

Article

Modeling Insolation, Multi-Spectral Imagery and LiDAR Point-Cloud Metrics to Predict Plant Diversity in a Temperate Montane Forest

Paul Christian Dunn * and Leonhard Blesius

Department of Geography and Environment, San Francisco State University, San Francisco, CA 94132, USA; lblesius@sfsu.edu

* Correspondence: pchristiandunn@alumni.sfsu.edu

Abstract: Incident solar radiation (insolation) passing through the forest canopy to the ground surface is either absorbed or scattered. This phenomenon, known as radiation attenuation, is measured using the extinction coefficient (K). The amount of radiation reaching the ground surface of a given site is effectively controlled by the canopy's surface and structure, determining its suitability for plant species. Menhinick's and Simpson's biodiversity indexes were selected as spatially explicit response variables for the regression equation using canopy structure metrics as predictors. Independent variables include modeled area solar radiation, LiDAR-derived canopy height, effective leaf area index data derived from multi-spectral imagery and canopy strata metrics derived from LiDAR point-cloud data. The results support the hypothesis that (1) canopy surface and strata variability may be associated with understory species diversity due to radiation attenuation and the resultant habitat partitioning and that, (2) such a model can predict both this relationship and biodiversity clustering. The study data yielded significant correlations between predictor and response variables and were used to produce a multiple-linear model comprising canopy relief, the texture of heights, and vegetation density to predict understory plant diversity. When analyzed for spatial autocorrelation, the predicted biodiversity data exhibited non-random spatial continuity.

Keywords: biodiversity; insolation; biogeography; lidar; point-cloud; multi-spectral imagery; spatial prediction model; forest canopy



check for updates

Citation: Dunn, P.C.; Blesius, L. Modeling Insolation, Multi-Spectral Imagery and LiDAR Point-Cloud Metrics to Predict Plant Diversity in a Temperate Montane Forest. *Geographies* **2021**, *1*, 79–103. <https://doi.org/10.3390/geographies1020006>

Academic Editor: Jorge Rocha

Received: 2 August 2021

Accepted: 18 August 2021

Published: 23 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Spatial ecologists and biogeographers study the variation of plant species across spatial scales and latitudinal gradients. This spatial variability and its link to ecological and biogeochemical processes are considered by some researchers to be fundamental to biological inquiry [1] and understanding biological diversity a central research problem in modern biology. The spatial analysis of biota often focuses on either species distribution or species biodiversity within communities, landscapes and ecosystems, or genetic systems associated with spatial scales such as species populations. The discipline of phytogeography provides one such example, combining geography and botany to investigate the spatial distributions of plant species and their communities. Alexander von Humboldt and Aimé Jacques Bonpland's "Essai sur la géographie des plantes" [2], published in 1807, was a groundbreaking work in this respect. In it, Humboldt and Bonpland considered the variation of ecological gradients and put forth a theory of the geographical repartition of species, visualizing these concepts in portraiture. Works such as the "Tableau physique des Andes et pays voisins" contributed significantly to the formation of modern biogeography [2].

Biodiversity measures describe species and trait richness and evenness on different scales. Alpha biodiversity (α -diversity), or species diversity in habitats at a local scale, is influenced both by the number of types of habitat and ecological processes [3]. 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 [4–6]. These plant communities also tend to be less vulnerable to pathogens, the negative effects of high intensity wildfire and wind related disturbances [7]. It has been shown that the rates of such processes are affected by the spatial configuration of the environment. This study develops a hypothesis that α -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 [8].

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 [8], and also within the vertical and horizontal structural variation in forest landscape. The asymmetric competition for light, for instance, is noted in a forest structure meta-study as a key negative influence on overall productivity [9].

A forest's structure is composed of 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 [10]. On a global scale, forest canopies support approximately 40% of extant species, 10% of which are predicted to be canopy specialists [11]. 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 [11].

Landscape scale meta-studies have observed that stand structure is a biodiversity indicator. Gao et al. [12] noted 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 [12].

Chronosequential field studies of forest succession related α -diversity near the study area 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 [13]. In contrast to comparatively uniform mechanical stand replacement disturbances, Donato et al. [14] note that in forest ecosystems characterized by variable natural disturbances such as windfall and mixed-severity fire regimes, native biota possess functional traits that provide alternate successional rates and pathways that provide resilience to recurrent fire.

Due to its three-dimensional distribution of leaves, branches and stems, as well as its physical relationship to incoming solar radiation, canopy structure is a source of habitat niche partitioning for plant species [10]. Canopy structure can be considered in terms of both vertical components, such as strata, and horizontal ones, such as age related or other natural disturbance patches and fragments. Stands undisturbed by landscape-level events such as a 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 [15]. Properties of this structure can be described by metrics such as stand height, density, distribution and volume—each serve as a proxy for the more complex and unmeasurable distribution of forest canopy structure itself [16]. Derived point-cloud metrics measuring these variables with a high degree of accuracy can be developed with data sensed with laser arrays and subsequently “gridded” to raster formats for spatial modeling in a geographic information system.

The purpose of this study is to evaluate spatially explicit area solar radiation and structural variables in the forest canopy that influence its distribution in order to develop a diversity prediction model using the most significant canopy metrics in the available

data. Evaluating the statistically significant clustering of the predicted data is a secondary objective that assists in the observation of biodiversity hotspots.

Remotely sensed and field data acquisition from randomly selected biodiversity plots located in a natural, temperate, montane conifer forest in Northern California were used to develop a multivariate regression model and clustering analysis of predicated points to visualize the spatial distribution of plant α -diversity. The results are believed to be significant and also adequate enough to initiate a more comprehensive study to aid forest management and the development of ecosystem services prescriptions with plant diversity components.

The study area landscape and floristic background are discussed first, followed by a detailed review of the technology and methods used to derive metrics for the study. Analysis of the summary data, regression model and spatially clustering are provided prior to concluding remarks.

2. Materials and Methods

2.1. Study Area

The study area is located in the center of the Klamath Ecoregion of Northwestern 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 study area contains a unique conifer diversity concentration within a region that represents global maxima, comprising 30 native and endemic species [17].

The analysis extent and the study area are two distinct geographic areas due to solar radiation modeling requirements. The analysis extent is bounded by the topographic horizon captured in the bare earth digital elevation model (DEM). It consists of an arid east-slope canyon at the base of Russian Peak in the Upper Sugar Creek Watershed.

The study area used for canopy analysis is an approximately 18.5-hectare area located at 507090 E, 4572336 N (UTM) Zone 10, (NAD 1983) as shown in Figure 1. The study area elevation was between approximately 1500 m and 1700 m above mean sea level, with the plot samples observed primarily near 1500 m in elevation. Loamy inceptisols are the dominant soil taxon in the study area. These soils are common on landscapes that are relatively active, such as mountain slopes, where erosional processes expose unweathered materials. All plots were sampled within a Klamath Mixed Conifer (KMC) plant community within an average canopy height of one standard deviation from the mean of the plant community boundary at a 30 m resolution. Plots were selected from within a uniform climate, elevation, soil type and vegetation profile. Surface water quantity varied considerably, even outside excluded riparian habitat, due to water point-source locations and ground surface structure variability associated with biomass accumulation.

2.2. Biogeography

The Klamath Ecoregion is characterized by “complex biogeographic patterns, high endemism, and unusual community assemblages” [17]. 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 [17]. Relic, endemic and regionally dominant species coexist in among abrupt alternes and a diverse series of xeric, hydric and mesic habitats defined by variability in elevation and microclimate [18]. The taxonomic biodiversity present in the study area itself is thought to be the unique result of its topographic-climate variability and biogeographic history [18].

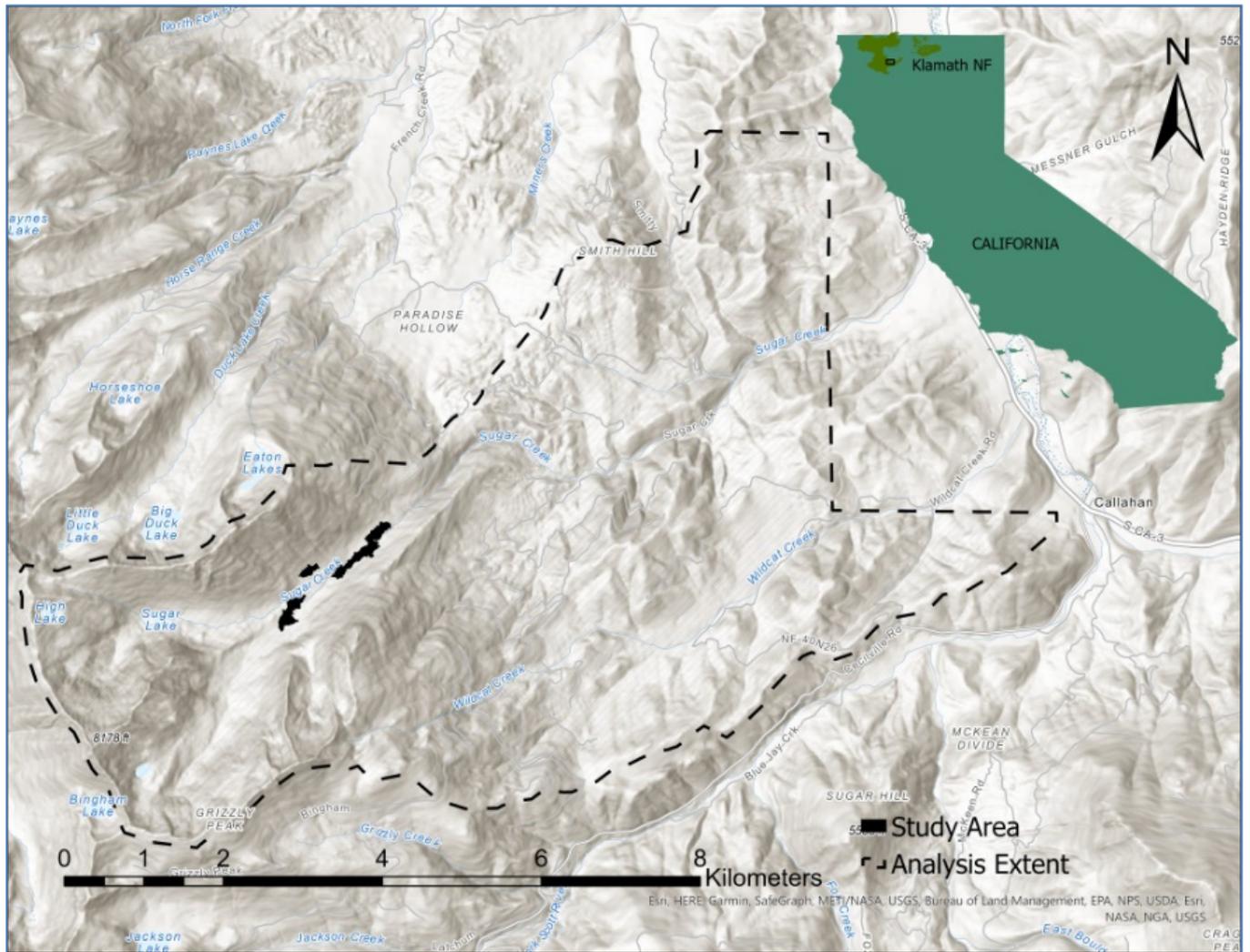


Figure 1. The study area (bounded by montane mixed-conifer forest) and the analysis extent, located in the Klamath National Forest in Northern California. (Source: Existing Vegetation-CALVEG 2007, QSI).

2.3. Forest Species

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 located on the middle Klamath River reach is one of the last remaining unfragmented lowland forests of old-growth Klamath Mixed Conifer [17]. The montane 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 [19]. It is dominated by evergreen conifers up to 60 m 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 [20]. Figure 2 shows xeric and mesic understory plant communities observed in the study area.

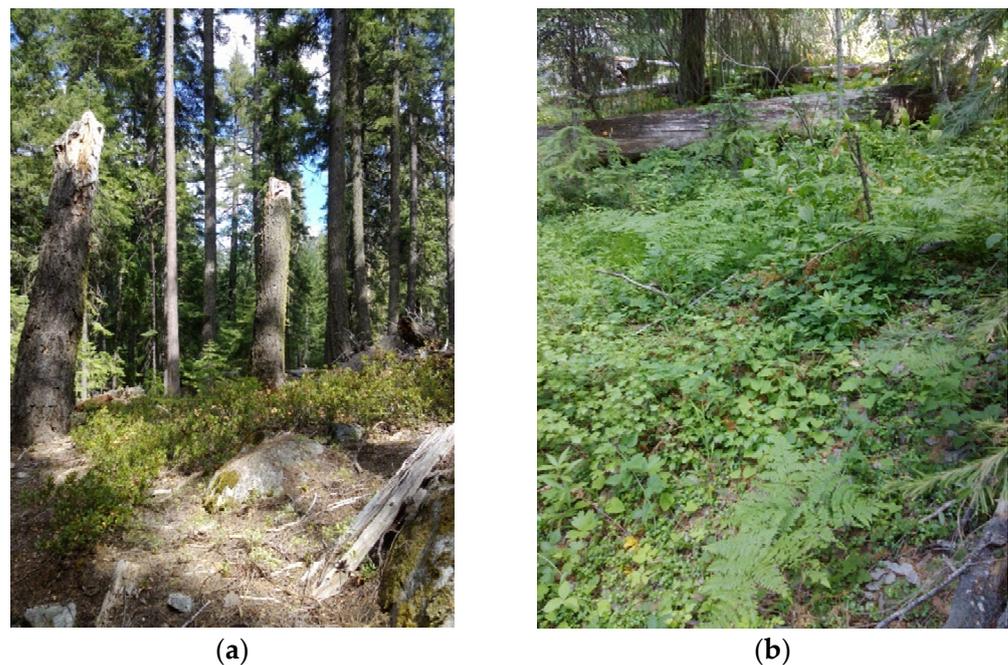


Figure 2. Montane Klamath Mixed Conifer stand (a) and the lower strata herbaceous community (b) in the study area.

2.4. Meteorology and Fire Regime

The region's weather patterns are a key source of canopy disturbance in its forest ecosystems. 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 [21]. The Klamath Ecoregion is noted to have a mixed-severity fire regime, with a fire return interval of approximately 15 years in lower montane conifer forests. Return interval describes a fire frequency at the scale of a stand or relatively small landscape area, while the fire rotation interval presents a more nuanced view of the local study area fire regime in that it describes the fire cycle over the larger scale landscape with variable spatio-temporal frequency and intensity. Notably, the study area interval has been influenced by fire suppression [22].

2.5. Area Solar Radiation, 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 [4]. In the context of the canopy surface radiation balance insolation is shortwave radiation influenced by atmospheric water vapor, aerosols and ozone, as shown in Table 1.

Table 1. Radiation balance based on average global incident solar radiation of 342 Wm^2 (Source: Schulz 2005).

Location	Incoming Solar Radiation (Absorption)	Outgoing Shortwave Radiation (Reflectance)
Space	100%	31%
Atmosphere	Absorbed by water vapor, aerosols and O_3 : 16% Absorbed by clouds: 4%	Back scattered: 6%
Surface	49%	Reflected by clouds: 16% 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 sur-

rounding terrain and the diffuse fraction of radiant flux [23,24]. Therefore, the spatial variability of annual photosynthetically active radiation (PAR) irradiation on the terrain surface is substantial in complex topography [23]. Seasonal variation due to atmospheric opaqueness also distributes radiation spatially due to topographic effects, with aspects receiving direct radiant fluxes contrasting with aspects that receive primarily diffuse solar radiation [25].

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 [26]. Study metrics are more likely associated with cover than closure per se, and additional closure ground measurements for each modeled point would be useful to understand attenuation. Sub-canopy radiant flux is consequential to habitat diversity and plant species persistence. Additionally, while the upper strata of the canopy determine the type of vegetation on the forest floor [4], understory microclimate is determined by the topographic variability of the ground surface and availability of water, production in the overstory canopy, and to the distribution of understory species and the maintenance of subsurface processes [27].

2.6. Data

2.6.1. Multispectral Imagery Data Specification

WorldView-2 is a high-resolution 8-band multispectral satellite operating at an altitude of 770 km. A corrected early season “leaf-on” image acquired in 2013 by this system 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) [28]. The imagery provides 1.85 m multispectral resolution (source scale), and a swath width is 16.4 km at nadir with a demonstrated geolocation accuracy of less than 3.5 m without ground control.

2.6.2. LiDAR Data Acquisition and Specification

The 2015 Sugar Creek LiDAR survey used the NAD83 (CORS96) datum (epoch 2002.00) and a Leica ALS50 system (Leica Geosystems, Heerbrugg, Switzerland, part of Hexagon group) for acquisition in late season “leaf-on” conditions. Actual first return average point density for the survey averaged 16.11 points/m² for the analysis extent while ground classified point density, used for accuracy assessments with ground control survey data, was 2.36 points/m².

Absolute accuracy was determined using Fundamental Vertical Accuracy (FVA) methods outlined in the FGDC National Standard for Spatial Data Accuracy. FVA compares known real-time kinematic (RTK) ground control data collected on open, bare earth surfaces with level slope (less than 20°) to the triangulated surface generated by the LiDAR points. Quantum Spatial acquired and performed post-acquisition processing to derive the LiDAR products used in the study. This included the high-resolution digital elevation model (DEM) and point-cloud dataset used to develop canopy metrics and surface models with the Boise Center Aerospace Laboratory (BCAL) library.

2.7. Methods

2.7.1. Data Processing Summary

Survey grade LiDAR products were used as inputs to georeference canopy and solar radiation models. The ERDAS Spatial Modeler (ERDAS, Atlanta, GA, USA, part of Hexagon Group) was used to derive leaf area index (LAI) data from the multispectral image, and Boise Center Aerospace Laboratory (BCAL) Vegetation and Intensity tools developed in ENVI (Harris Geospatial, Boulder, CO, USA) were used to derive canopy metrics from the LiDAR point-cloud data that were subsequently gridded to a 30 m resolution raster format from raster data scales less than 30 m.

The Klamath Mixed Conifer (KMC) forest-type was selected as the focal plant community to delineate the study area. KMC is a California Wildlife Habitat Relationship System

(CWHRS) attribute “crosswalked” in the Classification and Assessment with LANDSAT of Visible Ecological Groupings (CALVEG) data layer [29]. All LiDAR point-cloud and imagery data were resampled to 30 m resolution using bilinear or nearest neighbor methods for discrete and continuous data and then clipped to the KMC vegetation type within the analysis extent and elevation boundary in ArcGIS (ESRI, Redlands, CA, USA). This step established the study area, yielding 144 data points for each independent variable. The raster resolution and the corresponding centroids provided plot size and location coordinates for the field analysis.

The pre-processing of survey-grade data products required to analyze the study area and select plots for biodiversity sampling is shown in a diagram (Figure 3).

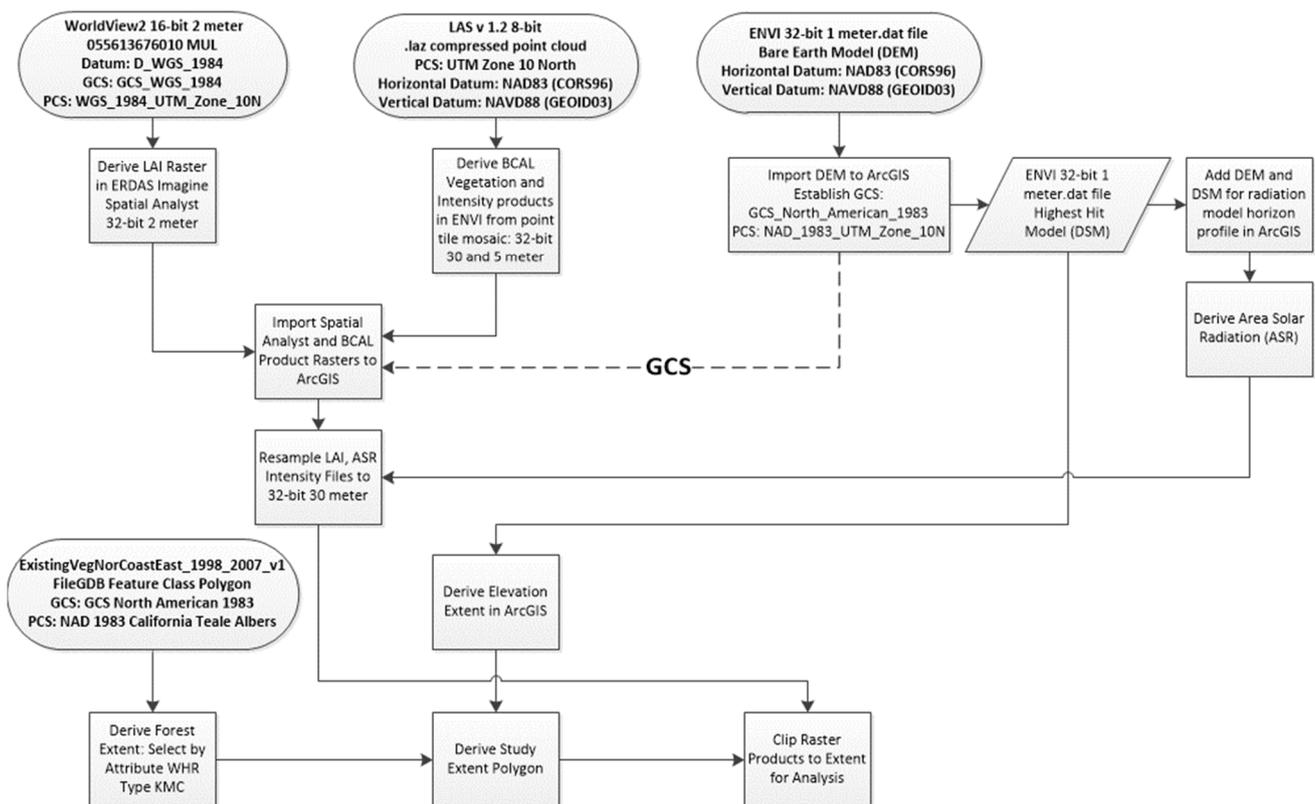


Figure 3. Data processing steps for the analysis extent and study area.

Each 30 m field plot references a pixel of gridded LiDAR data with the same resolution. All diversity values resulting from plot data are numerically ordinal and assumed to be spatially continuous. As ArcGIS transforms the geographic co-ordinate system (GCS) of layers to match the map’s reference GCS, all gridded raster and vector data used in the analysis were transformed to match the survey grade bare earth DEM provided in the post-acquisition LiDAR products. The WorldView-2 imagery was also transformed to these co-ordinate systems [30] using the ArcGIS transformation tools.

Forest dynamics (disturbance types and succession) are spatio-temporal phenomena that were not the direct focus of the study. There is a time-lag of five years between the LiDAR and multispectral data acquisition and the final field sampling of biodiversity plots. For this reason, sites evidencing plot-scale canopy disturbance were excluded from the study to exclude remotely sensed data that did not align with the observed vegetation height.

2.7.2. Canopy Structure and Area Solar Radiation 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 [31].

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

LAI is also a key variable for regional and global models of biosphere–atmosphere exchanges of energy, carbon dioxide, water vapor and other materials [32]. Canopy photosynthesis and its equivalent gross primary productivity (GPP) should theoretically reach a maximum as leaf area (*LAI*) increases to a value where spectral radiation in the 0.4 to 0.7 μm visible wavelength is totally intercepted [33].

It can be challenging to quantify *LAI* accurately due to significant spatio-temporal variability and measurement limitations inherent in current methodologies [26]. These methods utilize “direct” approaches employing field sampling or “indirect” approaches involving optical instruments combined with modeling [26]. 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*” [26].

LAI was derived in ERDAS Imagine Spatial Modeler by first deriving a Normalized Difference Vegetation Index (*NDVI*) from the multispectral image and then using a linear model to determine *LAI* [34]. *NDVI* is calculated from the visible red and near-infrared light (*NIR*) [35].

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2)$$

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* [35]. The biophysical interpretation of *NDVI* for Figure 4 is the fraction of absorbed photosynthetically active radiation (*fPAR*). Although derived *NDVI* values can range from -1.0 to $+1.0$, values less than zero do not have any biological meaning. The image data is classified for ease of interpreting *NDVI* and *LAI* values.

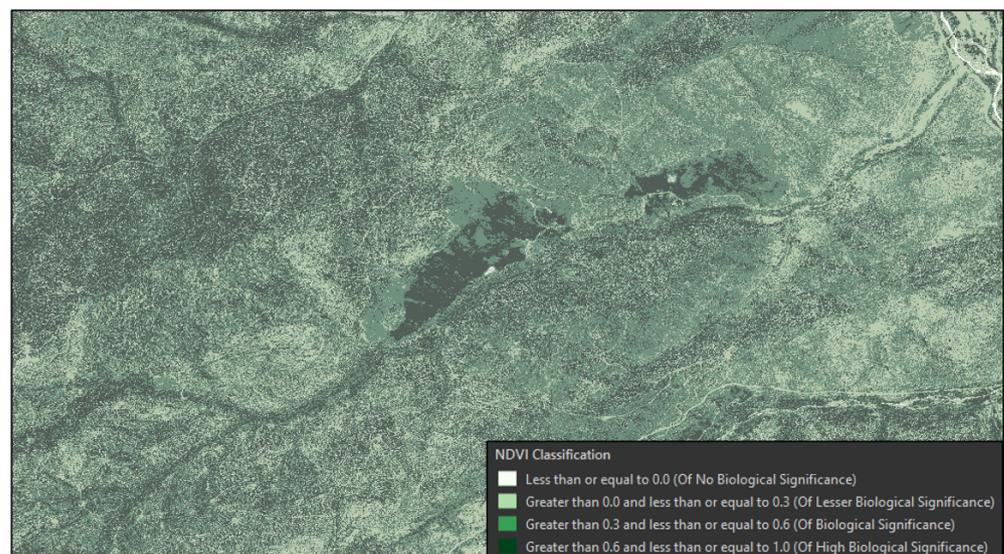


Figure 4. Classified *NDVI* map centered on the study area landscape in 2 m resolution (Source: DigitalGlobe).

The empirical conversion from the *NDVI* described by Wulder et al. [34] was then used to derive the canopy leaf area in the ERDAS Spatial Modeler tool (Figure 5) [24].

$$LAI = (17.35 * NDVI) - 9.01 \quad (3)$$

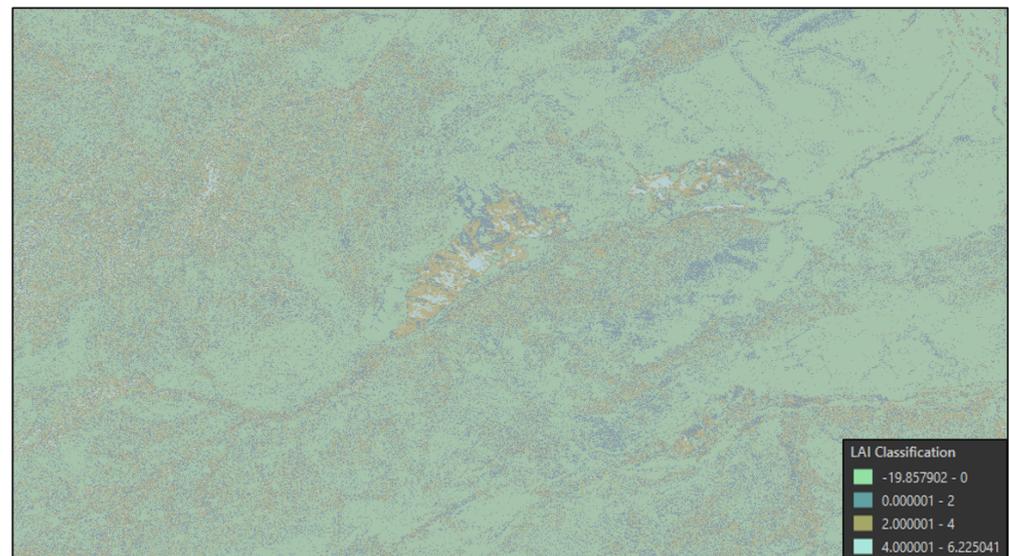


Figure 5. Classified optical *LAI* map centered on the study area landscape in 2 m resolution (Source: DigitalGlobe).

Photosynthetically available radiation (*PAR*) and its relationship to *LAI* is well established in the literature [26]. The theoretical basis for the absorption of light as it passes through a forest canopy is described by the Beer–Lambert law:

$$t(x) = \exp[-K(LAI(x))] \quad (4)$$

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 indicates that light intensity decreases exponentially as it passes through each canopy layer [4,33]. 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 [36].

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 [37]. A meta-study of canopy light extinction showed significant intra-annual negative correlations between extinction coefficient (*K*) and seasonal changes in *LAI* in natural ecosystems [38]. In another study, a *K* value of 0.48 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* [33], 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 [39]. The basis for calculating *fPAR* using an exponential function is based on Beer's law:

$$fPAR = P_{\infty}[1 - \exp(-LAI)] \quad (5)$$

where P_{∞} 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 [24]. This assumption was made in the *LAI* calculation for this study as well. Both parameters are used for evaluating surface photosynthesis, evapotranspiration and primary productivity for the analysis of ecosystem functions such as terrestrial energy, carbon, water cycle processes and the biogeochemistry of vegetation [39].

In addition to *LAI* derived from satellite imagery, insolation and canopy metrics were derived from LiDAR surface and point-cloud data to measure localized strata variables effecting solar radiation in the understory.

1. Canopy Height (CHM) in meters was determined from the LiDAR digital surface model (DSM) and the first returns data at 1 m resolution. It was utilized 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 m resolution matched to the gridded canopy metrics raster [40–42].
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 m 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 [43,44]. 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 WH/m^2 . Irradiance is the instantaneous density of solar radiation on a unit area expressed in W/m^2 . It comprises a global radiation value ($Global_{tot}$), or the sum of direct (Dir_{tot}) and diffuse (Dif_{tot}) radiation of all sun map and sky map sectors as shown in the solar model equation [43,44]:

$$ASR = Global_{tot} = Dir_{tot} + Dif_{tot} \quad (6)$$

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

$$Dir_{\theta,\alpha} = S_{Const} \times \beta^{m(\theta)} \times SunDur_{\theta,\alpha} \times SunGap_{\theta,\alpha} \times \cos(AngIn_{\theta,\alpha}) \quad (7)$$

where:

- S_{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 $1367 W/m^2$. This is consistent with the World Radiation Center (WRC) solar constant;
- β 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;

- $SunDur_{\theta,\alpha}$ 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;
- $SunGap_{\theta,\alpha}$ is the gap fraction for the sun map sector;
- $AngIn_{\theta,\alpha}$ the angle of incidence between the centroid of the sky sector and the axis normal to the surface [43,44].

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:

$$Dif_{\theta,\alpha} = R_{glb} \times P_{dif} \times Dur \times SkyGap_{\theta,\alpha} \times Weight_{\theta,\alpha} \times \cos(AngIn_{\theta,\alpha}) \quad (8)$$

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 [43,44].
3. Intensity of return (IR) is a measure of 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 m pixel resolution before being processed using the BCAL vegetation intensity tools and output to a 30 m resolution for the vegetation excluding bare earth data [40–42]. 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 [45].
 4. Total Vegetation Density is a derived percent ratio of vegetation to ground returns per m² or:

$$TVD = \frac{[n_i | nonground]}{N} 100 \quad (9)$$

where n_i is the number of vegetation returns in the i th layer [41,42].

5. Canopy Relief Ratio is a derived mean height less the minimum height divided by the maximum height less the minimum height per m² or:

$$CRR = \frac{\mu(\text{height}) - \min(\text{height})}{\max(\text{height}) - \min(\text{height})} \quad (10)$$

This ratio represents 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 [41,42,46,47]

6. Texture of Heights (TH) is the variance of height of points per pixel equal to the SD of height $>$ ground threshold and height $<$ crown threshold per m² or:

$$TH = \sigma (n_i | > \text{height}(0) \text{ and } \leq \text{height}(1) |) \quad (11)$$

where n_i is the number of returns in the i th layer [41,42].

7. Foliage Height Diversity is a derived Shannon diversity index statistic calculated as percentage cover at different heights per m² or:

$$FHD = - \sum p_i \ln p_i \quad (12)$$

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) [41,42].

2.7.3. Measures of Biodiversity

Biodiversity is the quantitative measure of the variety of distinct species in an area and their relative abundance, a biological phenomenon described with consideration of both spatial and temporal scales [4]. Alpha diversity (α -diversity) was developed by R. H. Whittaker during his study of plant communities in the Klamath region to describe habitat (local) level diversity characteristics. It was introduced with the complementary statistics, β -diversity and γ -diversity. Beta diversity is dimensionless, as a comparative statistic, i.e., the ratio between γ (regional) and α (local) diversities [48]. α -diversity and γ -diversity are limited to discrete units of space—roughly communities and ecosystems, respectively [4], but the ratio describes the number of unique habitats utilized by species in a region [49].

Whittaker postulated that total species diversity in a landscape, described by γ -diversity, is determined by the mean species diversity at a community scale (α -diversity) and also by the differentiation among those communities (β -diversity) [48,50]. Whittaker's subsequent usage of α -diversity implies the application of the statistic across multiple sites in a landscape, strongly influencing its primary use at the assemblage and community scale [48,50].

Insofar as the describing an observation of diversity present in a discrete unit of space, three broad types of biodiversity measures in the literature indicate an evolution in the understanding of biodiversity and related ecological processes. These are taxa based (taxonomic), trait based (functional, phenotypical expression of traits) and phylogenetic diversity, the latter defined as the minimum total length of phylogenetic branches required to span a given set of taxa on the phylogenetic tree [51].

We used taxonomic diversity indexes for relative ease of calculation and the extensive theoretical background in the literature, but the others are mentioned in the discussion as they represent improvements on the usefulness of a spatial model. That said, a fundamental ecological axiom is that habitats differ in the richness and abundance of species associated with them. Therefore, some issues arise in the most basic taxonomic diversity calculations which must be addressed. These are (1) that the total number of species in a sample varies both within habitats and between different habitats and that unless adjustments are made it is difficult to compare taxonomic α -diversity values, and (2) species of lesser abundance should not carry as much weight as those with higher abundance as their functional role is less [49].

Point samples were acquired within a uniform mixed-conifer habitat type, and Menhinick's Index was used to provide an adjustment to species richness data. It measures point diversity as the ratio of number of species (s) and the square root of the total number of individuals (N). It is notated:

$$MDI = \frac{s}{\sqrt{N}} \quad (13)$$

where:

- s is the number of different species in your sample;
- 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. This is notated:

$$SDI = 1 - \sum_{i=1}^s p_i^2 \quad (14)$$

where p_i is the proportion of individuals found in species i .

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

2.7.4. Multiple Regression Modeling and Cluster Analysis Methods

Multiple linear regression was used to model the selected predictor variables and a biodiversity response variable [52]. This is notated:

$$Y_i = b_0 + b_1x_{1,i} + b_2x_{2,i} + \dots + b_px_{p,i} + \varepsilon_i \quad (15)$$

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);
- $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$ represents the random error.

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 producing biodiversity prediction maps, 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 the predicted dataset. 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 dataset, 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:

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

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;
- S_0 is the aggregate of spatial weights $\sum_{i=1}^n \sum_{j=1}^n w_{i,j}$.

The Getis-Ord General G is denoted:

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

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;
- $\forall_j \neq i$ Indicates that features i and j cannot be the same feature.

2.7.5. Sample Plot Selection

Six random plots were sampled without adjusting for variation in slope, aspect or relative soil hydrology of the site (xeric, mesic, or hydric). Each sample plot provided 900 square meters of taxonomic field data with a total survey area of 5400 m². The survey was conducted by quantifying species richness and abundance within the plot perimeter. Absolute quantitative numbers were not the focus; species richness was estimated in areas of significant vegetation abundance. Surveyed taxa included visible vascular plants, bryophytes and canopy lichens. Overall accuracy of species identification was high, with few visible species remaining unidentified or uncounted. The data are representative of the forest vegetation in the study area, but likely understate actual richness and abundance.

The plot map consisted of a web feature service published to a hosted cloud server, downloaded and synchronized to a field data collection tablet paired with a map grade GNSS receiver. The mobile field data collection device utilized a satellite based real-time kinematic (RTK) network for plot position corrections (or SBAS if the RTK network was unavailable). This method achieved an estimated horizontal plot accuracy of approximately 2 m with respect to the LiDAR survey data used to produce the web-gis map service.

2.7.6. Biodiversity Response Variables

Richness (number of specific taxa), abundance (quantity of each specific taxa) and two additional measures of biodiversity, *MDI* and *SDI*, were derived for the model as shown in Table 2. Figure 6 shows the distribution of the data.

Table 2. Derived plot survey data for richness (R), abundance (A), the *MDI* and *SDI*.

Plot	Richness (R)	Abundance (A)	<i>MDI</i>	<i>SDI</i>
1	14	361	0.737	0.829
2	15	222	1.007	0.629
3	18	655	0.703	0.866
4	18	601	0.734	0.723
5	11	171	0.841	0.852
6	22	662	0.855	0.771

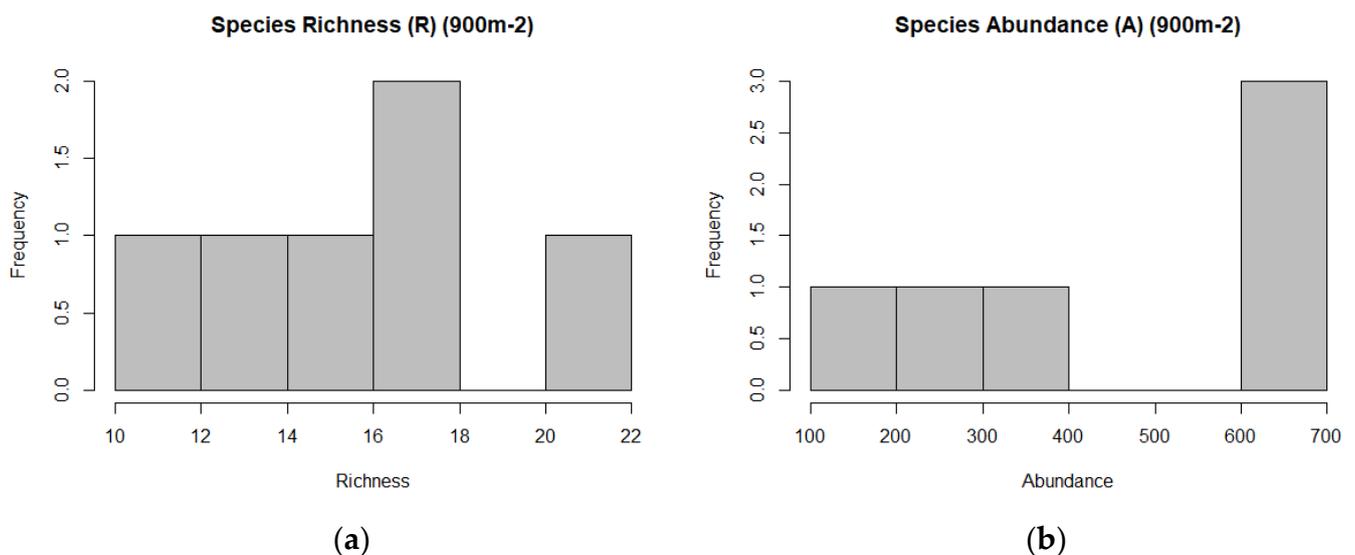


Figure 6. Cont.

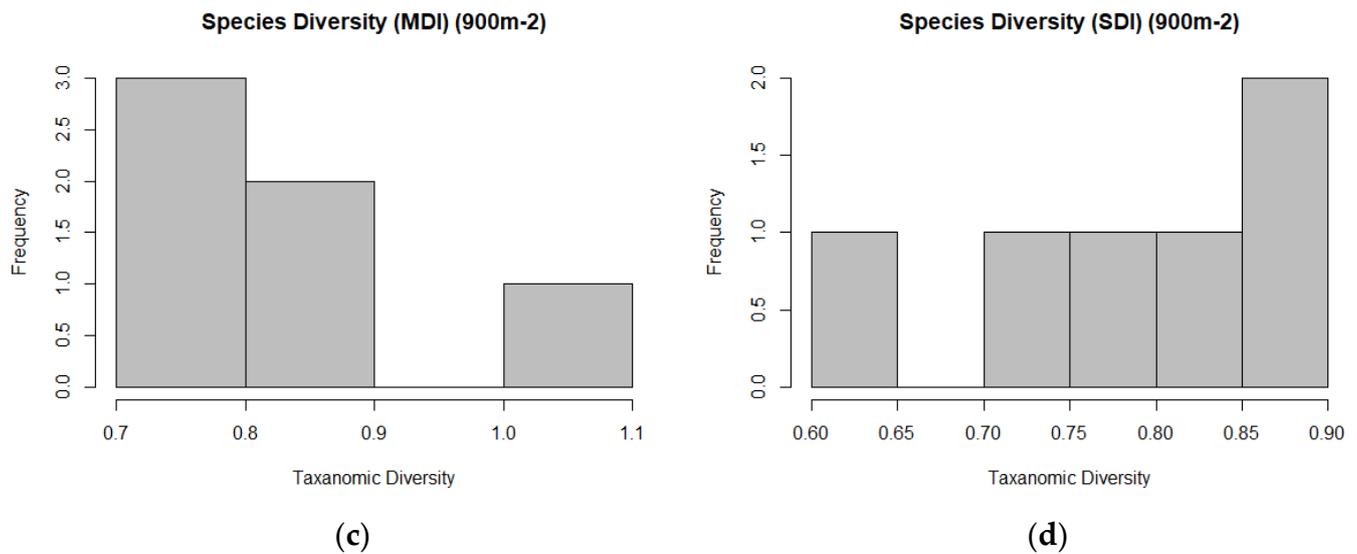


Figure 6. Distribution of plot values for species richness (a), abundance (b), the MDI (c) and SDI (d).

3. Results

Study Area Predictor Variable Distributions

Area solar radiation derived from the DSM and the point-cloud and imagery metrics gridded to raster yielded 144 data points for analysis as predictors. The histograms in Figure 7 depict distributions of the predictor variable data points used in the regression equation.

A collinearity analysis of selected predictor variables is shown in Figure 8. It displays coefficient values in the upper half of the matrix and non-parametric trend lines (loess smoothed fits) and correlation ellipses below the diagonal. Modeled irradiation (*ASR*) within the study area dataset exhibited a correlation with canopy height. A significant positive correlation coefficient (r) of 0.61 provided a credibility test of the irradiation data since canopy height should reasonably have some association to WH/m^2 due to the direct relationship between photosynthetic processes to the quantity of radiant energy in one area of space relative to another area of space.

Effective leaf area and vegetation density relationships were not significant. It is worth noting that *LAI* and *ASR* were negatively correlated, and although *TVD* and *ASR* exhibited minimal correlation, *LAI* and *TVD* interestingly did were negatively correlated as well. The acquisition date for the multi-spectral data was early in the growth season, and although the forest canopy is dominated by evergreen species, the full lower strata canopy was not at a seasonal peak. Much of the analysis extent exhibited significant canopy height and structural variation with dead biomass in all strata. 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.

Collinearity analysis was subsequently performed on all predictors and the four dependent variables as well. Abundance of individuals (*A*) was correlated with the texture of heights (*TH*, $r = 0.59$) and the intensity of the returns (*IR*, $r = 0.70$). Taxonomic diversity (*MDI*) was significantly correlated with leaf area (*LAI*, $r = 0.69$). Taxonomic diversity adjusted for evenness (*SDI*) was modestly correlated with canopy height (*CH*, $r = 0.42$) and the intensity of the returns (*IR*, $r = 0.24$).

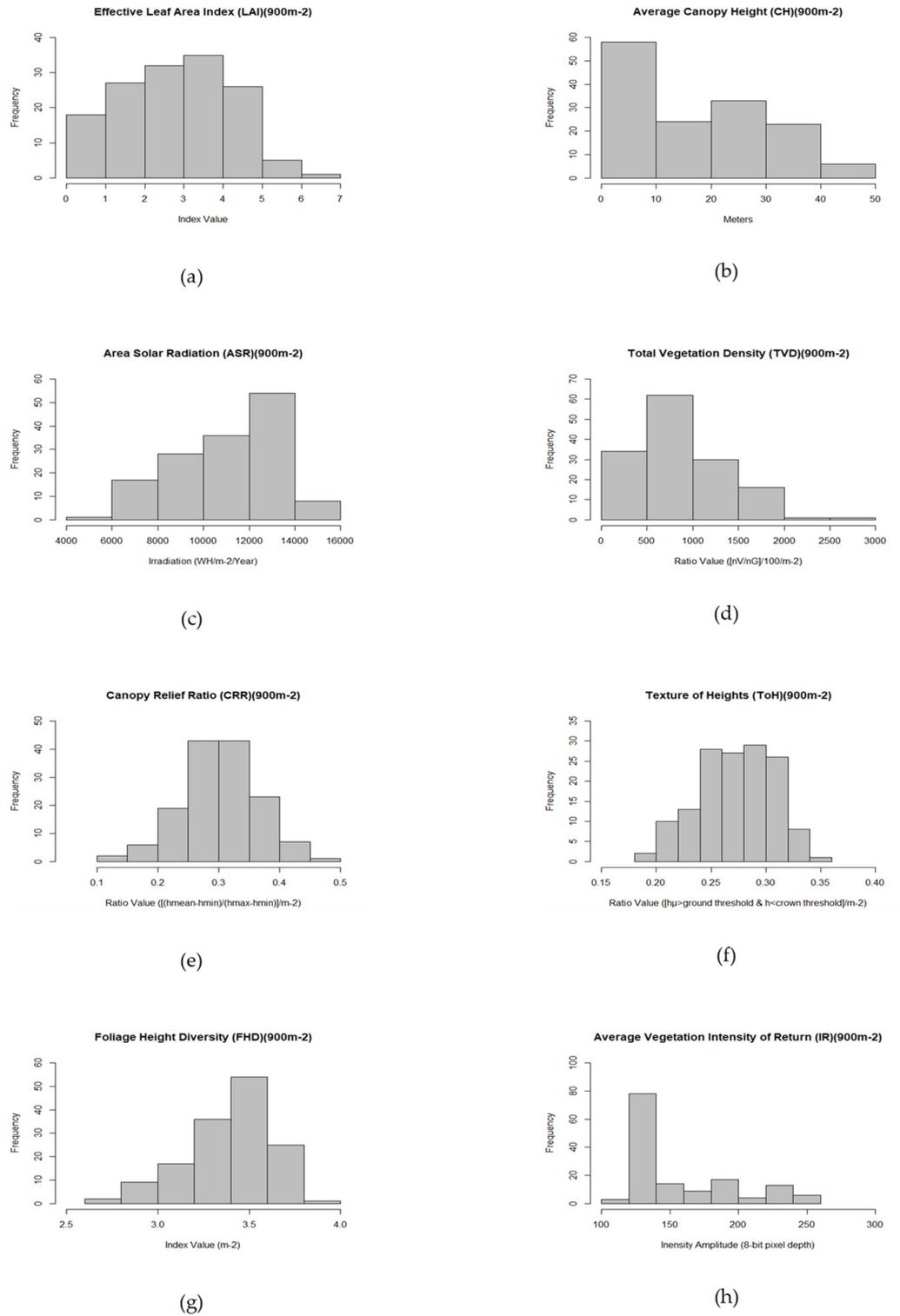


Figure 7. Study area predictor variable distributions for effective leaf area index (a), canopy height (b), area solar radiation (c), total vegetation density (d), canopy relief ratio (e), texture of heights (f), foliage height diversity (g) and average intensity of return (h).

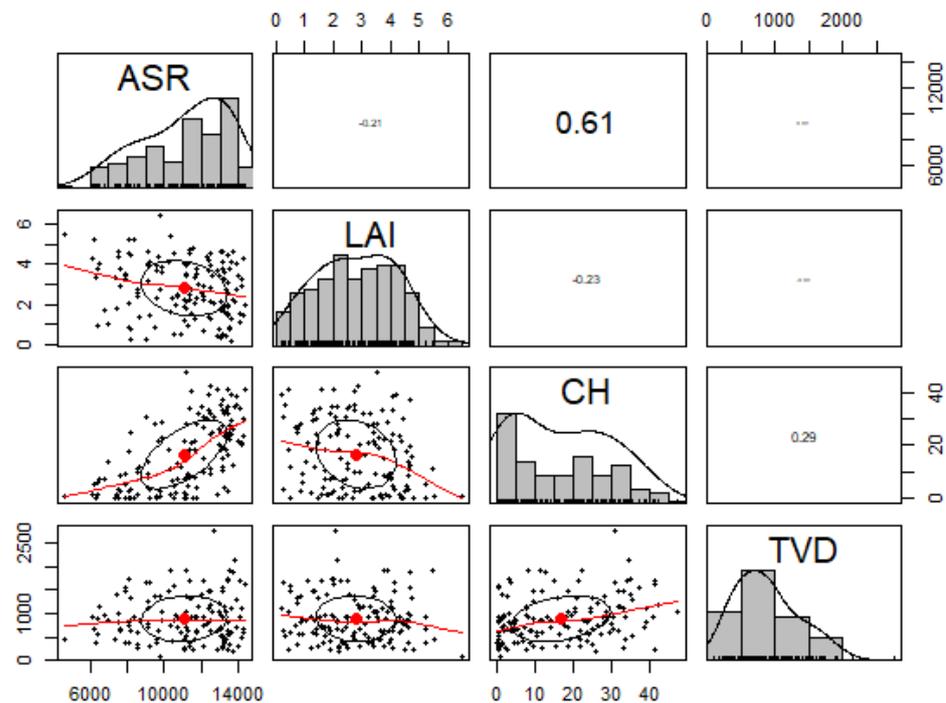


Figure 8. Density matrix of area solar radiation (ASR), leaf area index (LAI), canopy height (CH) and total vegetation density (TVD).

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:

$$\binom{p}{k} = \frac{p(p-1)}{k} \quad (18)$$

where:

- p is the specific predictor variable;
- 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.

Using R linear modeling functions, several variable subsets with significant r^2 values were evaluated to estimate the parameters of the linear model for all measures of diversity. A cross-validation analysis is shown in Figure 9 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 (0.9) is associated with two of the variables for richness (R). The results indicate a similar portion of the variance (0.91) is associated with two of the variables for abundance (A), but the variables differ. The MDI diversity cross-validation indicates that much of the variance (0.99) is associated with three of the variables, while the SDI diversity analysis indicates that much of the variance (0.99) is attributed to three of the predictors and a similar proportion (0.94) is associated with two of the variables.

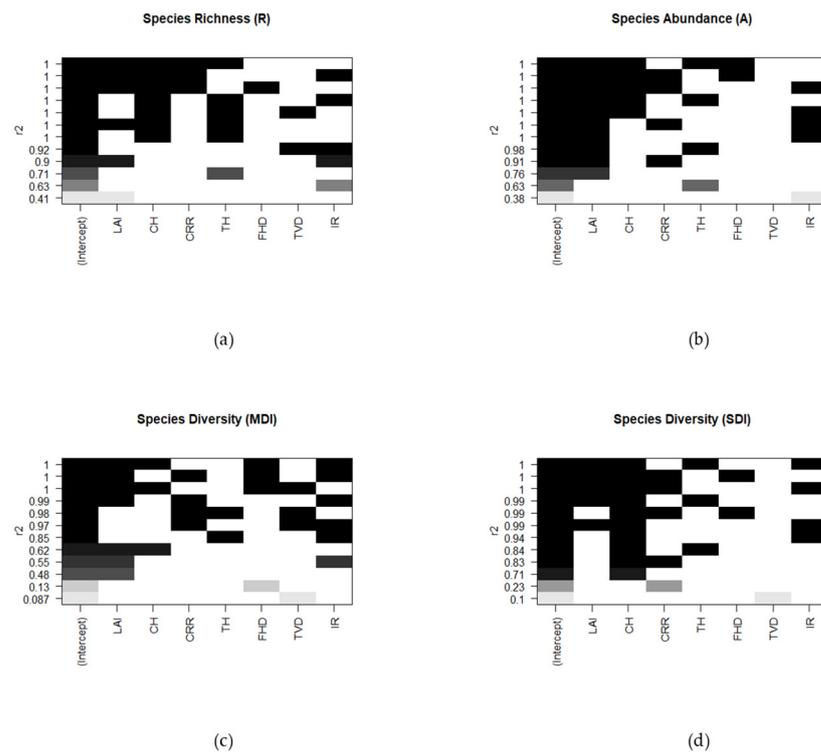


Figure 9. Cross-validation graphic displaying variance relationships for predictor variables as applied to richness (a), abundance (b), the MDI (c) and SDI (d).

Development of the prediction model focused the two diversity indexes, MDI and SDI. Table 3 compares the results of three SDI models (Lm0,1,2) to see if the available data would yield a good predictive model. The Lm0 F statistic and p -value do not indicate the model itself is significant, while Lm1 was created with two variables having the best p -values from the first model, and a third model (Lm2) contains an interaction term that increases the total variance associated with the predictors.

Table 3. Model Result Summary.

Model	LAI + CH + IR (Lm0)	LAI + CH (Lm1)	CH + IR + CH:IR (Lm2)	CRR + TH + TVD (Lm3)
Adjusted r^2	0.94	−0.3252	0.693	0.9687
F statistic	22.24	0.38	4.763	52.53
p -value	0.1544	0.709	0.1784	0.01874
Coefficient significance	CH/0.05	N/A	CH, IR, CH:IR/0.05	CRR, TVD/0.001; TH/0.01

Linear model 3 (Lm3) for MDI summary statistics present the best scenario for predictive modeling given the available data. The model p -value is below 0.05 and the adjusted r^2 value is greater than 0.96. In addition, each of the predictor variables strongly contributes to the significance of the model without adjustment.

The validity of linear regression modelling for prediction rests on four assumptions:

- A linear relationship between the dependent and predictor variables;
- The model errors are independent;
- The model errors are normally distributed;
- The model errors have a constant variance with respect to the predictor variables [53].

The plot in Figure 10a graphs the model errors (residuals) vs. the predicted values (fitted) and tests for non-linearity 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 [53]. The QQ plot (Figure 10b) displays a satisfactory positive diagonal trendline for the available data.

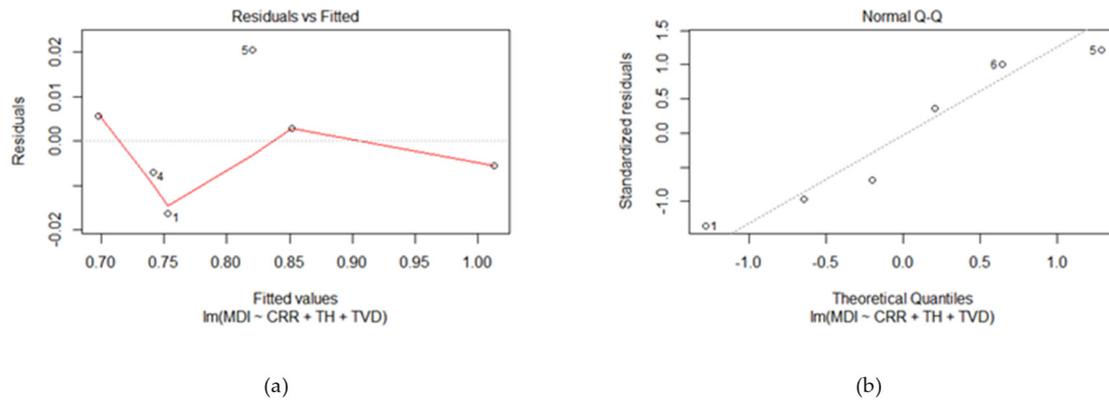


Figure 10. MDI linear model (Lm3) residuals vs. fitted values (a). Some row values are flagged for a discrepancy in the fitted vs. residual relationship (1,4–5). QQ plot of linear model (Lm3) errors (b). Variables 1, 5–6 are flagged.

Coefficients from the model were used to display the prediction spatially over the study area extent (Figure 11):

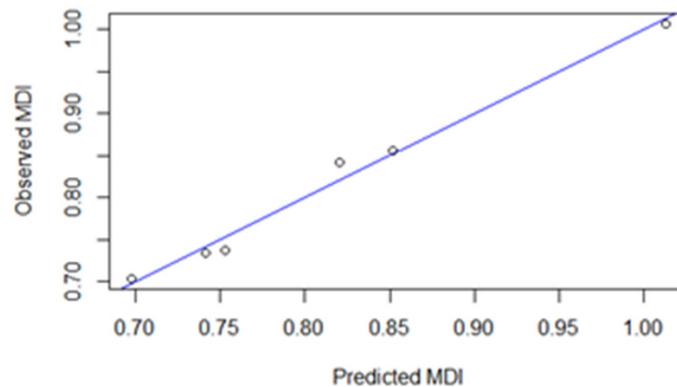


Figure 11. Graphic displaying predictive scatterplot for MDI (Lm3).

Multicollinearity, such as that exhibited between CRR and TVD ($r^2 = 0.80$), makes 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.

Since predicting beyond the ranges of the original data will result in model extrapolation and subsequent prediction errors, areas where the model extrapolated beyond the range of observation were “masked out”. Average CRR at a 30 m resolution ranged in value from 0.27 to 0.40, the average TH from 0.23 to 0.31 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.

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, these raster data are 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 [54] shown in Figure 12.

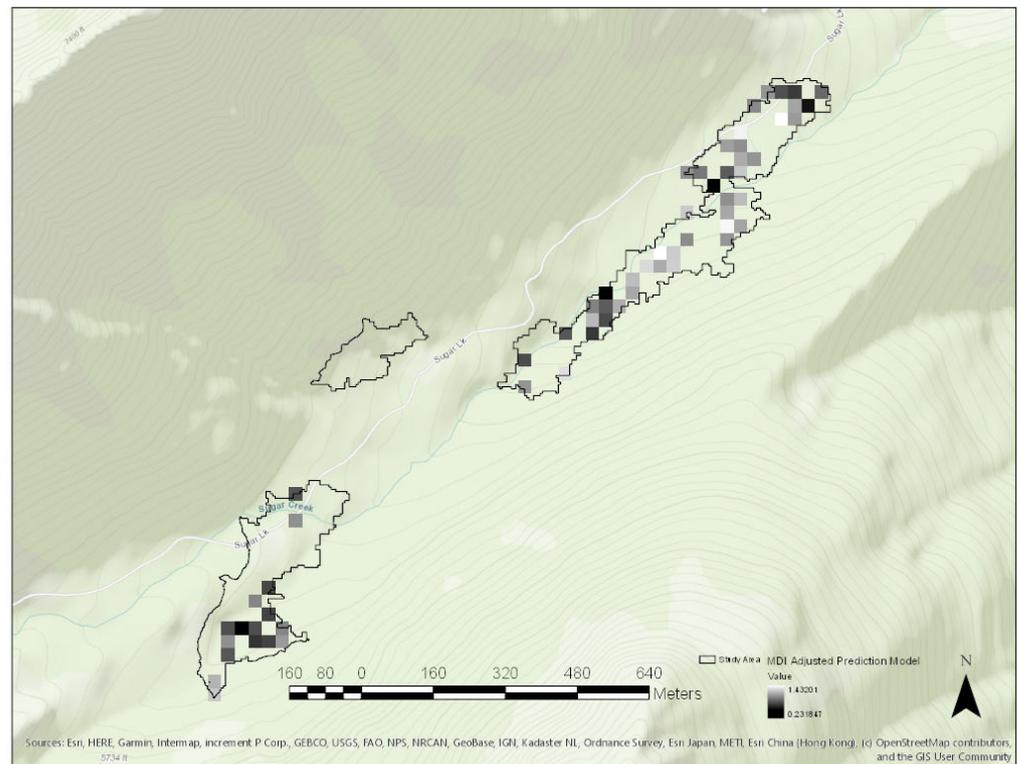


Figure 12. Map depicting spatial prediction raster of MDI in 30 m resolution, adjusted for predictor variable ranges.

Table 4 provides a comparative summary for the Moran's I and Getis-Ord General G values. A 100 m Euclidian distance parameter was estimated for spatial statistics, although an estimated value based on nearest neighbor distances is considered as a default practice when not specified.

Table 4. Global Moran's I and General G Results for Predicted MDI.

Global Moran's I		Getis-Ord General G	
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

The resulting statistics imply 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 affirms 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.

The Getis-Ord G_i^* statistic is 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 hotspot analysis map from the predicted raster values (Figure 13). Hotspots derived from this statistic are both positive and negative, like the General G, and both types have significance for biodiversity prediction, based on the study parameters.

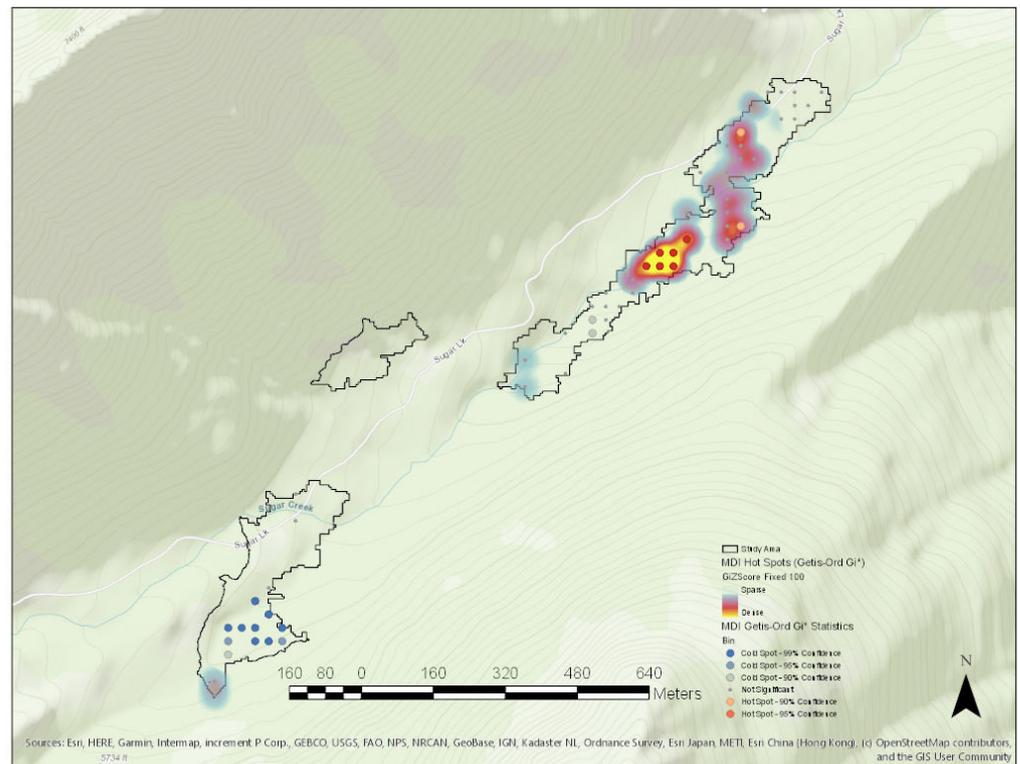


Figure 13. Map depicting Getis-Ord G_i^* clustering of the predicted MDI raster centroid values.

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

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

where:

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

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

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (19)$$

The statistic in this case used a fixed distance band of 100 m 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 specified distance for the “distance band” and neighbors within that distance are weighted equally. The hotspot map data frame contained 61 points each with 2–13 neighboring data points depending on its spatial position. These hotspots varied from 90–99% based on the concentration of nearest neighbors within 100 m of each analyzed data point.

4. Conclusions

Area solar radiation (*ASR*) and canopy height (*CH*) exhibited a significant positive correlation, and canopy height also exhibited positive correlation to total vegetation density (*TVD*). Statistically significant relationships were evident between the canopy structure and biodiversity data as well. While the available data did not indicate a relationship of significant magnitude between modeled radiant flux and observed biodiversity, there was a significant relationship between *CH* and diversity, depending on how diversity was measured.

The *MDI* linear model (*Lm3*) comprising the canopy relief ratio (*CRR*), the texture of heights (*TH*) and total vegetation density (*TVD*) was the most significant of the four models analyzed, given the available data. This set of predictors consists of metrics that would influence the variability of irradiance and the extinction of *PAR* in the lower strata. Unlike canopy height (*CH*), there was no significant positive correlation between *ASR* and vegetation density (*TVD*). The fact that *TVD* would have some relationship to *MDI* stands to reason, as does the relationship canopy variability represented by *CRR* and *TH*.

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 determined to be useful for predictive spatial modeling by providing over 60 data points after adjusting the prediction model for the range of canopy metrics surveyed.

Spatial statistics at the global level indicated the predicted range of *MDI* values with respect to proximity was not random overall, and the local hotspot map produced from this predicted data indicated several areas of both high and low biodiversity concentrations based on the predicted values. Within the constraints of the available data, this was an outcome validated by a field survey of the study area.

5. Discussion

The forest stands in the study area are relatively undisturbed by human activity except for recent historical fire suppression. It is important to remember that natural succession and regeneration in managed stands can differ considerably from these conditions. Modern even-age and commercial species focused silvicultural methods have often effectively promoted biodiversity decline. This has subsequently generated interest in managing for increased structural complexity to enhance stand productivity by promoting complementary resource utilization through spatial, temporal and physiological differentiation [4–6]. Structural variation in the canopy creates habitat differentiation by allowing for variation in sub-canopy insolation and subsequently the diversity and richness of species. The model results indicated that the study method may yield useful biodiversity prediction modeling protocols, especially with additional point samples and greater focus on horizontal and vertical vegetation density metrics with respect to insolation attenuation.

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 the abundance of vascular plants. Additional fieldwork could be considered in order to develop point sample variation and determine a more precise relationship of structural metrics that can be associated with measurement of insolation extinction (*K*). This would lead to a more coherent understanding of canopy insolation partitioning of the radiant flux and its impact on both diversity and productivity.

Forest productivity research indicates that both functional and phylogenetic diversity significantly influence biomass productivity on small scales, while taxonomic diversity evidenced “only indirect effects” at that scale [54]. Many researchers focused on the topic believe that the classification of species into functional types rather than higher taxonomic identity may improve the understanding of processes at the ecosystem scale, including vegetation responses and effects on climate, atmospheric chemistry, land use and

disturbances [55]. Although there appears to be no consensus on a uniform definition for functional diversity in the literature, there are contrasts in areas of its measurement: those that use trait values directly and those that use distance-based and dendrogram-based constructs [56].

It has also been proposed 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 [56]. Low functional richness, for instance, indicates that some alpha niches (i.e., resources) potentially available to the community remain unused, reducing productivity [57]. The number of traits included in future 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 [58].

Given the many leaf traits available to research that impact productivity, such as specific leaf area (SLA) in <SEE ARTICLE FORMAT>, these traits represent a practical area of focus for future application to this model. Given that trait-based information is more descriptive of biological phenomena than taxonomic information, this could have a significant impact on the prediction and understanding of spatially continuous ecological phenomena.

Our regression model focuses on generality with precision to predict 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 [52]. Guisan and Zimmermann [52] 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 habitat preferences [52].

Borrowing from the field of spatial econometrics, it may also be useful to consider spatial weights to adjust for proximity to abiotic temperature and precipitation gradients, or structurally related surface water and soil moisture in future iterations of the linear model. Whereas canopy structure provides niche partitioning associated with richness, additional abiotic factors likely influence abundance of plant species. Overall, biodiversity in the study landscape appears to be a clustered biotic phenomenon resulting from variation in the range of sub-surface to canopy forest structure, and consideration of limiting abiotic factors that impact plant species abundance would improve the model’s ability to predict and more accurate outcomes verifiable by field observations.

Author Contributions: Conceptualization, methodology, visualization, writing—original draft preparation, P.C.D.; writing—review and editing, L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Levin, S. The problem of pattern and scale in ecology. *Ecology* **1992**, *73*, 1943–1967. [CrossRef]
2. von Humboldt, A.; Bonpland, A. *Essai sur la Géographie des Plantes*; Levrault, Schoelle et Cie: Paris, France, 1807.
3. Dufour, A.; Gadallah, F.; Wagner, H.; Guisan, A.; Buttler, A. Plant species richness and environmental heterogeneity in a mountain landscape: Effects of variability and spatial configuration. *Ecography* **2006**, *29*, 573–584. [CrossRef]
4. Schulze, E.-D. *Plant Ecology*; Springer: Berlin/Heidelberg, Germany, 2005.
5. Zhang, Y.; Chen, H.; Reich, P. Forest productivity increases with evenness, species richness and trait variation: A global meta-analysis. *J. Ecol.* **2012**, *100*, 742–749. [CrossRef]
6. Ishii, H.; Tanabe, S.; Hiura, T. Exploring the Relationships Among Canopy Structure, Stand Productivity and Biodiversity of Temperate Forest Ecosystems. *For. Sci.* **2004**, *50*, 342.
7. Schneider, F.; Morsdorf, F.; Schmid, B.; Petchey, O.; Hueni, A.; Schimel, D.; Schaepman, M. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nat. Commun.* **2017**, *8*, 1441. [CrossRef]
8. Hillenbrand, H. On the generality of the latitudinal diversity gradient. *Am. Nat.* **2004**, *163*, 192–211. [CrossRef]
9. Ali, A. Forest stand structure and functioning: Current knowledge and future challenges. *Ecol. Indic.* **2019**, *98*, 665–677. [CrossRef]
10. Pan, Y.; Birdsey, R.; Phillips, O.; Jackson, R. The Structure, Distribution, and Biomass of the World's Forests. *Annu. Rev. Ecol. Evol. Syst.* **2013**, *44*, 593. [CrossRef]
11. Ozanne, C.; Anhuf, D.; Boulter, S.; Keller, M.; Kitching, R.; Körner, C.; Melnzer, F.; Mitchell, A.; Nakashizuka, T.; Silva Dias, P.; et al. Biodiversity Meets the Atmosphere: A Global View of Forest Canopies. *Science* **2003**, *301*, 183–186. [CrossRef]
12. Gao, T.; Hedblom, M.; Emilsson, T.; Busse Nielsen, A. The role of forest stand structure as biodiversity indicator. *For. Ecol. Manag.* **2014**, *330*, 82–93. [CrossRef]
13. Schoonmaker, P.; McKee, A. Species Composition and Diversity During Secondary Succession of Coniferous Forests in the Western Cascade Mountains of Oregon. *For. Sci.* **1988**, *34*, 960–979.
14. Donato, D.; Fontaine, J.; Robinson, W.; Kauffman, J.; Law, B. Vegetation response to a short interval between high-severity wildfires in a mixed-evergreen forest. *J. Ecol.* **2009**, *97*, 142–154. [CrossRef]
15. Pan, Y.; Chen, J.M.; Birdsey, R.; McCullough, K.; He, L.; Deng, F. Age structure and disturbance legacy of North American forests. *Biogeosciences* **2011**, *8*, 715–732. [CrossRef]
16. Leiterer, R.; Furrer, R.; Schaepman, M.; Morsdorf, F. Forest canopy-structure characterization: A data-driven approach. *For. Ecol. Manag.* **2015**, *358*, 48–61. [CrossRef]
17. Klamath-Siskiyou. Available online: <http://www.worldwildlife.org/ecoregions/na0516> (accessed on 1 December 2016).
18. Keeler-Wolf, T. *Ecological Surveys of Forest Service Research Natural Areas in California*; Gen. Tech. Rep. PSW-125; Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture: Berkeley, CA, USA, 1990; 177p.
19. Küchler, A.W. Appendix: The map of the natural vegetation of California. In *Terrestrial Vegetation of California*; Barbour, M., Major, J., Eds.; John Wiley and Sons: New York, NY, USA, 1977; pp. 909–938.
20. Sawyer, J.; Thornburgh, D.; Griffin, J. Mixed evergreen forest. In *Terrestrial Vegetation of California*; Barbour, M., Major, J., Eds.; John Wiley and Sons: New York, NY, USA, 1977; pp. 359–382.
21. Skinner, C.; Taylor, A.; Agee, J. Klamath Mountains bioregion. In *Fire in California's Ecosystems*; Sugihara, N., van Wagtenonk, J., Fites-Kaufman, J., Shaffer, K., Thode, A., Eds.; University of California Press: Berkeley, CA, USA, 2006; pp. 170–194.
22. DeSiervo, M.; Jules, E.; Kauffmann, M.; Bost, D.; Butz, R. Revisiting John Sawyer and Dale Thornburgh's 1969 Vegetation Plots in the Russian Wilderness: A Legacy Continued. *Fremontia* **2016**, *44*, 20–25.
23. Olseth, J.; Skartveit, A. Spatial distribution of photosynthetically active radiation over complex topography. *Agric. For. Meteorol.* **1997**, *86*, 205–214. [CrossRef]
24. Oliphant, A.; Susan, C.; Grimmond, B.; Schmid, H.-P.; Wayson, C. 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 Sens. Environ.* **2006**, *103*, 324–337. [CrossRef]
25. Nunez, M. The calculation of solar and net radiation in mountainous terrain. *J. Biogeogr.* **1980**, *7*, 173–186. [CrossRef]
26. Smith, M.-L.; Anderson, J.; Fladeland, M. Forest Canopy Structural Properties. In *Field Measurements for Forest Carbon Monitoring*; Hoover, C.M., Ed.; Springer Science+Business Media B.V.: New York, NY, USA, 2008; pp. 179–196.
27. Geiger, R. *The Climate near the Ground*; Harvard University Press: Cambridge, MA, USA, 1965.
28. Hamilton, R.; Cushman, S.; McCallum, K.; McCusker, N.; Mellin, T.; Nigrelli, M.; Williamson, M. *Multiscale Landscape Pattern Monitoring Using Remote Sensing: The Four-Forest Restoration Initiative*; RSAC-10022-RPT1; U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center: Salt Lake City, UT, USA, 2013; 24p.
29. Pacific Southwest Region. *Existing Vegetation-CALVEG*; ESRI geodatabase; USDA-Forest Service, Pacific Southwest Region: McClellan, CA, USA, 2017.
30. DigitalGlobe. *WorldView2 Image 055613676010*; DigitalGlobe: Longmont, CO, USA, 2013.
31. Chen, J.; Black, T. Defining leaf area index for non-flat leaves. *Agric. For. Meteorol.* **1992**, *57*, 1–12. [CrossRef]
32. Asner, G.; Scurlock, J.; Hicke, J. Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Glob. Ecol. Biogeogr.* **2003**, *12*, 191–205. [CrossRef]
33. Perry, D.; Oren, R.; Hart, S. *Forest Ecosystems*; Johns Hopkins University Press: Baltimore, MD, USA, 2008.

34. Wulder, M.; LeDrew, E.; Franklin, S.; Lavigne, M. Aerial image texture information in the estimation of northern deciduous and mixed wood forest leaf area index (LAI). *Remote Sens. Environ.* **1998**, *64*, 64–76. [CrossRef]
35. Normalized Difference Vegetation Index (NDVI) Analysis for Forestry and Crop Management. Available online: http://simwright.com/downloads/SimWright_NDVI.pdfSlide2. (accessed on 1 May 2019).
36. Attenuation Coefficient. Available online: https://en.wikipedia.org/w/index.php?title=Attenuation_coefficient&oldid=925582822 (accessed on 1 December 2019).
37. Saitoh, T.; Shin, N.; Noda, H.; Muraoka, H.; Nasahara, K. Examination of the extinction coefficient in the Beer–Lambert law for an accurate estimation of the forest canopy leaf area index. *For. Sci. Technol.* **2012**, *8*, 67–76. [CrossRef]
38. Zhang, L.; Hu, Z.; Fan, J.; Zhou, D.; Tang, F. A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems. *Front. Earth Sci.* **2014**, *8*, 599–609. [CrossRef]
39. “Leaf Area Index—Fraction of Photosynthetically Active Radiation 8-Day L4 Global 1km.” MOD15A2 | LP DAAC NASA Land Data Products and Services. Available online: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod15a2 (accessed on 2 December 2017).
40. QSI Environmental. *Technical Data Report-USFS PSW Region 5 LiDAR*; USDA-Forest Service, Pacific Southwest Region: McClellan, CA, USA, 2015.
41. Boise Center Aerospace Laboratory Lidar Tools. Available online: <http://bcsl.boisestate.edu/tools/lidar> (accessed on 30 April 2017).
42. Fu, P.; Rich, P. *The Solar Analyst 1.0 Manual*; Helios Environmental Modeling Institute: Lawrence, KS, USA, 2000.
43. Fu, P.; Rich, P. A Geometric Solar Radiation Model with Applications in Agriculture and Forestry. *Comput. Electron. Agric.* **2002**, *37*, 25–35. [CrossRef]
44. Evans, J.; Hudak, A.; Faux, R.; Smith, A. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sens.* **2009**, *1*, 776–794. [CrossRef]
45. Kim, Y.; Yang, Z.; Cohen, W.; Pflugmacher, D.; Lauer, C.; Vankat, J. Distinguishing between Live and Dead Standing Tree Biomass on the North Rim of Grand Canyon National Park, USA Using Small-Footprint Lidar Data. *Remote Sens. Environ.* **2009**, *113*, 2499–2510. [CrossRef]
46. Pike, R.; Wilson, S. Elevation-Relief Ratio, Hypsometric Integral and Geomorphic Area-Altitude Analysis. *Geol. Soc. Am. Bull.* **1971**, *82*, 1079–1084. [CrossRef]
47. Parker, G.; Russ, M. The canopy surface and stand development: Assessing forest canopy structure and complexity with near-surface altimetry. *For. Ecol. Manag.* **2004**, *189*, 307–315. [CrossRef]
48. Whittaker, R. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* **1960**, *30*, 279–338. [CrossRef]
49. Ricklefs, R.; Miller, G. *Ecology*, 4th ed.; W.H. Freeman: New York, NY, USA, 2000.
50. Whittaker, R. Evolution and Measurement of Species Diversity. *Taxon* **1972**, *21*, 213–251. [CrossRef]
51. Faith, D. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **1992**, *61*, 1–10. [CrossRef]
52. Guisan, A.; Zimmermann, N. Predictive habitat distribution models in ecology. *Ecol. Model.* **2000**, *135*, 147–186. [CrossRef]
53. USDA. *Technical Report-USFS Remote Sensing Applications Center*; USDA-Forest Service-RSAC: Salt Lake City, UT, USA, 2016.
54. Hao, M.; Zhang, C.; Zhao, X.; von Gadow, K. Functional and phylogenetic diversity determine woody productivity in a temperate forest. *Ecol. Evol.* **2018**, *2018*, 1–12. [CrossRef] [PubMed]
55. Cornelissen, J.; Lavorel, S.; Garnier, E.; Diaz, S.; Buchmann, N.; Gurvich, D.; Reich, P.; ter Steege, H.; Morgan, H.; van der Heijden, M.; et al. A handbook of protocols for standardized and easy measurement of plant functional traits worldwide. *Aust. J. Bot.* **2003**, *51*, 335–380. [CrossRef]
56. Petchey, O.; O’Gorman, E.; Flynn, D. A functional guide to functional diversity measures. In *Biodiversity, Ecosystem Functioning and Human Wellbeing: An Ecological and Economic Perspective*; Oxford University Press: Oxford, UK, 2009.
57. Mason, N.; Mouillot, D.; William, L.; Wilson, J. Functional richness, functional evenness and functional divergence: The primary components of functional diversity. *Oikos* **2005**, *111*, 112–118. [CrossRef]
58. Laureto, L.; Cianciaruso, M.; Samia, D. Functional diversity: An overview of its history and applicability. *Nat. Conserv.* **2015**, *13*, 112–116. [CrossRef]