
- $b_0, \dots, b_p$  are unknown regression coefficients (fixed) and
- $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$  represents the random error

Guisan & Zimmermann (2000) describe general patterns in the geographical distribution of species that originated in the field of biogeography to characterize the usefulness of predictive models. They note that the importance of abiotic and biotic factors at the margins of a species' range of biological and physical stresses, respectively, as well as physical limits caused by environmental gradients and physiological constraints more generally determine the habitat preferences (Guisan & Zimmermann 2000). The regression model therefore focuses on “generality with precision” to predict and accurate response within a finite set of limitations (or a “simplified reality”) that implies natural phenomena are too complex and heterogeneous to be predicted accurately in every spatio-temporal aspect (Guisan & Zimmermann 2000).

Canopy metrics, radiation and diversity were evaluated using R language functions for correlation, predictor collinearity, regression, model fit and predictive modeling within the study area. After prediction maps were created, the global inferential statistics Moran's-I, Getis Global G and the local Getis-Ord  $G_i^*$  statistics were used to evaluate spatial autocorrelation and clustering of continuous biodiversity data.

## Results

Richness of species, abundance of individuals and taxonomic diversity were initially plotted to ASR to confirm any relationships that might exist in the data using a paired-plots diagram (Figure 13). Abundance (A) and diversity (MDI) exhibited modest negative and positive correlations, respectively, of similar magnitude. Diversity (SDI) exhibited the strongest correlation, with a correlation coefficient of  $-0.49$ .



**Figure 13: Density matrix display of species richness, abundance, MDI and SDI with respect to ASR.**

When collinearity of all predictors was analyzed with respect to diversity in the paired-plots matrix, species richness (R) was most significantly correlated with the texture of heights (TH,  $r = .76$ ) and the intensity of the returns (IR,  $r = .78$ ) (Figure 14).

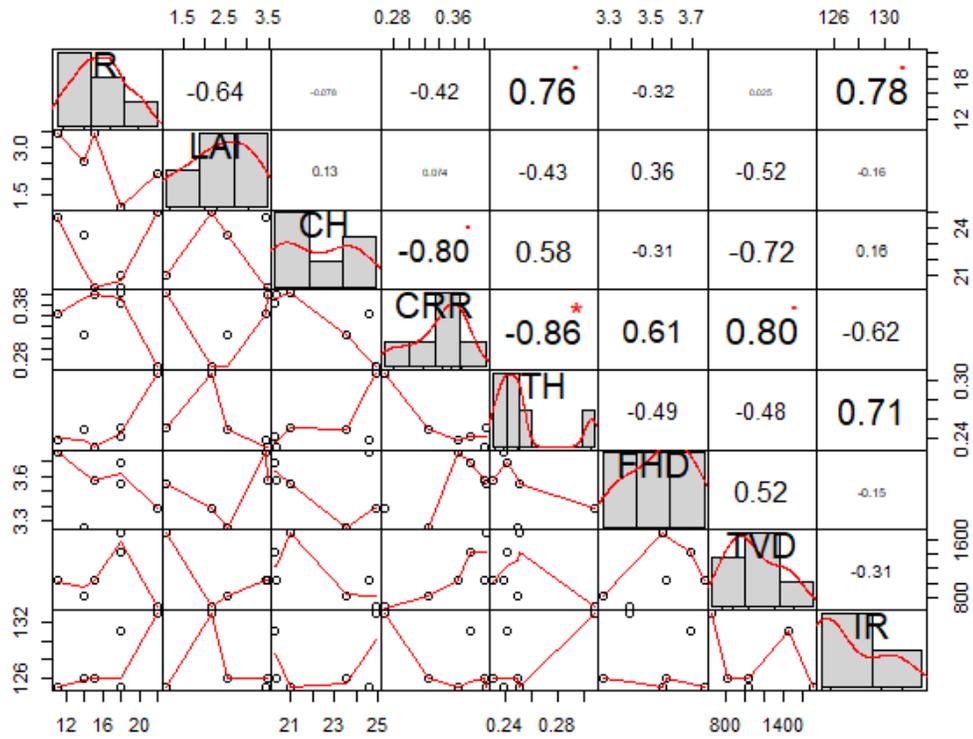


Figure 14: Density matrix display of richness and canopy variables. Significance signals for each coefficient value are illustrated by the asterisks in the upper portion of the matrix.

Abundance of individuals (A) was correlated with the texture of heights (TH,  $r = .59$ ) and the intensity of the returns (IR,  $r = .70$ ) (Figure 15).

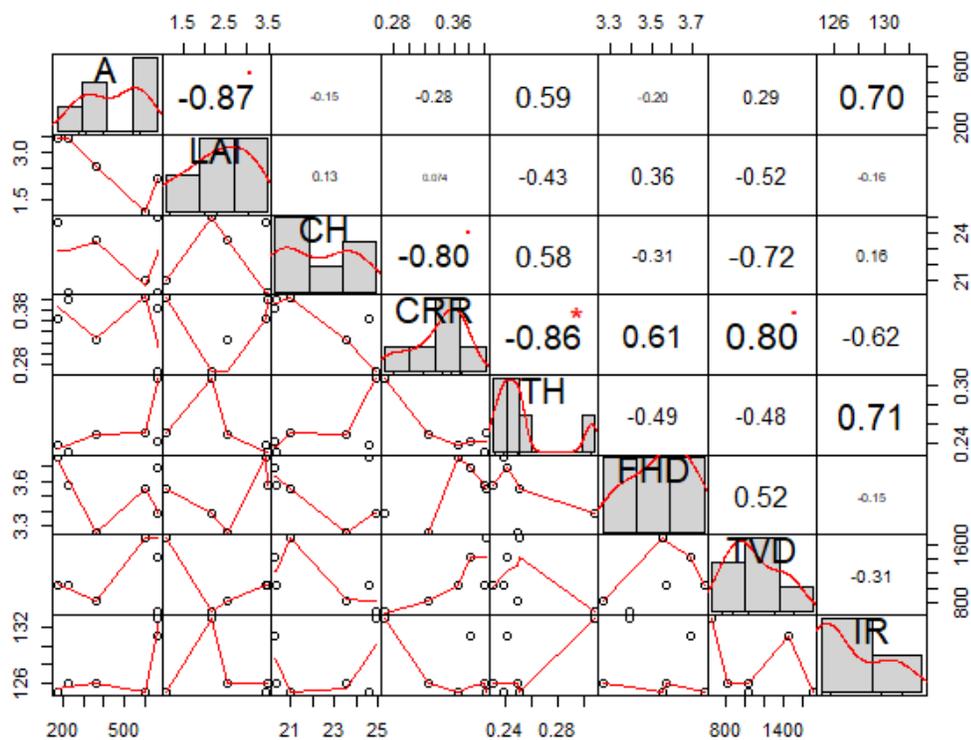


Figure 15: Density matrix display of abundance and canopy variables.

Diversity (MDI), was significantly correlated with leaf area (LAI,  $r = .69$ ). Given the size of the plot dataset, and the early seasonality of the multi-spectral image acquisition, additional plot data and peak ‘leaf-on’ imagery would assist in confirming the potential usefulness of this finding. Vegetation density was negatively correlated as well (TVD,  $r = -.43$ ) as shown in Figure 16.

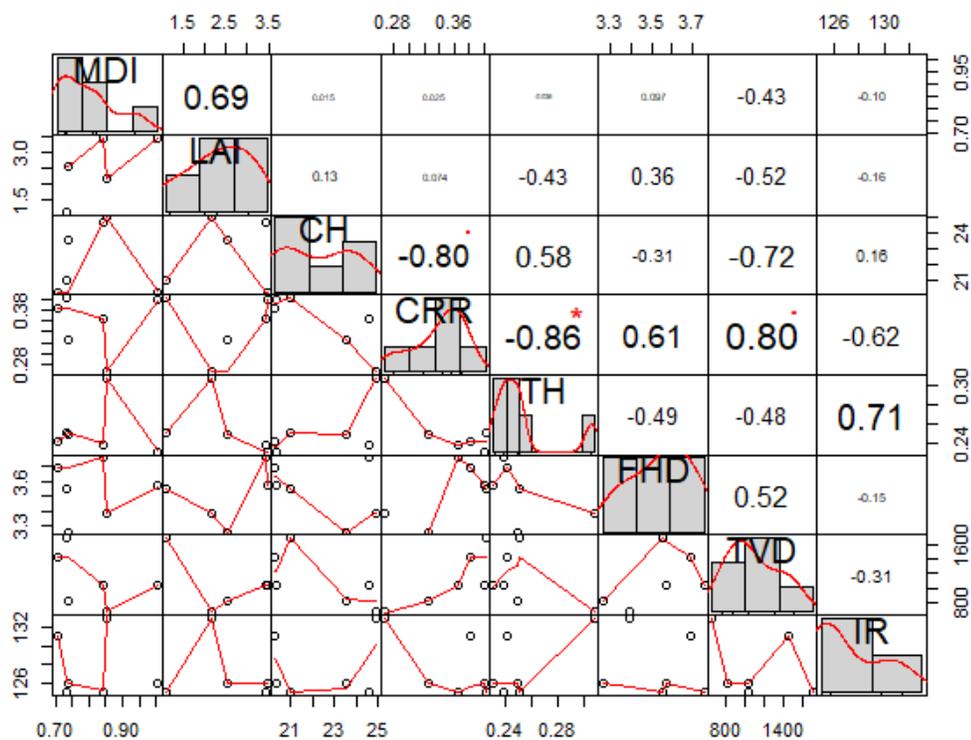


Figure 16: Density matrix display of MDI and canopy variables.

Total vegetation density (TVD), representative of the total LiDAR vegetation returns to ground returns, bore a negative relationship to MDI. This correlation was the second most significant in the available data. Although it is not immediately discernable what this relationship represents, the density is in one sense most representative of all predictor variables as an indicator of relative canopy closure.

Diversity adjusted for evenness (SDI), was modestly correlated with canopy height (CH,  $r = .42$ ) and the intensity of the returns (IR,  $r = .24$ ) within the available dataset (Figure 17).

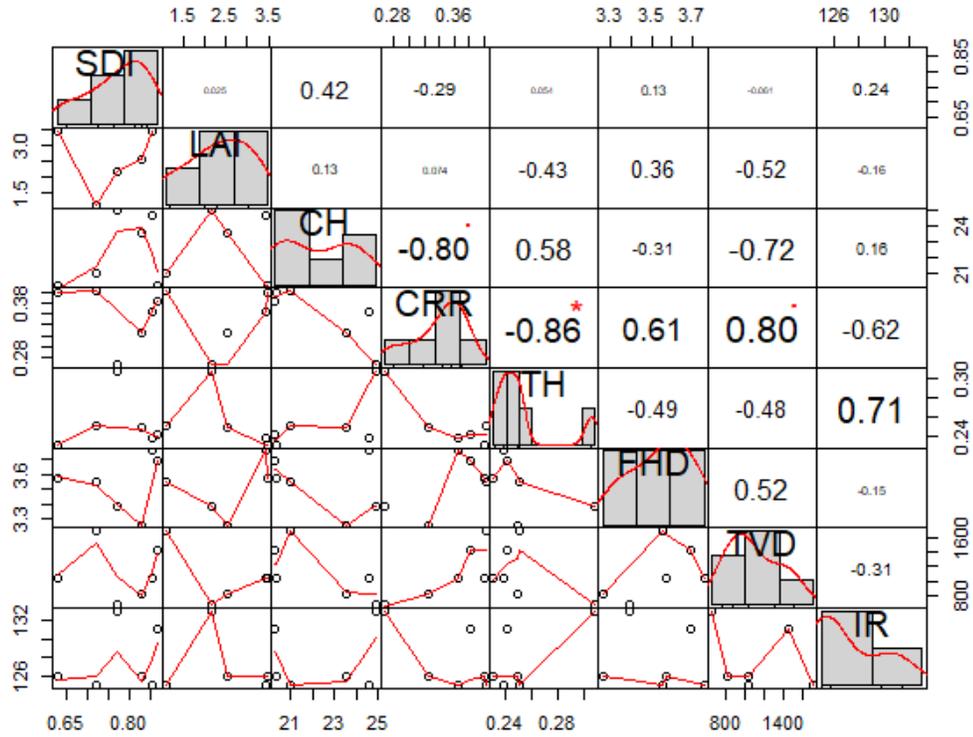


Figure 17: Density matrix display of SDI and canopy variables.

The regression model was created by first identifying the best variable subset using cross-validation statistics (the ‘regsubsets’ function in the R ‘leaps’ library), and then selecting predictors with high  $r^2$ . To summarize, the cross-validation statistical notation is:

$$\binom{p}{k} = \frac{p(p-1)}{k} \quad (15)$$

where

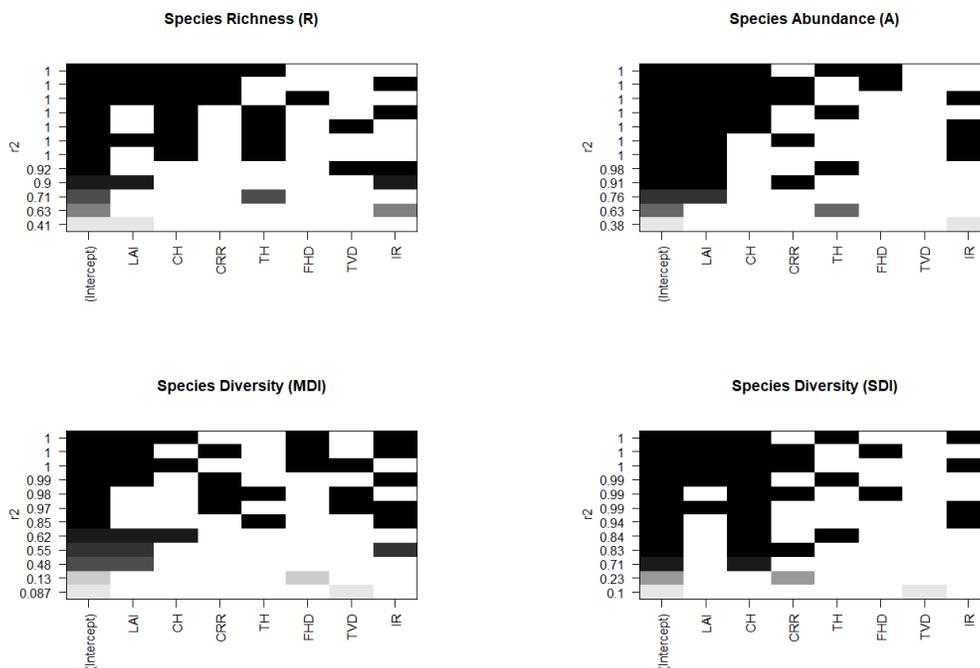
- $p$  is the specific predictor variable and
- $k$  is the total number of predictor variables.

If  $M_0$  is the null model containing 0 predictor values, the best model ( $M_k$ ) from a series of models, denoted as  $M_0 \dots M_p$ , is derived by fitting  $\binom{p}{k}$  models containing  $k$  predictors.

Biodiversity is analyzed as a function of all the variables in the canopy metrics data frame. The function evaluates the predictor variables using an exhaustive selection algorithm, selecting a maximum of four predictors for each variable subset.

The richness (R) cross-validation analysis is shown in Figure 16 below with the coefficient of determination ( $r^2$ ) values on the y-axis, canopy parameters on the x-axis and values indicating the variables included in the validation. Results indicate that the greater portion of the variance (.9) is associated with two of the variables. The abundance (A) results indicate a similar portion of the variance (.91) associated with two of the variables, but the variables differ.

The MDI diversity cross-validation indicates that much of the variance (.99) is associated with three of the variables, while the SDI diversity analysis indicates that much of the variance (.99) is attributed to three of the predictors and a similar proportion (.94) is associated with two of the variables. Complete results of each are shown in Figure 18.



**Figure 18: Model cross-validation analysis results for all predictor variables. Correlation coefficient is shown on the y-axis and predictor variables on the x-axis. The intercept is calculated as an output of the regression equation and is not significant for analysis.**

Using R linear modeling functions, we evaluated several variable subsets with significant  $r^2$  values to estimate the parameters of the linear model for SDI and MDI.

Figure 19 shows the results displayed for an SDI model with variables that have an  $r^2$  of .94. The p-values for each variable are insignificant (except for CH) and the model F statistic and p-value do not indicate the model itself is significant.

```
Lm0(formula = SDI ~ LAI + CH + IR, data = thesisModelSDI)
```

```
Residuals:
```

```
      1      2      4      5      6
0.0181328 0.0001679 -0.0057449 -0.0102682 -0.0022876
```

```
Coefficients:
```

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.743356  0.437239  3.987  0.1564
LAI         -0.021046  0.011619 -1.811  0.3211
CH           0.050570  0.006195  8.163  0.0776 .
IR          -0.016426  0.003853 -4.263  0.1467
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.02174 on 1 degrees of freedom
```

```
(1 observation deleted due to missingness)
```

```
Multiple R-squared:  0.9852,    Adjusted R-squared:  0.9409
```

```
F-statistic: 22.24 on 3 and 1 DF,  p-value: 0.1544
```

**Figure 19: Summary statistics for LAI + CH + IR model (Lm0).**

A second model with two variables having the best p-values from the first model yielded the following results (Figure 20).

```
Lm1(formula = SDI ~ CH + IR, data = thesisModelSDI)
```

```
Residuals:
```

```
      1      2      3      4      5      6
0.04041 -0.10677 0.10747 -0.01932 0.05052 -0.07230
```

```
Coefficients:
```

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.175880  1.745028 -0.101  0.926
CH           0.016505  0.021998  0.750  0.508
IR           0.004572  0.013737  0.333  0.761
```

```
Residual standard error: 0.1045 on 3 degrees of freedom
```

```
Multiple R-squared:  0.2049,    Adjusted R-squared: -0.3252
```

```
F-statistic: 0.3865 on 2 and 3 DF,  p-value: 0.709
```

**Figure 20: Summary statistics for CH + IR model (Lm1).**

The addition of the interaction term increases the total variance associated with the predictors to 0.693 as shown in Figure 21.

```
Lm2(formula = SDI ~ CH + IR +CH*IR, data = thesisModelSDI)

Residuals:
    1     2     3     4     5     6
0.015387 -0.049387  0.007445  0.045034 -0.017145 -0.001334

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -32.68231   9.85867  -3.315  0.0802 .
CH             1.41816   0.42369   3.347  0.0788 .
IR             0.25797   0.07686   3.356  0.0785 .
CH:IR        -0.01092   0.00330  -3.309  0.0804 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05027 on 2 degrees of freedom
Multiple R-squared:  0.8772,    Adjusted R-squared:  0.693
F-statistic: 4.763 on 3 and 2 DF,  p-value: 0.1784
```

Figure 21: Summary statistics for CH + IR + interaction model (Lm2).

An ANOVA was output to determine whether the improved fit was worth the additional model complexity. This method can be used to compare models with the same response variable. (Figure 22).

```
Analysis of Variance Table

Model 1: SDI ~ CH + IR
Model 2: SDI ~ CH + IR + CH * IR
  Res.Df  RSS Df Sum of Sq  F Pr(>F)
1     3 0.032735
2     2 0.005055  1  0.02768 10.951 0.08045 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 22: ANOVA R output for Lm1 and Lm2.

A Likelihood Ratio test (ANOVA) compares nested models increasing in complexity (Bolker 2008). The resulting test p-value does not indicate that the model as a whole is significant or should be used despite its improved  $r^2$  value and significance of the predictor variables in the Lm2 model summary.

Linear model 3 summary statistics (shown in Figure 23) yield the best data frame for predictive modeling given the available data.

```
Lm3(formula = MDI ~ CRR + TH + TVD, data = thesisModelMDI)
```

Residuals:

```
  1    2    3    4    5    6
-0.016272 -0.005572  0.005557 -0.007054  0.020468  0.002873
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-2.904e+00	4.080e-01	-7.116	0.01918 *
CRR	7.291e+00	6.777e-01	10.759	0.00853 **
TH	7.266e+00	8.754e-01	8.301	0.01421 *
TVD	-6.492e-04	5.175e-05	-12.544	0.00630 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02005 on 2 degrees of freedom

Multiple R-squared: 0.9875, Adjusted R-squared: 0.9687

F-statistic: 52.53 on 3 and 2 DF, p-value: 0.01874

**Figure 23: Summary statistics for CRR + TH + TVD model (Lm3).**

The model p-value is below .05 and the adjusted  $r^2$  value is greater than .96. In addition, each of the predictor variables contribute to the significance of the model.

The validity of linear regression modelling for prediction rests on four assumptions. They are that there is,

1. a linear relationship between the dependent and predictor variables,
2. the model errors are independent,
3. the model errors are normally distributed, and
4. the model errors have a constant variance with respect to the predictor variables (USDA 2016).

The plot in Figure 24 graphs the model errors (residuals) vs. the predicted values (fitted) and tests for nonlinearity as well as heteroscedacity (non-constant variance) in errors. The plotted points should be symmetrically distributed around zero (a horizontal line) indicating that the model doesn't make systematic errors (USDA 2016). The QQ plot (Figure 25) displays a satisfactory positive diagonal trendline for the available data.

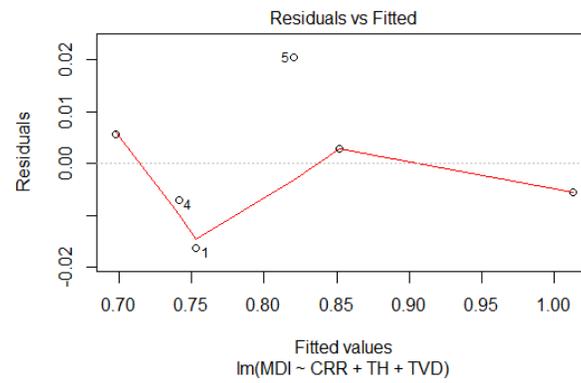


Figure 24: MDI linear model (Lm3) residuals vs. fitted values. Some row values are flagged for a discrepancy in the fitted vs. residual relationship (1,4-5).

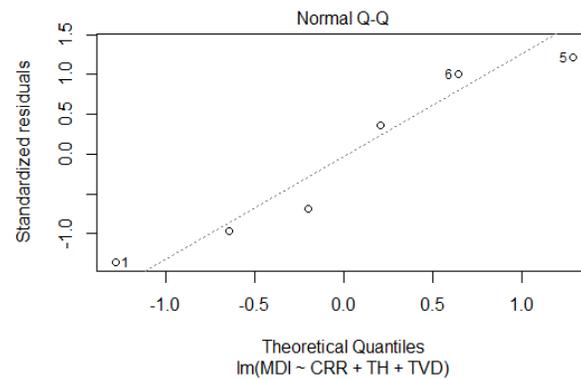


Figure 25: QQ plot of linear model errors (Lm3) Variables 1, 5-6 are flagged.

Coefficients from the model were used to display the prediction spatially over the study area extent as shown in Figure 26.

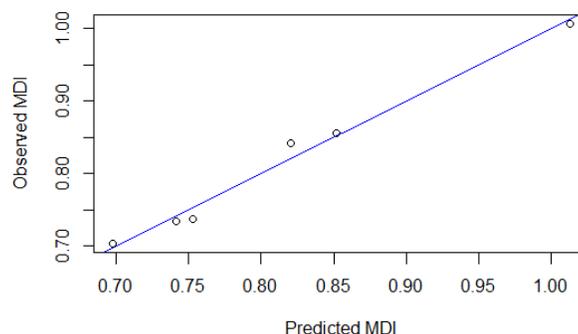


Figure 26: Graphic displaying predictive scatterplot for MDI (Lm3).

The prediction model can be used to develop a raster with the predictive data in a few steps.

**Coefficients from linear model:**

(Intercept)	CRR	TH	TVD
-2.9038666336	7.2912875982	7.2662259085	-0.0006491474

**Model equation for predicting MDI in R:**

$$\text{MDI} = -2.9038666336 + 7.2912875982 * \text{CRR} + 7.2662259085 * \text{TH} + (-0.0006491474) * \text{TVD}$$

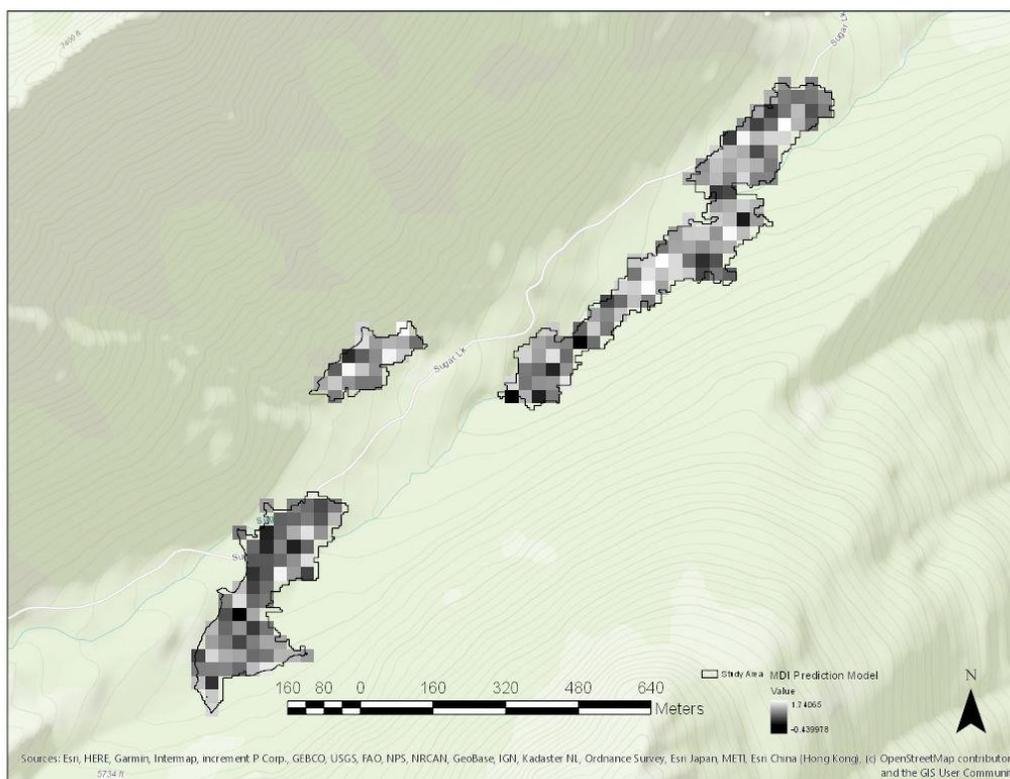
**Model equation formatted in map algebra for spatial prediction with gridded raster layers:**

$$\text{MDI} = -2.9038666336 + 7.2912875982 * \text{"CRR30mStat"} + 7.2662259085 * \text{"ToH30mStat"} + (-0.0006491474 * \text{"TotVegDen30mStat"})$$

Multicollinearity, such as that exhibited between CRR and TVD ( $r^2 = .80$ ), does make it difficult to assess the relative importance of independent variables if they are both used in the model, but it does not impact the usefulness of the regression equation for prediction. Even when multicollinearity is great, the least-squares regression equation can be highly predictive.

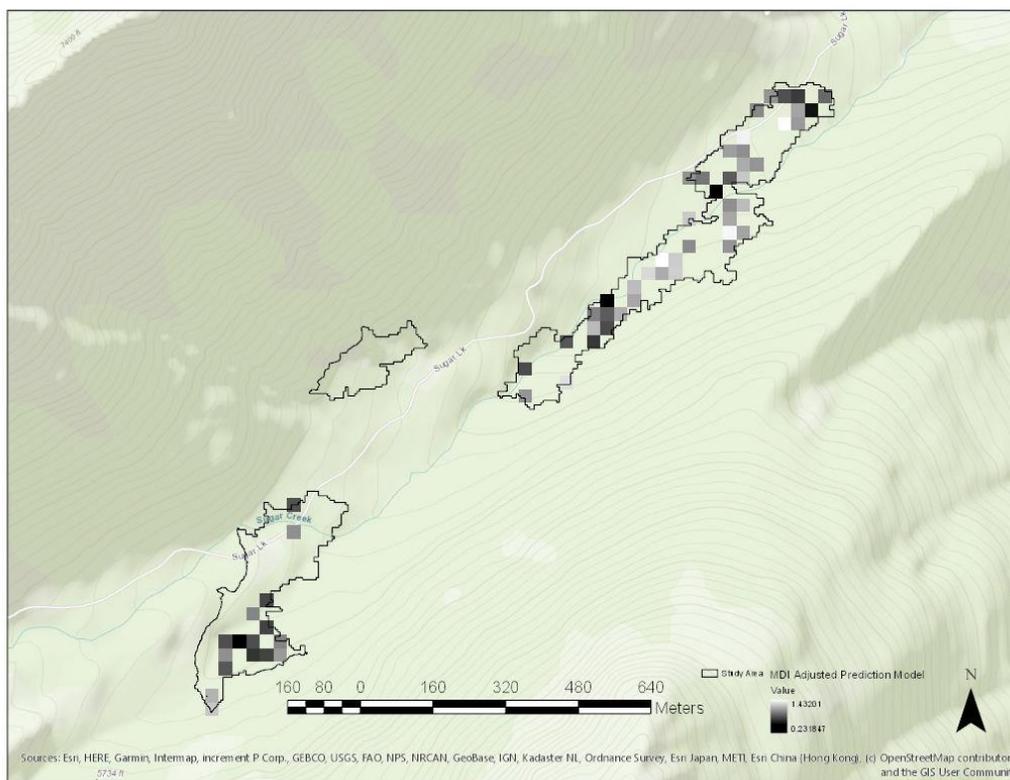
Figure 27 shows the results of the unadjusted prediction results mapped over the study area. Since predicting beyond the ranges of the original data will result in model

extrapolating and subsequent prediction errors, areas where the model extrapolated beyond the range of observation were ‘masked out’. Average CRR at a 30-meter resolution ranged in value from .27 to .40, the average TH from .23 to .31 and TVD from 656 to 1704 per meter square. The Con tool in ArcGIS creates a raster where ‘1’ is the assigned value if the canopy parameters are met, and if false, ‘0’ is the assigned value.



**Figure 27: Map depicting spatial prediction raster of MDI in 30-meter resolution.**

Combining three mask extrapolation layers – one for each variable – by deriving the product (using the Times tool) of raster layers with 0 or 1 pixel values creates an output raster consisting of pixels where variables are within the range of observations (a value of 1). Finally, this raster data is used to create a combined raster extrapolation mask applied to the original prediction file with the Set Null tool to derive a new predicted range of values (USDA 2016) shown in Figure 28.



**Figure 28: Map depicting spatial prediction raster of MDI in 30-meter resolution, adjusted for predictor variable ranges.**

Two global spatial (inferential) statistics were used to evaluate the raster surface values for spatial autocorrelation and high-low value clustering of the predicted data set. Global Moran's I statistic uses a z-score and p-value that test the null hypothesis that MDI is randomly distributed across the study area. Similarly, the Getis-Ord General G statistic uses a p-value and z-score to determine that there is no spatial clustering of feature values. However, whereas the Moran's I statistic evaluates high and low values together for purposes of spatial autocorrelation, the General G distinguishes between clustering of high and low values in the data set, determining which, if either, is significant and non-random. Therefore, if the p-value is statistically significant, the null hypothesis can be rejected, and the z-score is used to determine if the high or low values are contributing to the clustering.

The Global Moran's I statistic is denoted as,

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (16)$$

where

- $z_i$  is the deviation of an attribute for feature  $i$  from its mean,
- $w_{i,j}$  is the spatial weight between feature  $i$  and  $j$ ,
- $n$  is the number of features in the dataset and
- $S_0$  is the aggregate of spatial weights  $\sum_{i=1}^n \sum_{j=1}^n w_{i,j}$ . (Rogerson 2015)

The Getis-Ord General G is denoted as,

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j}, \forall_j \neq i \quad (17)$$

where

- $x_i$  and  $x_j$  are attribute values for features  $i$  and  $j$ ,
- $w_{i,j}$  is the spatial weight between feature  $i$  and  $j$ ,
- $n$  is the number of features in the dataset and
- $\forall_j \neq i$  indicates that features  $i$  and  $j$  cannot be the same feature.

A 100 meter Euclidian distance parameter was estimated for spatial statistics (although formal methods exist to derive this distance, this would add unnecessary complexity for the purposes of the current study as the assumption is within reason for sub-canopy surface light penetration).

**Table 8: Global Moran's I and General G Results for Predicted MDI**

Global Moran's I Summary		Getis-Ord General G Summary	
Moran's Index:	0.261149	Observed General G:	0.002234
Expected Index:	-0.016667	Expected General G:	0.002245
Variance:	0.004745	Variance:	0.000000
z-score:	4.033045	z-score:	-0.162268
p-value:	0.000055	p-value:	0.871095

These statistics indicate that the data as a whole are non-random with respect to spatial autocorrelation – the Global Moran's I p-value indicates significance, so the null

hypothesis is rejected. However, the Getis-Ord General G p-value indicates we should accept the null hypothesis that there is randomness in the data with respect to the clustering not being attributable to either high or low MDI biodiversity values (Getis & Ord 1992, Ord & Getis 1995).

The Getis-Ord  $G_i^*$  statistic can be used to determine statistically significant spatial clustering of data with an influence threshold distance. The resulting data can be used to create an MDI biodiversity hot-spot analysis map from the predicted raster values (Figure 29). ‘Hot spots derived from this statistic are both positive and negative, like the General G, and both types have significance for biodiversity prediction. It may be useful for a forest ecologist, for instance, to predict biodiversity ‘cold-spots’ as well as high biodiversity areas to see why this might be the case or for stand biodiversity and productivity improvement.

The notation for the local Gi-star (z-score) statistic is,

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{[n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2]}{n-1}}}$$
(18)

where  $x_j$  is the attribute for feature  $j$ ,  $w_{i,j}$  is the spatial weight between feature  $i$  and  $j$ ,  $n$  is the number of features in the dataset and:

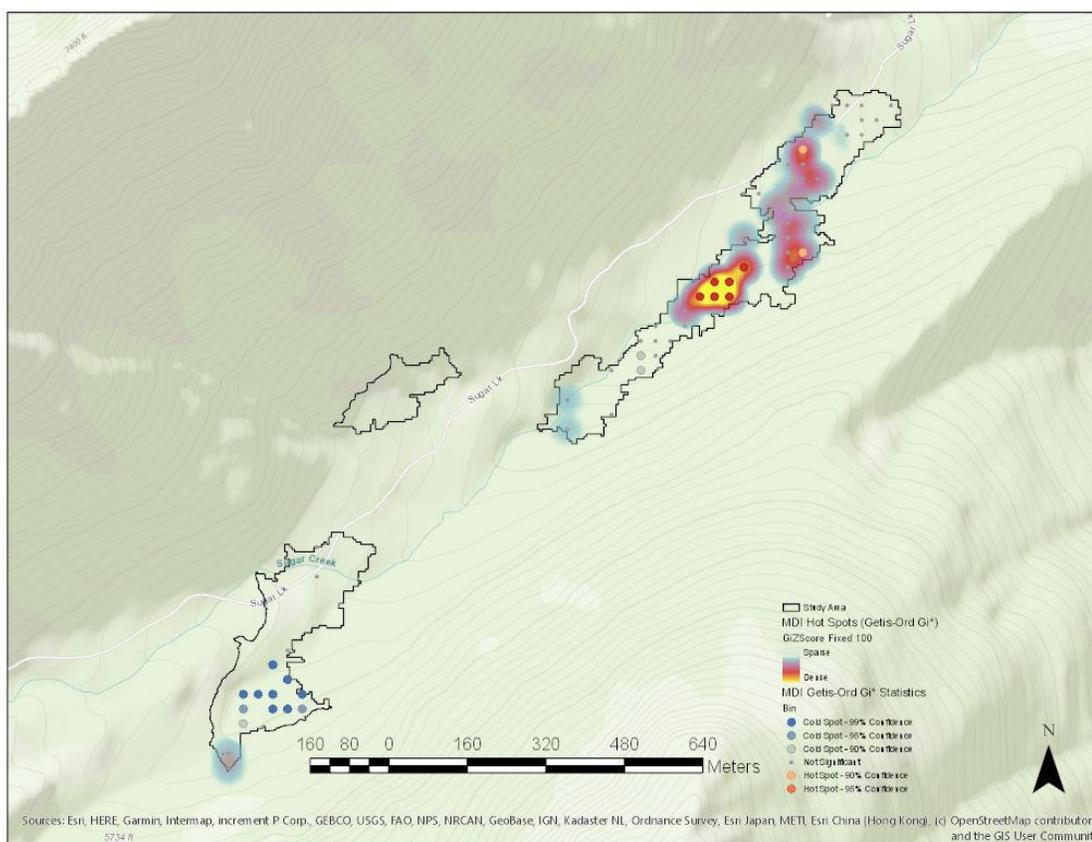
$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2}$$
(19)

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n}$$
(20)

(Rogerson 2015)

The statistic in this case used a fixed distance band of 100 meters to impose a “sphere of influence” or “moving window” conceptual model of spatial interactions onto the data.

Each feature is analyzed within the context of those neighboring features located within the distance you specify for the distance band and neighbors within the specified distance are weighted equally. The hot-spot map data frame contained 61 points each with 2-13 neighboring data points depending on its spatial position (Appendix 4).



**Figure 29: Map depicting spatial prediction of MDI hot spots based on Gi-star z-scores. High-value z-score hot spots are enhanced visually; both hot and cold statistics are displayed with confidence internals.**

## Discussion

Geospatial analysis of the radiant flux and canopy metrics focused on three general areas. 1.) exploratory analysis of collinearity between variables, 2.) selecting predictor variables for regression and predictive models, and 3.) examination of the predicted values with spatial statistics to determine if the predicated values were random. Regarding exploratory analysis, area solar radiation and canopy height exhibited a significant positive correlation (validating in some sense the radiation model with respect to the LiDAR canopy mensuration data). Statistically significant relationships were evident

between the canopy structure and biodiversity data as well. Prior to regression, a review of area solar radiation and biodiversity (MDI) data exhibited a modest association with an  $r$  value of .2, the only positive relationship of magnitude evident between radiant flux and biodiversity in the available data.

Of the four models analyzed, linear model (Lm3) comprising the canopy relief ratio (CRR), the texture of heights (TH) and total vegetation density (TVD) and its relationship to biodiversity (MDI) was the most significant given the available data. Interestingly, this set of predictors is intuitively descriptive of the category of metrics that would influence the variability of irradiance and the extinction of PAR in the lower strata. Unlike with canopy height (CH), there was no significant positive correlation between area radiation (ASR) and vegetation density (TVD). TVD's negative relationship to MDI may represent a spatial pattern whereby canopy density is associated with less diversity, much like patch dynamics in horizontal forest patterns generally suggest greater heterogeneity. I.e., canopy closure could suggest fewer, dominant species in the area (patch).

The MDI linear regression model had a p-value of 0.01874 and an adjusted  $r^2$  of 0.9687. As a predictor variable, CRR had a p-value of 0.00853 (a 0.01 level of significance), TH a p-value of 0.01412 (a 0.05 level of significance), and TVD a value of 0.00630 (a 0.01 level of significance). This model was used for study area prediction based on 144 data points in the base data set. It yielded over 60 data points after adjusting the prediction model for the range of canopy metrics in the specific plots that were surveyed.

The application of spatial statistics to the predicted data was the final component of the study. Global spatial statistics indicated the predicted range of MDI was not-random overall, and the local hot-spot map produced from this predicted data indicated several areas of both high and low biodiversity concentrations based on the prediction model data. The statistical significance of each of these hotspots varied from 90-99% based on the concentration of nearest neighbors within 100 meters of each analyzed data point. Schneider et al. (2017) undertook a similar study in a European conifer forest supported by the University of Zurich Research Priority Program on Global Change and Biodiversity (URPP GCB) and the working group of the National Center for Ecological Analysis and Synthesis on 'Prospects and priorities for satellite monitoring of global terrestrial biodiversity'. They employed a "spatially simultaneous autoregressive error model estimation based on first order neighbors" (using a spatial econometrics R package `spdep`, `errorsarm`) "to fit a generalized linear model". This allowed for the spatial autocorrelation to be applied in the linear model method itself and could be calculated for this study with the available data as an alternate method.

Although the plant biodiversity point-sample size was not as large as initially planned, the forest area sampled was nonetheless significant and varied in terms of area solar radiation and canopy structural diversity. Even in the relatively small number of plots surveyed, there were examples of hydric, mesic and dry-mesic soils which impacted vascular plant species abundance given the availability of water. Schneider et al. (2017) use abiotic factors to test for “patterns of functional traits and trait diversity” related to topography, soil type, water and incident solar radiation. This study does not analyze the same biogeographic scale or use the same diversity assumptions (although they are not contradictory), however, the use of these abiotic factors in the list of explanatory variables would prove useful in future research.

The model results provided some level of confidence that current or more topographically precise modeling methods might yield useful prediction models with additional point samples. However, more fieldwork is also required to determine the relationship of structural metrics used in the study to extinction ( $K$ ) to develop a more coherent understanding of canopy insolation partitioning of the radiant flux in the context of plant communities. These relationships, being more highly descriptive of understory PAR, would likely add more insight into the interpretation of modeled results of larger data sets.

Based on facts and probabilities in the technology innovation domain, it may be inferred that the increased availability of LiDAR and multi-spectral data acquired from UAVs, the quantity of products useful to practitioners may increase significantly as acquisition costs decrease in coming years – provided ALS has a data acquisition competitor in UAV for forest landscapes, this would likely impact LiDAR use trends. Even if satellite based acquisition becomes the norm for larger studies, there will be a need for small scale analysis. Additionally, the market for accurate biodiversity assessments has a basis in wildlife management and forest biometrics and is mandated under ecosystem services management requirements in some jurisdictions. One example of the use of this model is the implementation of biodiversity deficiency potential logic into long-term carbon sequestration and yield models that might prove advantageous if managing for multiple objectives, and lead to decreasing the effective fragmentation of horizontal structure across landscapes. It is possible that agroforestry professionals who seek to increase total yield or forestry fund managers looking for sustainable carbon offset investments would benefit from methods that build on this study. In any case, the development of a flexible and reasonably priced collection of tools for practitioners analyzing complex forest canopy models is within the means of a broad spectrum of researchers and institutions.

## References

- Ali, Arshad. 2019. Forest stand structure and functioning: Current knowledge and future challenges. *Ecological Indicators* 98: 665-677.
- Asner, G., J. Scurlock & J. Hicke. 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology & Biogeography* 12:191–205.
- Boise Center Aerospace Laboratory Lidar Tools version 5.2. Boise State University, Department of Geosciences, 1910 University Drive, Boise Idaho. Accessed April 30, 2017. URL: <http://bcsl.boisestate.edu/tools/lidar>
- Bolker, B. 2008. *Ecological Models and Data in R*. Princeton University Press, Princeton NJ.
- Chen, J. & Black, T. 1992. "Defining leaf area index for non-flat leaves". *Agricultural and Forest Meteorology* 57:1–12.
- Cornelissen, J., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D. Gurvich, P. Reich, H. ter Steege, H. Morgan, M. van der Heijden, J. Pausas, & H. Poorter. 2003. A handbook of protocols for standardized and easy measurement of plant functional traits worldwide. *Australian Journal of Botany* 51:335-380.
- DeSiervo, M., Jules, E., M. Kauffmann, D. Bost & R. Butz. 2016. Revisiting John Sawyer and Dale Thornburgh's 1969 Vegetation Plots in the Russian Wilderness: A Legacy Continued. *Fremontia* 44,1:20-25.
- DigitalGlobe. 2015. WorldView2 image 055613676010, DigitalGlobe, Longmont, Colorado, 5/4/2013.
- Donato, D., J. Fontaine, W. Robinson, J. Kauffman & B. Law. 2009. Vegetation response to a short interval between high-severity wildfires in a mixed-evergreen forest. *Journal of Ecology*. 97:142–154
- Dufour, A., F. Gadallah, H. Wagner, A. Guisan & A. Buttler. 2006. Plant species richness and environmental heterogeneity in a mountain landscape: effects of variability and spatial configuration. *Ecography* 29:573-584.

- Evans, J., A. Hudak, R. Faux, & A. Smith. 2009. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing* 1,4:776-794.
- Existing Vegetation-CALVEG, [ESRI geodatabase]. 2007. McClellan, CA: USDA-Forest Service, Pacific Southwest Region. EvgeFileGeodatabaseFeatureClassExistingVegNorCoastEast\_1998\_2007\_v1. [2017].
- Faith, D. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation*. 61:1–10.
- In: Fire Effects Information System, [Online]. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). Accessed May 2, 2019  
[URL:https://www.fs.fed.us/database/feis/plants/tree/pinlam/all.html](https://www.fs.fed.us/database/feis/plants/tree/pinlam/all.html)
- Fu, P. & P. Rich. 2000. *The Solar Analyst 1.0 Manual*. Helios Environmental Modeling Institute (HEMI), USA.
- Fu, P. & P. Rich. 2002. A Geometric Solar Radiation Model with Applications in Agriculture and Forestry. *Computers and Electronics in Agriculture* 37:25–35.
- Gao, T., M. Hedblom, T. Emilsson & A. Nielsen. 2014. The role of forest stand structure as biodiversity indicator. *Forest Ecology and Management* 330:82-93.
- Geiger, R. 1965. *The Climate Near the Ground*. Cambridge (MA): Harvard University Press.
- Getis, A. & J. Ord. 1992. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis* 24:3.
- Guisan, A. & N. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147–186.
- Hamilton, R., S. Cushman, K. McCallum, N. McCusker, T. Mellin, M. Nigrelli & M. Williamson. 2013. Multiscale landscape pattern monitoring using remote sensing: The four forest restoration initiative. RSAC-10022-RPT1. Salt Lake City, UT:

U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 24 p.

Hearne, L. 2008. "Normalized Difference Vegetation Index (NDVI) Analysis for Forestry and Crop Management" SimWright Website, May. 2019, [http://simwright.com/downloads/SimWright\\_NDVI.pdf](http://simwright.com/downloads/SimWright_NDVI.pdf) Slide 2.

Hillenbrand, H. 2004. "On the generality of the latitudinal diversity gradient". *The American Naturalist* 163,2:192-211.

Hao, M., C. Zhang, X. Zhao & K. von Gadow. 2018. Functional and phylogenetic diversity determine woody productivity in a temperate forest. *Ecology and Evolution*. 2018,1–12

von Humboldt, A. & A. Bonpland. 1807. *Essai sur la géographie des plantes*. Paris

Ishii, H., S. Tanabe & T. Hiura. 2004. Exploring the Relationships Among Canopy Structure, Stand Productivity and Biodiversity of Temperate Forest Ecosystems. *Forest Science* 50,3:342.

Kim, Y., Z. Q. Yang, W. B. Cohen, D. Pflugmacher, C. L. Lauver, & J. L. Vankat. 2009. "Distinguishing between Live and Dead Standing Tree Biomass on the North Rim of Grand Canyon National Park, USA Using Small-Footprint Lidar Data." *Remote Sensing of Environment* 113:2499–510.

Küchler, A. 1977. Appendix: the map of the natural vegetation of California. Pages 909-938 In: M. G. Barbour and J. Major, eds, *Terrestrial vegetation of California*. John Wiley and Sons, New York.

Kauffmann, M. "Miracle Mile." *Plant Explorations*, 23 Oct. 2014, URL:[blog.conifercountry.com/2014/07/miracle-mile-russian-wilderness/](http://blog.conifercountry.com/2014/07/miracle-mile-russian-wilderness/).

Keeler-Wolf, T. 1990. *Ecological surveys of Forest Service Research Natural Areas in California*. Gen. Tech. Rep. PSW-125. Berkeley, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 177 p.

Leiterer, R., R. Furrer, M. Schaepman, & F. Morsdorf. 2015. Forest canopy-structure characterization: A data-driven approach. *Forest Ecology and Management* 358:48–61.

- "Leaf Area Index - Fraction of Photosynthetically Active Radiation 8-Day L4 Global 1km." MOD15A2 | LP DAAC NASA Land Data Products and Services. [https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/mod15a2](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod15a2) Accessed December 02, 2017.
- Laureto, L., M. Cianciaruso & D. Samia. 2015. Functional diversity: an overview of its history and applicability. *Natureza & Conservação* 13:122-116.
- Levin, S. 1992. The problem of pattern and scale in ecology. *Ecology* 73:1943-1967.
- Nunez, M. 1980. The calculation of solar and net radiation in mountainous terrain. *Journal of Biogeography* 7:173-186.
- Oliphant, A., C. Susan, B. Grimmond, H-P. Schmid & C. Wayson. 2006. Local-scale heterogeneity of photosynthetically active radiation (PAR), absorbed PAR and net radiation as a function of topography, sky conditions and leaf area index. *Remote Sensing of Environment* 103:324–337.
- Olseth, J. & A. Skartveit. 1997 Spatial distribution of photosynthetically active radiation over complex topography. *Agricultural and Forest Meteorology* 86:205-214.
- Olson, D., R. Noss, G. Orians, J. Stritholt, C. Williams, & J. Sawyer. "Klamath-Siskiyou." WWF. Accessed December 01, 2016. URL:<http://www.worldwildlife.org/ecoregions/na0516>.
- Ord, J. & A. Getis. 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* 27:4.
- Ozanne, C. D. Anhuf, S. Boulter, M. Keller., R. Kitching, C. Körner, F. Melnzer, A. Mitchell, T. Nakashizuka, P. Silva Dias, N. Stork, S. Wright & M. Yoshimura. 2003. Biodiversity Meets the Atmosphere: A Global View of Forest Canopies. *Science* 301:183-186.
- Pan, Y., J. M. Chen, R. Birdsey, K. McCullough, L. He & F. Deng. 2011. Age structure and disturbance legacy of North American forests. *Biogeosciences* 8:715-732.
- Pan, Y., R. Birdsey, O. Phillips & R. Jackson. 2013. The Structure, Distribution, and Biomass of the World's Forests. *Annual Review of Ecology, Evolution, and Systematics* 44:593.

- Parker, G. & M. Russ. 2004. The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. *Forest Ecology and Management* 189:307-315.
- Paquette, A. & C. Messier. 2015. The effect of biodiversity on tree productivity: from temperate to boreal forests. *Global Ecology and Biogeography* 20:170-180.
- Perry, D., R. Oren, & S. Hart. 2008. *Forest ecosystems*. Baltimore: Johns Hopkins University Press.
- Petchey, O., E. O'Gorman & D. Flynn. 2009. A functional guide to functional diversity measures. In: *Biodiversity, Ecosystem Functioning and Human Wellbeing: An Ecological and Economic Perspective*. Oxford University Press
- Pike, R.J. & S. Wilson. 1971. Elevation-Relief Ratio, Hypsometric Integral and Geomorphic Area-Altitude Analysis. *Geological Society of America Bulletin* 82:1079-1084.
- Pla L., F. Casanoves & J. Di Rienzo. 2012. Functional Diversity Indexes. In: *Quantifying Functional Biodiversity*. Springer Briefs in Environmental Science. Springer, Dordrecht
- QSI Environmental. 2015. Technical Data Report-USFS PSW Region 5 LiDAR. Accessed October 2016.
- Rogerson, P. 2015. *Statistical methods for geography*. Los Angeles: Sage.
- Saitoh, T., N. Shin, H. Noda, H. Muraoka & K. Nasahara. 2012. Examination of the extinction coefficient in the Beer-Lambert law for an accurate estimation of the forest canopy leaf area index. *Forest Science and Technology* 8,2:67-76
- Sawyer, J., D. Thornburgh & J. Griffin. 1977. Mixed evergreen forest. Pages 359-382 in M. Barbour and J. Major, Eds, *Terrestrial Vegetation of California*. John Wiley and Sons, New York.
- Sawyer, J. & D. Thornburgh. 1974. Sub-alpine and Montane Forests on Granodiorite in the Central Klamath Mountains of California. Unpubl. Rep. to Pacific Southwest Forest and Range Exp. Sta., Berkeley, Calif.
- Sawyer, J. 2006. *Northwest California: A natural history*. Berkeley: University of California Press.

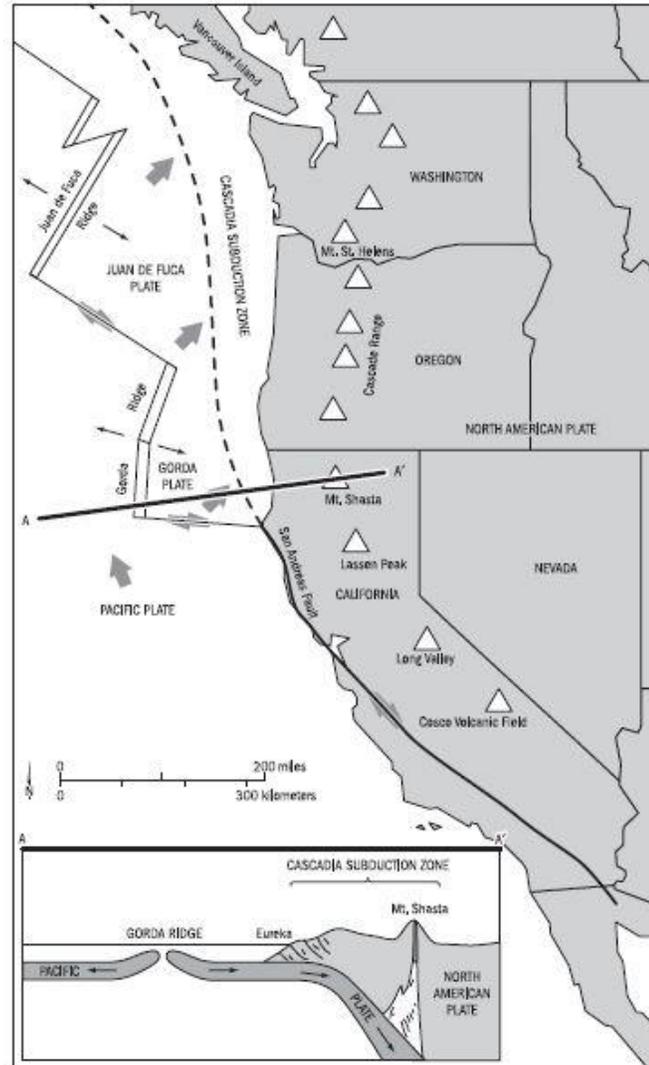
- Schneider, F., F. Morsdorf, B. Schmid, O. Petchey, A. Hueni, D. Schimel, & M. Schaepman. 2017. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications* 8:1441.
- Schoonmaker, P., A. McKee. 1988. Species Composition and Diversity During Secondary Succession of Coniferous Forests in the Western Cascade Mountains of Oregon. *Forest Science* 34,4:960-979.
- Schulze, E-D. 2005. *Plant ecology*. Berlin: Springer.
- Skinner, C., A. Taylor, & J. Agee. 2006. Klamath Mountains bioregion. In: Fire in California's Ecosystems. Edited by N.G. Sugihara, J.W. van Wagtendonk, J. Fites-Kaufman, K. Shaffer, A. Thode. University of California Press, Berkeley. pp. 170-194.
- Smith, M-L., J. Anderson, & M. Fladeland. 2008 Forest Canopy Structural Properties. In: Field Measurements for Forest Carbon Monitoring. Edited by C.M. Hoover., Springer Science+Business Media B.V. pp.179-196.
- Thorson, T., S. Bryce, D. Lammers, A. Woods, J. Omernik, J. Kagan, D. Pater, & J. Comstock. 2003. Ecoregions of Oregon (color poster with map, descriptive text, summary tables, and photographs): Reston, Virginia, U.S. Geological Survey (map scale 1:1,500,000).
- USDA. 2016. Technical Report-USFS Remote Sensing Applications Center. Accessed November 2018.
- USDA. 1982. Soil Survey Klamath National Forest Area California. USFS PSW Region. Accessed November 2019.
- Whittaker, R. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- 1972. Evolution and Measurement of Species Diversity. *Taxon* 21:213-251.
- Wikipedia contributors, "Attenuation coefficient," *Wikipedia, The Free Encyclopedia*, [https://en.wikipedia.org/w/index.php?title=Attenuation\\_coefficient&oldid=925582822](https://en.wikipedia.org/w/index.php?title=Attenuation_coefficient&oldid=925582822) Accessed November 25, 2019.

- Wulder, M., E. LeDrew, S. Franklin & M. Lavigne. 1998. Aerial image texture information in the estimation of northern deciduous and mixed wood forest leaf area index (LAI), *Remote Sensing of Environment* 64:64–76.
- Zhang, L., Z. Hu, J. Fan, D. Zhou, & F Tang. 2014. A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems. *Frontiers of Earth Science* 8,4:599-609.
- Zhang, Y., H. Chen, & P. Reich. 2012. Forest productivity increases with evenness, species richness and trait variation: a global meta-analysis. *Journal of Ecology* 100:742–749.

## Appendices

### Appendix 1.

Plate tectonic setting for the Klamath Ecoregion. A—A' indicates the location of the profile of the Gorda Plate subduction under the North American Plate between Eureka and Mount Shasta, the arc in direct proximity to the study area near Russian Peak (Sawyer 2006).



## Appendix 2.

### LiDAR Data Specification

- Projection: UTM Zone 10 North
- Horizontal Datum: NAD83 (CORS96)
- Vertical Datum: NAVD88 (GEOID03)
- Units: Meters
- LAS v 1.2 - All Returns (.laz compressed)

- Raster (1.0 Meter ENVI .dat files)
  - Bare Earth Model (1/4 USGS quads)
  - Highest Hit Model (1/4 USGS quads) 0.5 Meter ENVI .dat files
  - Normalized Intensity Images (1/4 USGS quads)

#### LiDAR Acquisition Specification

- Survey Altitude (AGL) 900 meters
- Target Pulse Rate 600 kHz
- Pulse Mode Single Pulse in Air (SPiA)
- Laser Pulse Diameter 21 cm
- Mirror Scan Rate 52.0 Hz
- Field of View 30 degrees
- GPS Baselines less than or equal to 13 nm
- GPS Position Dilution of Precision (PDOP) less than or equal to 3.0
- Intensity 8-bit
- Resolution/Density Average 8 pulses/m<sup>2</sup>
- Accuracy RMSEZ  $\leq$  15 cm

## Appendix 3.

Biodiversity Sample Plot Data

Plot: #32

Map Coordinates: 10T 507793 4572949

GNSS correction method: SBAS

Estimated horizontal accuracy: 2 meters (approximate)

Elevation: N/A

Densitometer reading at plot center:  $100 - (20 * 1.04) = 79.2\%$  canopy closure

Note: Lower strata abundance concentrated in mesic soils; xeric soils also observed on plot.

Upper and mid-strata species quantity estimate:

<i>Abies concolor</i>	50
<i>Pinus contorta</i>	5
<i>Pseudotsuga menziesii</i>	15
<i>Calocedrus decurrens</i>	5
<i>Pinus ponderosa</i>	9
<i>Letharia vulpine</i>	100

Lower-strata species quantity estimate:

<i>Abies concolor</i>	30
<i>Pseudotsuga menziesii</i>	10
<i>Picea engelmannii</i>	3
<i>Chrysolepi sempervirens</i>	40
<i>Chimaphila umbellata</i>	30
<i>Adenocaulon bicolor</i>	10
<i>Linnaea borealis</i> var. <i>longiflora</i>	50
<i>Rosa gymnocarpa</i>	2
<i>Prosartes hookeri</i>	1
<i>Prunus emarginata</i>	1

Plot: #30

Map Coordinates: 10T 507733 4572919

GNSS correction method: SBAS

Estimated horizontal accuracy: 4 meters (approximate)

Elevation: 1521 HAE

Densitometer reading at plot center:  $100 - (49 * 1.04) = 49.0\%$  canopy closure

Upper and mid-strata species quantity estimate:

<i>Abies concolor</i>	4
<i>Pinus monticola</i>	1
<i>Pseudotsuga menziesii</i>	1
<i>Calocedrus decurrens</i>	1
<i>Letharia vulpine</i>	2
<i>Lobaria hallii</i>	15

Lower-strata species quantity estimate:

<i>Abies concolor</i>	2
<i>Chrysolepi sempervirens</i>	6
<i>Chimaphila umbellata</i>	130
<i>Adenocaulon bicolor</i>	1
<i>Linnaea borealis var. longiflora</i>	30
<i>Rosa gymnocarpa</i>	4
<i>Prosartes hookeri</i>	7
<i>Pteridium aquilinum</i>	5
<i>Galium triflorum</i>	9
<i>Apocynum androsaemifolium</i>	4

Plot: #26

Map Coordinates: 10T 507584 4572768

GNSS correction method: SBAS

Estimated horizontal accuracy: 1.5 meters (approximate)

Elevation: 1550 HAE

Densitometer reading at plot center:  $100 - (12 * 1.04) = 87.5\%$  canopy closure

Upper and mid-strata species quantity estimate:

<i>Abies concolor</i>	18
<i>Pseudotsuga menziesii</i>	1
<i>Picea engelmannii</i>	11
<i>Letharia vulpine</i>	2
<i>Lobaria hallii</i>	2

Lower-strata species quantity estimate:

<i>Abies concolor</i>	28
<i>Pseudotsuga menziesii</i>	10
<i>Calocedrus decurrens</i>	3
<i>Picea engelmannii</i>	2
<i>Chrysolepi sempervirens</i>	3
<i>Chimaphila umbellata</i>	175
<i>Adenocaulon bicolor</i>	40
<i>Linnaea borealis</i> var. <i>longiflora</i>	50
<i>Rosa gymnocarpa</i>	10
<i>Prosartes hookeri</i>	80
<i>Pteridium aquilinum</i>	20
<i>Clintonia uniflora</i>	5
<i>Galium triflorium</i>	50
<i>Apocynum androsaemifolium</i>	25
<i>Calamagrostis koelerioides</i>	10
<i>Letharia vulpine</i>	10
Unidentified plant	100

Plot: #29

Map Coordinates: 10T 507613 4572888

GNSS correction method: SBAS

Estimated horizontal accuracy: 1.5 meters (approximate)

Elevation: 1543 HAE

Densitometer reading at plot center:  $100 - (12 * 1.04) = 87.5\%$  canopy closure

Upper and mid-strata species quantity estimate:

<i>Pseudotsuga menziesii</i>	18
<i>Calocedrus decurrens</i>	2
<i>Pinus ponderosa</i>	1
<i>Lobaria hallii</i>	5

Lower-strata species quantity estimate:

<i>Abies concolor</i>	35
<i>Pseudotsuga menziesii</i>	13
<i>Calocedrus decurrens</i>	1
<i>Chimaphila umbellata</i>	50
<i>Adenocaulon bicolor</i>	25
<i>Linnaea borealis var. longiflora</i>	50
<i>Rosa gymnocarpa</i>	5
<i>Prosartes hookeri</i>	30
<i>Pteridium aquilinum</i>	1
<i>Clintonia uniflora</i>	25
<i>Rubus parviflorus</i>	2
<i>Galium triflorum</i>	10
<i>Prunus emarginata</i>	1
<i>Apocynum androsaemifolium</i>	25
<i>Letharia vulpine</i>	1
<i>Lobaria hallii</i>	1
Unidentified plant	300

Plot: #25

Map Coordinates: 10T 507524 4572768

GNSS correction method: SBAS

Estimated horizontal accuracy: 3.5 meters (approximate)

Elevation: 1530 HAE

Densitometer reading at plot center:  $100 - (17 * 1.04) = 82.3\%$  canopy closure

Upper and mid-strata species quantity estimate:

<i>Abies concolor</i>	2
<i>Pseudotsuga menziesii</i>	11
<i>Calocedrus decurrens</i>	1
<i>Pinus ponderosa</i>	1
<i>Letharia vulpine</i>	25

Lower-strata species quantity estimate:

<i>Abies concolor</i>	26
<i>Pseudotsuga menziesii</i>	1
<i>Calocedrus decurrens</i>	3
<i>Chrysolepi sempervirens</i>	5
<i>Chimaphila umbellate</i>	20
<i>Adenocaulon bicolor</i>	25
<i>Berberis aquifolium</i>	1
<i>Rosa gymnocarpa</i>	40
<i>Prosartes hookeri</i>	10

Plot: #23

Map Coordinates: 10T 507433 4572530

GNSS correction method: SBAS

Estimated horizontal accuracy: 1 meter (approximate)

Elevation: 1569 HAE

Densitometer reading at plot center:  $100 - (50 * 1.04) = 48\%$  canopy closure

Note: Plot soil hydrology is hydric.

Upper and mid-strata species quantity estimate:

<i>Abies concolor</i>	2
<i>Pseudotsuga menziesii</i>	4
<i>Picea engelmannii</i>	5
<i>Pinus ponderosa</i>	1
<i>Letharia vulpine</i>	25

Lower-strata species quantity estimate:

<i>Abies concolor</i>	70
<i>Picea engelmannii</i>	2
<i>Chrysolepi sempervirens</i>	1
<i>Prunus emarginata</i>	2
<i>Rubus parviflorus</i>	5
<i>Adenocaulon bicolor</i>	45
<i>Amelanchier pumila</i>	2
<i>Rosa gymnocarpa</i>	6
<i>Chimaphila umbellata</i>	25
<i>Pteridium aquilinum</i>	75
<i>Clintonia uniflora</i>	25
<i>Lilium columbianum</i>	5
<i>Galium triflorum</i>	25
<i>Apocynum androsaemifolium</i>	25
<i>Acer macrophyllum</i>	1
<i>Symphoricarpos hesperius</i>	2
<i>Veratrum californicum</i>	3
<i>Alnus incana</i>	6
<i>Linnaea borealis</i> var. <i>longiflora</i>	300

## Appendix 4.

MDI, Z-score and p-values for data points used to produce the local Gi-star map.

POINT ID	MDI VALUE	GI* Z-SCORE	GI*P-VALUE	N NEIGHBORS
1	0.950523436	-0.379137603	0.704585685	8
2	0.64489764	-0.536237105	0.591794695	9
3	0.551392734	-0.536237105	0.591794695	9
4	0.753116131	-1.57725639	0.114736542	6
5	0.895265818	0.945622418	0.344341217	9
6	1.012304306	-0.536237105	0.591794695	9
7	0.374205112	-0.571862952	0.567414827	8
8	1.432014585	0.239518571	0.810703498	9
9	0.971790373	-0.536237105	0.591794695	9
10	1.354251146	1.880310147	0.060065821	9
11	0.992714345	0.235316107	0.813963394	9
12	0.921366751	1.239257339	0.215250236	8
13	1.047469139	0.532882945	0.594114614	11
14	0.946267366	1.53915907	0.123765467	8
15	0.934375107	-0.968195169	0.332946919	5
16	0.820723474	0.002165397	0.998272265	9
17	0.697763085	0.707607701	0.479188914	13
18	1.170986414	0.676712505	0.498588386	12
19	0.231846929	0.983866671	0.325181086	12
20	0.926844597	1.027307879	0.304275528	12
21	1.116466284	0.865254017	0.386899392	11
22	1.17999649	0.766481926	0.443389609	9
23	1.029138923	0.991587646	0.321398731	10
24	1.37430346	0.942997484	0.345682201	9
25	1.014053702	1.655980718	0.097725724	6
26	0.908354163	2.914650473	0.003560871	8
27	0.950768173	1.57505523	0.115243697	7
28	1.411396027	2.726129093	0.006408193	7
29	1.189407229	2.810175623	0.004951447	7
30	1.243346453	2.898078894	0.003754561	7
31	1.048261285	2.76259358	0.005734412	8
32	1.199327826	2.726129093	0.006408193	7
33	1.121464968	1.436313129	0.15091326	9
34	0.301283121	-0.834310748	0.404105905	9

35	1.028103113	-0.004708383	0.996243267	9
36	0.867261827	-1.408520534	0.158976991	9
37	0.698297679	-0.834310748	0.404105905	9
38	1.062563777	-0.834310748	0.404105905	9
39	1.107253671	-1.6683891	0.095238511	8
40	0.626810849	-1.408520534	0.158976991	9
41	0.676508367	-0.368937706	0.712174158	6
42	0.541493118	-1.6683891	0.095238511	8
43	0.62604028	0.388567646	0.697596005	3
44	1.274631977	-0.053431365	0.957388216	4
45	0.949938536	0.388567646	0.697596005	3
46	0.656762123	-0.509226676	0.610593349	2
47	0.92310524	-0.509226676	0.610593349	2
48	0.60632205	-1.420676396	0.155410858	5
49	0.881251276	-3.04520218	0.002325239	9
50	0.59173876	-2.811447688	0.004931911	10
51	0.638733983	-2.890760729	0.003843106	9
52	0.296034575	-2.576138402	0.009991065	11
53	0.634688079	-2.816957101	0.0048481	12
54	0.879496157	-2.61192181	0.009003484	9
55	0.925068915	-2.363721262	0.018092419	8
56	0.523034692	-2.576138402	0.009991065	11
57	0.590007246	-2.576138402	0.009991065	11
58	0.974484026	-2.489235818	0.012801802	7
59	0.680426359	-1.886055345	0.059287497	9
60	1.080617905	0.556541235	0.577840916	4
61	1.170920849	0.560747311	0.574969813	3