5-2022

Multi-trophic Biodiversity Increases with Increasing Structural Complexity of Forest Canopy

Ayanna St. Rose
University of Arkansas, Fayetteville

Multi-trophic Biodiversity Increases with Increasing Structural Complexity of Forest Canopy

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Science in Statistics and Analytics

by

Ayanna St Rose
University of Arkansas
Bachelor of Science in Biological Sciences, 2018

May 2022
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____
Kusum J. Naithani, PhD
Thesis Director

_____
Mark Arnold, PhD
Committee Member

_____
Jack Cothren, PhD
Committee Member

Abstract

Understanding the effects of forest canopy structural complexity on multi-trophic diversity is critical for conserving biodiversity and managing land sustainably. But multi-trophic diversity is often ignored when making decisions about land management due to lack of cost- and time-effective methods to evaluate it. Here, we explored a new method based on widely available remote sensing data to quantify canopy structural complexity and its relationships with multi-trophic biodiversity at landscape scale using 32 forested sites of the National Ecological Observatory Network. We investigated the influence of vertical and horizontal structural complexity of forest canopy on multi-trophic (primary producers, herbivores (beetles), omnivores (birds)) diversity in forested ecosystems. We used plant presence, beetle pitfall trap, and bird count data to calculate species richness and species diversity, and high density LiDAR data for calculating structural complexity metrics of forest canopy. Our results show that species richness and diversity across all trophic levels generally increase with increasing vertical and horizontal structural complexity with highest diversity at intermediate levels of structural complexity, but these relationships differ across different forest types (deciduous, mixed, and evergreen). Our results highlight the importance of maintaining structural complexity in forest canopies for conserving multi-trophic biodiversity.

Acknowledgements

The National Science Foundation award (EAGER: 2026815) to KN funded this study. All data used in this study are available at the NEON Data Portal <https://data.neonscience.org/home> and data processing code is provided in the appendix.

Table of Content

Introduction

About 1.78 trillion m^2 of forested land has been lost globally from 1990 to 2020 (FAO & UNEP 2020). This global decline of forested ecosystems has created an urgency to re-evaluate our approach to forest conservation as forests provide key ecosystem services including supplies for medicine, food, fodder, fuel and construction, water filtration, air purification, carbon sequestration, climate change mitigation, recreational and cultural uses, and habitat for the majority of terrestrial biodiversity (Houghton 2005; Millennium Ecosystem Assessment (Program) 2005; Neary *et al.* 2009; Angelsen *et al.* 2014). In recent years land managers are moving from single species conservation (Fleishman *et al.* 2000; Poiani *et al.* 2001; Suter *et al.* 2002) to multi-species conservation (Barrows *et al.* 2005; Critchlow *et al.* 2022), but the progress has been slow due to complex techniques and lack of cost- and time-effective methods to evaluate it (Suter *et al.* 2002). Prior work has highlighted the role of the structural complexity of forest canopy, diversity of the physical attributes and spatial distribution of canopy (hereafter, structural complexity), in estimating forest age, species richness, and primary productivity (Franklin 1981; Ishii *et al.* 2004; Hardiman *et al.* 2011). Climate and topography (e.g. precipitation and latitude) shape the structural complexity of forested systems (Ehbrecht *et al.* 2021), which generally increases with forest age (MacArthur & MacArthur 1961) and can be maintained through sustainable forest management practices (Molina *et al.* 2006). Thus, older forests tend to show greater structural diversity and support greater animal diversity (Franklin 1981; Ishii *et al.* 2004; Hardiman *et al.* 2011; Carrasco *et al.* 2019). In this study, we explore the relationship between structural complexity (both vertical and horizontal), derived from Light Detection and Ranging (LiDAR) data, and multi-trophic biodiversity, derived from field data, at landscape scale to identify patterns that may support forest management decisions to promote

multi-trophic biodiversity conservation. Forest management and conservation efforts are often conducted at the landscape scale, so this landscape study is crucial to identify important indicators of multi-trophic diversity for forest management and biodiversity conservation.

Here, we refer to structural complexity as the diversity of the physical attributes and spatial distribution of the forest canopy including both vertical and horizontal complexity (McElhinny *et al.* 2005; Atkins *et al.* 2018). Vertical complexity refers to the attributes related to the height of the canopy, such as mean maximum canopy height (MOCH), maximum canopy height (MCH), vertical complexity index (VCI), entropy, and top rugosity (Zimble *et al.* 2003; Zellweger *et al.* 2013). Horizontal complexity refers to the variation in the canopy surface density and layout, such as deep-gap fraction, cover fraction, vegetative area index (VAI), and rumple (Zimble *et al.* 2003; Zellweger *et al.* 2013). The role of vertical complexity in supporting species diversity has received greater attention in prior work (Wolf *et al.* 2012; Guo *et al.* 2017; Camargo *et al.* 2018; Müller *et al.* 2018; Carrasco *et al.* 2019) including the foundational work on niche partitioning (MacArthur & MacArthur 1961). In contrast, the role of horizontal complexity in supporting species diversity is poorly understood, possibly due to the difficulty in attaining these measurements from ground data that have been historically used in site specific studies of forest structural complexity. However; recent advancements in LiDAR measurements has opened new possibilities for exploring horizontal complexity with a less labor-intensive alternative which yields comparable results to ground data (Zimble et al. 2003; Stark et al. 2012; Zellweger et al. 2013; Guo et al. 2017; Hardiman et al. 2018; LaRue et al. 2020). One of the major drawbacks of using LiDAR to study structural complexity is greater variance in low density LiDAR data, possibly due to shadows created by the canopy, but this can be rectified with high density data (Zimble et al. 2003; Chow & Hodgson 2009).

Previous studies on the effect of structural complexity on species diversity were mainly site specific (e.g. study of plant diversity and structural complexity in the Great Smoky Mountains (Walter *et al.* 2021)) or focused on single trophic levels (i.e. avian diversity in (MacArthur & MacArthur 1961; Zellweger *et al.* 2013; Carrasco *et al.* 2019)). Here, we take a landscape scale approach to investigate the effects of vertical and horizontal structural complexity on the diversity of primary producers (plants) and consumers (beetles and birds) in forested ecosystems using high density airborne LiDAR data and field data (plant presence, beetle pitfall trap, and bird count) from the National Ecological Observatory Network (NEON). We asked: (1) Does vertical structural complexity affect multi-trophic diversity in forested ecosystems? We expect that species diversity will increase with increasing vertical complexity as reported in previous studies (Franklin 1981; Kern *et al.* 2014; Carrasco *et al.* 2019; Walter *et al.* 2021) across all trophic levels. (2) Does horizontal structural complexity affect multi-trophic diversity similarly to vertical complexity in forested ecosystems? We expect that multi-trophic diversity will increase with increasing horizontal complexity, similar to vertical complexity (Carrasco *et al.* 2019; Walter *et al.* 2021). And (3) Do these relationships differ across different forest types (deciduous, evergreen, and mixed)? We expect similar patterns of increasing multi-trophic diversity with increasing structural complexity across different forest types.

Methods

We used openly available data from the National Ecological Observatory Network (NEON), operated by Battelle and funded by the National Science Foundation. The NEON field site metadata table was downloaded for each data product and the field dominant "nlcd class" column was filtered using the keyword "forest." There were 54 field sites that had "forested" listed as part of the name of their nlcd class. We selected 32 forested sites (Fig. 1) with available

data needed to answer the main questions of this study. Please refer to Table 1 in appendix for site specific information on siteID, domain, latitude, longitude, mean annual temperature, mean annual precipitation, elevation, forest type and year that LiDAR data was obtained for each site.

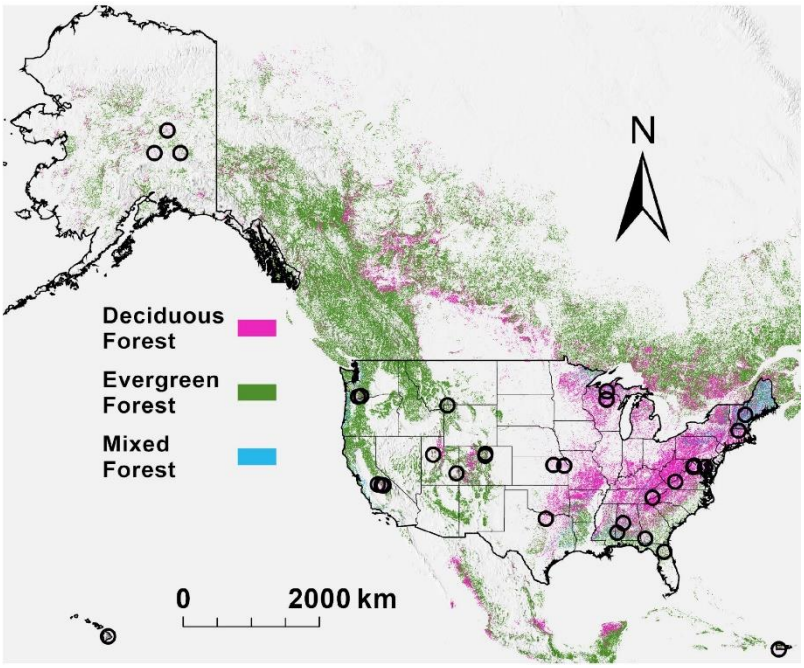

Figure 1: Map of the study area showing locations of 32 forested sites of the National Ecological Observatory Network. Darker circles indicate overlap of sites that appear closer due to scale of map. The World Forest (MDAUS BaseVue 2013) data in ArcGIS Pro is used for displaying different forest types (evergreen, deciduous, and mixed). Source: MDAUS, Airbus, USGS, NGA, NSA, CGIAR, NCEAS, NLS, OS, NMA, GSA, GSI, Geodatastylerelsen, and the GIS User Community, Esri, USGS, NEON.

NEON Data Products

We accessed all available data for Breeding landbird point counts (DP1.10003.001), ground beetles sampled from pitfall traps (DP1.10022.001), plant presence and percent cover (DP1.10058.001), and the most recent data for elevation - LiDAR (DP3.30024.001) for 32 sites across 17 ecoclimatic domains (NEON 2021a, 2021b, 2021c, 2021d).

Breeding Land Bird (DP1.10003.001)

This data product samples small birds that are only associated with terrestrial ecosystems during the first half of the breeding season. Observers used a point count method to collect data of the birds that were seen and heard within a 6 minute count (Thibault 2020). This presence data spanned from 2016 to 2020 for the 32 sites. The data was filtered by the siteID and taxonID column.

Ground Beetles Sampled From Pitfall Traps (DP1.10022.001)

This data product provides counts and taxa of beetles within each site. Traps are set up several meters from forest edges, roads, and buildings. Data were collected every two weeks in the growing season for every year (LeVan 2020). Each plot had about 30 traps set up outside of the biodiversity area which is a 20-m^2 area in the center of each plot. Available data for the 32 sites span from 2014 to 2019.

Plant Presence and Percent Cover (DP1.10058.001)

This data product contains % cover at a fine scale of 1-m^2 and plant presence in 400-m^2 plots. All unknown species were removed and then species richness and diversity metrics (Elmendorf 2021) were calculated. Available data for the 32 sites span from 2013 to 2020.

Elevation-LiDAR (DP3.30024.001)

LiDAR flights were conducted at peak greeness at each site. This data product contains processed Digital Terrain Model (DTM) and Digital Surface Models (DSM) of 1000-m^2 area of a single plot. There is a vertical resolution of 1.5 m and a spatial resolution of about 0.5 m (Krause and Goulden 2015). We subtracted the rasters of the DTM from DSM to create a Canopy Height Model (CHM) for each site that was further used for calculating the structural complexity metrics.

Statistical Analysis

We pulled data from the NEON data portal using the neonUtilities (v 2.1.3) package (Lunch et al. 2021), filtered all data products to remove unassigned taxa, calculated the shannon diversity index using the vegan (v 2.5-7) package (Oksanen et al. 2020), calculated forest structural complexity metrics using the lidR (v. 3.2.3, Roussel et al. 2020), raster (v. 3.5-11, Hijmans 2021), rgdal (v. 1.5-28, Roger et al. 2021), and gstat (Pebesma 2004) packages in R (RStudio Team, 2021) using the CyVerse computing platform, an open access cyber infrastructure funded by the National Science Foundation. We used MS Excel (v. 2202) for creating bivariate graphs and ArcGIS Pro (v. 2.9.2) for preparing spatial maps.

Canopy Structural Complexity Metrics Analysis

We calculated five vertical complexity metrics (top rugosity, entropy, maximum canopy height, mean maximum canopy height, and vertical complexity index (VCI)) and four horizontal complexity metrics (rumple, vegetative area index (VAI), deep-gap fraction, and gap fraction) following LaRue and O'Leary NEON tutorial (LaRue, O'Leary 2021). Top rugosity, or roughness due to variation in the heights, was calculated as the standard deviation of the CHM (Parker & Russ 2004; Hardiman *et al.* 2011; Atkins *et al.* 2018)). Maximum canopy height was calculated as the maximum value of the canopy heights of each 1-m^2 boxes of the CHM, and the mean maximum canopy height was calculated by calculating the mean canopy height of all the 1-m^2 boxes of the CHM (Atkins *et al.* 2018). Rumple, a ratio of the outer canopy surface to ground area (1-km^2 plot) of the CHM (Parker *et al.* 2004), is closely related to canopy closure and was calculated by running the "rumple_index()" command from the LidR package on the CHM. Entropy, a diversity and evenness measure across every 1-m vertical slice of the CHM, was calculated by following the Shannon Diversity Index (Shannon 1948) approach to yield a normalized canopy Shannon Diversity Index of the height profiles in the CHM. Vertical

Complexity Index (VCI) is the normalized entropy. Vegetative Area Index (VAI) is closely related to leaf area index (LAI) and is the sum of leaf area density values in 1-m horizontal slices of the CHM. Deep-gap fraction is calculated by counting the number of pixels with no return value (ground) relative to the 1-km^2 plot, and cover fraction (= 1- deep-gap fraction) is an estimate of the canopy density of the plot (Zhao *et al.* 2012).

Correlation Analysis

We created a correlation matrix of the nine structural complexity metrics, Shannon Diversity Index, and species richness to quantify correlation among variables (Fig.1 in appendix). Because of the high correlation of structural complexity metrics, we chose four vertical (rugosity, maximum canopy height, mean maximum canopy height, and VCI) and three horizontal (rumple, VAI, deep-gap fraction and cover fraction) complexity metrics to highlight these patterns followed by one vertical and one horizontal metric to highlight the differences across forest types.

Non-metric Multidimensional Scaling (NMDS) Analysis

We calculated species abundance from species count data and explored the relative impact of different environmental and structural variables on species abundance at each trophic level using the ecodist (Goslee and Urban 2007) and vegan (Oksanen et al. 2020) packages in R. We used the Bray-Curtis distance method to estimate the influence of environmental variables on multitrophic species abundance grouped by three forest types. Environmental variables include nine structural complexity metrics mentioned above, climate (mean annual precipitation (MAP), mean annual temperature (MAT)), topographical (mean elevation), and geographical (latitude and longitude) variables. We then used the vegan package (Oksanen et al. 2020) to run an

ANOSIM (Analysis of Similarities) test to investigate the differences (p value < 0.05) between three forest types (Chapman & Underwood 1999).

Results

Effect of Vertical Structural Complexity on Multi-trophic Species Diversity

Overall, vertical complexity metrics showed high correlation with plant, beetle, and bird diversity (Fig. 1 in appendix). Plant diversity was correlated with entropy (r = 0.34), VCI (r = 0.32), rugosity (r = 0.29), and mean maximum canopy height (r = 0.17), while beetle diversity was correlated to mean maximum canopy height (r = 0.46), rugosity (r = 0.45), VCI (r = 0.40), and entropy (r = 0.36). Similarly, bird diversity was correlated with maximum canopy height (r = 0.46), rugosity (0.45), mean maximum canopy height (0.40), and entropy (r = 0.14). Species richness showed similar correlation to entropy, rugosity, and VCI across all trophic levels (Fig. 1 in appendix), but species diversity showed greater correlation with vertical complexity metrics than species richness.

Based on the correlation analysis, we selected the top three vertical complexity metrics (rugosity, maximum canopy height, and VCI) to investigate the relationship between vertical complexity and multi-trophic diversity (Fig. 2a-c). Multi-trophic diversity and species richness generally increased with increasing vertical structural complexity. For example, plant (R^2 = 0.29), beetle (R^2 = 0.26), and bird (R^2 = 0.19) diversity increased with increasing rugosity, and showed saturation response at mid complexity for each trophic level (Fig. 2a). Similar patterns were observed in the saturating response of multi-trophic diversity to increasing maximum canopy height (Fig. 2b). Maximum canopy height explained greater variation in bird diversity (R^2 = 0.28), than plant (R^2 = 0.20) and beetle (R^2 = 0.13) diversity (Fig. 2c). Multi-trophic diversity increased with VCI with a saturating response and greater diversity at mid level VCI,

while birds and beetles diversity almost linearly increased with increasing VCI and didn't show saturating effect (Fig. 2c). Overall, rugosity was the strongest indicator of plant and beetle diversity, whereas maximum canopy height was the strongest indicator of bird diversity. (Fig. 2, Fig. 1 in appendix).

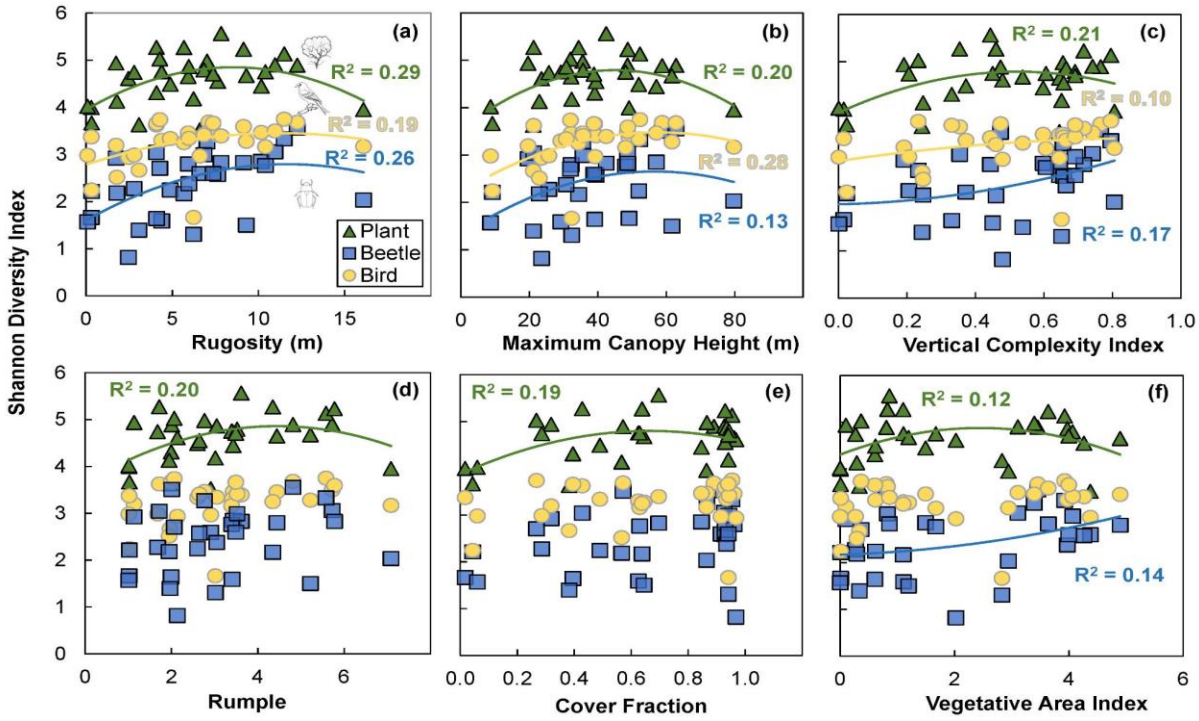

Figure 2: Relationship between structural complexity metrics (rugosity (a), maximum canopy height (b), vertical complexity index (c), rumple (d), cover fraction (e), and vegetative area index (f)) and multi-trophic diversity (Shannon Diversity Index of plants (a,d), beetles (b,e), and birds (c,f)) across 32 forested sites of the National Ecological Observatory Network. All models (see Table 2 in appendix for details) are significant at $p < 0.05$. Please see Table 1 in appendix for site information.

Effect of Horizontal Structural Complexity on Multi-trophic Species Diversity

Similar to vertical complexity metrics, multi-trophic diversity increased with increasing horizontal complexity metrics (Fig. 2d-f), but the relationships were weaker than the vertical complexity metrics (Fig. 2a-c) and differed across different trophic levels. For example, plant diversity was correlated with cover fraction ($r = 0.29$) and rumple ($r = 0.30$), whereas bird

diversity was correlated with rumple (r = 0.299) and VAI (r = 0. 23), and beetle diversity was correlated with VAI (r = 0.36) and cover fraction (0.31) (Fig. 1 in appendix).

Plant diversity showed strong positive relationships with horizontal complexity metrics with higher diversity at mid complexity (Fig. 2d-f). Beetle diversity showed a weak positive relationship with vegetative area index (Fig. 2f), but no relationship with rumple (Fig. 2d) and cover fraction (Fig. 2e), whereas bird diversity did not show any relationship with horizontal complexity metrics (Fig. 2d-f).

Effect of Forest Type on Structural Complexity and Multi-trophic Diversity Relationship

Plant and beetle diversity (and richness) showed greater variation across landscapes and distinct geographic patterns emerged (Fig. 3a-b) that highlighted the influence of forest types. For example, plant (Fig. 3a) and beetle (Fig. 3b) diversity (and richness) was lower in the western US, dominated by evergreen forests, and greater in the eastern US, dominated by deciduous and mixed forests. In contrast, bird diversity (and richness) was generally higher everywhere and lacked variation across sites, except four sites (three sites in Alaska and one in Hawaii) with smaller diversity (and richness) (Fig. 3c).

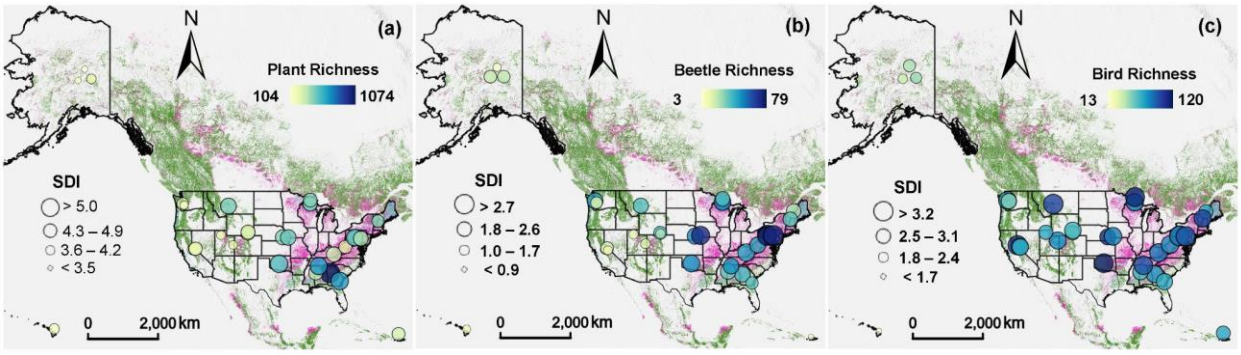

Figure 3: Spatial distribution of species richness and species diversity (Shannon Diversity Index, SDI) of plants (a), beetles (b), and birds (c) across 32 forested sites of the National Ecological Observatory Network. The World Forest (MDAUS BaseVue 2013) data in ArcGIS Pro is used for displaying different forest types (evergreen (dark green), deciduous (pink), and mixed (blue)).

The relationship between vertical complexity metric (e.g., rugosity) and plant diversity differed across different forest types ranging from strong (mixed forest-linear, evergreen forest-peaked) to no (deciduous-none) relationships (Fig. 4a). Bird and beetle diversity showed strong positive relationships with rugosity in deciduous (beetle: $R^2 = 0.76$, bird: $R^2 = 0.50$) and mixed (beetle: $R^2 = 0.74$, bird: $R^2 = 0.63$) forests, but no relationship in evergreen forests (Fig. 4b-c).

We found similar but weaker patterns in the relationships between horizontal complexity metrics (e.g., rumple) and multi-trophic diversity that differed across different forest types (Fig. 4d-f). Plant diversity showed strong positive relationships with rumple across all forest types (Fig. 4d), but beetle (Fig. 4e) and bird (Fig. 4f) diversity only showed strong relationships in mixed forest and no relationships in evergreen and deciduous forest (Fig. 4e,f).

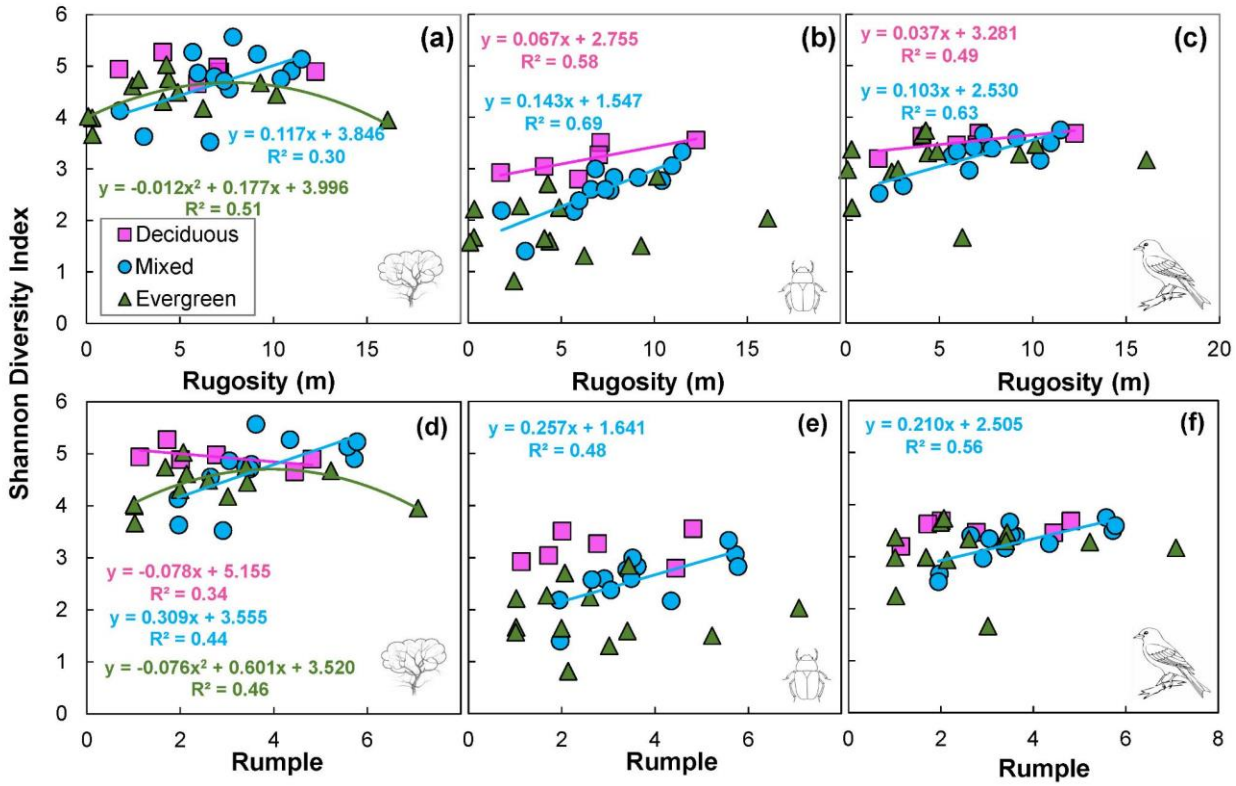

Figure 4: Relationship between structural complexity metric (rugosity (a-c), rumple (d-f)) and species diversity (Shannon Diversity Index) of plants (a,d), beetles (b,e), and birds (c,f) across 32 forested sites of the National Ecological Observatory Network. Regression models are specific to each forest type (equations colors match the forest type color coding), and only significant ($p < 0.05$) regressions are shown here. Please see Table 1 in appendix for site information.

Analysis of similarity (ANOSIM) showed that diversity in evergreen, mixed, and deciduous forests was significantly different ($p < 0.001$) across all trophic levels. The structural complexity metrics (entropy, mean maximum canopy height, VCI, cover fraction, VAI) and climatic variables (mean annual precipitation (MAP), mean annual temperature (MAT), and longitude) were clustered together and explained the clustering of deciduous and mixed forests (Fig. 5a-c). Evergreen forests were separated from mixed and deciduous forests, and this separation was explained by deep-gap fraction (DGF), mean elevation (ME), and latitude across all trophic levels (Fig. 5). MAT explained the most variation in plant diversity ($R^2 = 0.40$), followed by longitude ($R^2 = 0.34$) and structural complexity metrics (entropy ($R^2 = 0.32$), VAI ($R^2 = 0.32$), mean maximum canopy height ($R^2 = 0.29$), VCI ($R^2 = 0.29$), DGF ($R^2 = 0.26$), and cover fraction ($R^2 = 0.26$)) (Fig. 5a, NMDS stress = 0.17).

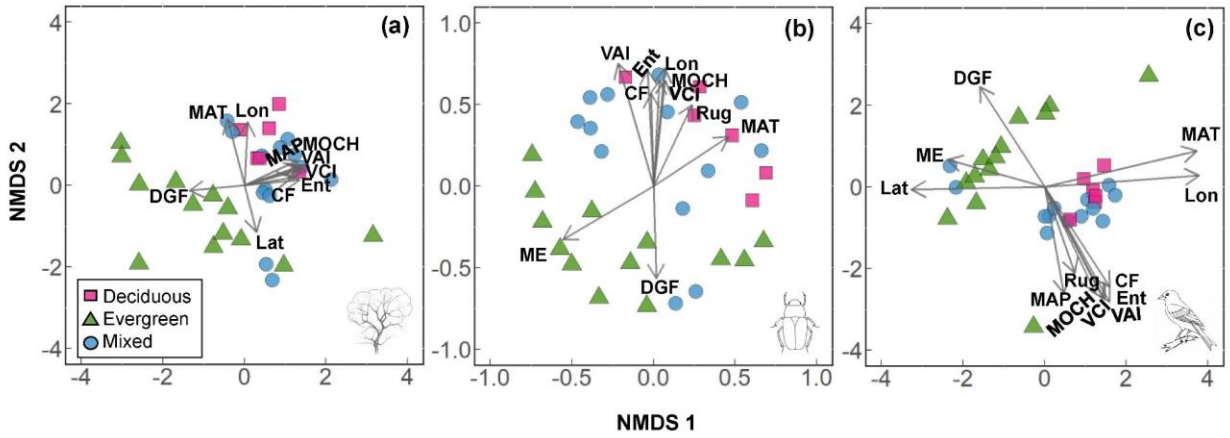

Figure 5: Non-metric multidimensional scaling (NMDS) plots of plants (a), beetles (b) and birds (c) abundance grouped by forest type across 32 forested sites of the National Ecological Observatory Network. The topographical (ME = Mean Elevation), geographical (Lat = Latitude and Long = Longitude), environmental (MAT = Mean Annual Temperature and MAP = Mean Annual Precipitation), and canopy structural complexity (VCI = Vertical Complexity Index, Ent = Entropy, Rug = Rugosity, MOCH = Mean Outer Canopy Height, and MCH = Max Canopy Height, VAL = Vegetative Area Index, DGF = Deep Gap Fraction, CF = Cover Fraction) variables are shown as vectors. Only significant ($p < 0.05$) variables are shown here.

The length of the arrows represent the strength of correlation. Structural complexity metrics explained 26 to 52 % of the variability in beetle diversity. VAI explained most of the variability (R^2 = 0.52) followed by longitude (R^2 = 0.45), mean maximum canopy height (R^2 = 0.43), VCI (R^2 = 0.36), entropy (R^2 = 0.36), mean elevation (R^2 = 0.35), canopy fraction (R^2 = 0.27), DGF (R^2 = 0.27), and MAT (R^2 = 0.26). Mean elevation, latitude, and DGF explained most variation in evergreen forests. The NMDS stress between the points was high (0.3) indicating lower confidence in results due to lack of enough data (Fig. 5b). MAT explained the most variation (R^2 = 0.64) in bird diversity (Fig. 5c, NMDS stress = 0.13) in deciduous and mixed forests, followed by longitude (R^2 = 0.63), entropy (R^2 = 0.46), VAI (R^2 = 0.43), VCI (R^2 = 0.42), canopy fraction (R^2 = 0.37), mean maximum canopy height (R^2 = 0.36), MAP (R^2 = 0.31), and rugosity (R^2 = 0.22). Latitude (R^2 = 0.47), mean elevation (R^2 = 0.26), and DGF (R^2 = 0.37) explained the variability in evergreen forest for bird diversity.

Discussion

Effect of Vertical Structural Complexity on Multi-trophic Species Diversity

Our results show that multi-trophic (plants, beetles, and birds) diversity generally increases with increasing vertical complexity (Fig. 2a-c), consistent with prior work showing that height-related complexity metrics explained most variation in plant (Gough *et al.* 2020; Torresani *et al.* 2020; Walter *et al.* 2021), beetle (Watts & Gibbs 2002), and avian species diversity (Zellweger *et al.* 2013; Carrasco *et al.* 2019). Previous work on height variation hypothesis suggests that increased variation in the height of the canopy is associated with greater diversity of tree species (Torresani *et al.* 2020) and avian species (MacArthur & MacArthur 1961; Pearson 1971; Yahner 1982; Zellweger *et al.* 2013; Carrasco *et al.* 2019) due to increased microhabitats. Our results extend these results to multi-trophic diversity and suggest

that the maximum species diversity across all trophic levels occurs at mid level of vertical complexity and not at the extreme ranges of complexity (Fig. 2a-c).

Effect of Horizontal Structural Complexity on Multi-trophic Species Diversity

Multi-trophic diversity increased with increasing horizontal complexity (Fig. 2d-f), similar to vertical complexity (Fig. 2a-c). Prior work on horizontal complexity effects on bird diversity showed mixed results from weaker (Zellweger *et al.* 2013) to stronger (Carrasco *et al.* 2019) relationships in comparison to vertical complexity. Similar patterns have been observed where beetle (Watts & Gibbs 2002) and bat (Erasmy *et al.* 2021) diversity increases with increasing canopy density, but some studies have shown that plant ((Vojík & Boublík 2018)) and bird (Gil-Tena *et al.* 2007) diversity decreases with increasing canopy density. Plant diversity had a better relationship with horizontal complexity than beetles, while bird diversity could not be explained by horizontal complexity (Fig. 2d-f). Plant and beetle diversity may be more affected by the amount of light and unique understory eco-climate that is promoted by high canopy cover or other horizontal complexity (Atkins *et al.* 2018) and does not affect bird diversity as much as vertical complexity that provides differential microhabitats.

Effects of Forest Type on Structural Complexity and Multi-trophic Diversity Relationship

Plant (Fig. 3a) and beetle (Fig. 3b) diversity (and richness) showed greater variation across eco-climatic regions and forest types, while bird (Fig. 3c) diversity (and richness) was less affected by ecoclimatic regions. Relationships between multi-trophic diversity and complexity differed in evergreen, deciduous and mixed forests (Fig. 4). We found a linear increase of species diversity with increasing structural complexity across all trophic levels in mixed forests. Mixed forests may have the optimal structural heterogeneity to support the increase of multi-trophic diversity. Previous work has shown that mixed forests are more resilient (Pretzsch *et al.* 2013)

and thus can sustain high diversity. In deciduous forests, beetle and bird diversity increased slowly with structural complexity, while plant diversity showed positive (Fig. 4a), negative (Fig. 4d), and no (Fig. 4b,c,e,f) relationship with increasing structural complexity. Structural complexity metrics and climatic variables explained the variation in species diversity in mixed and deciduous forests, but not evergreen forests.

Our study showed that longitude was the most important variable for multi-trophic diversity at continental level, while latitude explained the least variation. This may be because there is little difference in latitude as almost all sites were in the temperate zone, with the exception of sites in Puerto Rico and Hawaii. MAT and longitude explained the most variation in plant and bird diversity but had comparable strength with structural complexity in explaining beetle diversity. Not many studies have linked environmental and structural complexity metrics, but one study showed that canopy height was a better predictor than precipitation in primate species richness (Gouveia *et al.* 2014). More studies should be done to show how coupling climatic variables and structural complexity metrics may improve modeling techniques (Zellweger *et al.* 2013) for multi-trophic diversity, which has great implications as climate change pushes rapid changes in ecosystems.

Multi-trophic diversity and structural complexity showed no significant linear relationship with evergreen forests (Fig. 4). Our results suggest that structural complexity is less important for multi-trophic diversity in evergreen forests, but latitude, mean elevation, and deep-gap fraction can be used to explain the variability in multi-trophic diversity in evergreen forest. The lower species diversity in evergreen forests may be due to lower diversity and availability of food for consumers (e.g, diverse fruits, nuts, and worms for birds and litter for beetles) that are abundant in deciduous and mixed forests. We suspect that multi-trophic diversity in evergreen

forests may be more closely linked to topographical or climatic variables, though more research is needed to support this claim. Our results show that multi-trophic diversity has a better relationship with structural complexity metrics in deciduous and mixed forests and that the relationship of multi-trophic diversity with structural complexity becomes stronger when separated by forest type (Fig. 4).

Though LiDAR products do save time, money and resources, these data products are usually very large and require a lot of storage space and processing cores. Therefore, our use of 1000 m^2 study plots was suitable, convenient, and efficient as it helped us to bypass the aforementioned obstacles. We noticed the species accumulation curves for birds, beetles and plants did not plateau (Fig. 2 in appendix), which indicated that there was not enough data to just use one year for this study. We addressed this issue by using all available data from NEON. While conducting NMDS plots for beetles, we were forced to use another method as the lack of data for beetles inhibited us from running the metaMDS code from the vegan package. We used the ecodist package to conduct these plots, and it was revealed that the stress was high, another indication of lack of data. With time, NEON will continue to collect beetle information and this will hopefully curb some of the challenges we faced here when conducting future studies.

Conclusions

Our results show that multi-trophic diversity has a positive relationship with vertical and horizontal structural complexity as expected, but the relationships differ across different forest types. The strongest relationships between structural complexity and multi-trophic diversity were observed in mixed forests, which supported greater multi-trophic diversity compared to evergreen forests, but lower than deciduous forests across similar structural complexity gradients. Deciduous and evergreen forests showed weak or no relationship between structural

complexity metrics and species diversity. This work expanded the structural complexity metric research by including both horizontal and vertical complexity metrics as biodiversity indicators, and highlighted the differences across different forest types, which may be an asset when making management decisions.

Literature Cited

Angelsen, A., Jagger, P., Babigumira, R., Belcher, B., Hogarth, N.J., Bauch, S., *et al.* (2014). Environmental Income and Rural Livelihoods: A Global-Comparative Analysis. *World Dev.*, Forests, Livelihoods, and Conservation, 64, S12–S28.

Atkins, J.W., Fahey, R.T., Hardiman, B.S. & Gough, C.M. (2018). Forest Canopy Structural Complexity and Light Absorption Relationships at the Subcontinental Scale. *J. Geophys. Res. Biogeosciences*, 123, 1387–1405.

Barrows, C.W., Swartz, M.B., Hodges, W.L., Allen, M.F., Rotenberry, J.T., Li, B.-L., *et al.* (2005). A Framework for Monitoring Multiple-Species Conservation Plans. *J. Wildl. Manag.*, 69, 1333–1345.

Camargo, N.F. de, Sano, N.Y. & Vieira, E.M. (2018). Forest vertical complexity affects alpha and beta diversity of small mammals. *J. Mammal.*, 99, 1444–1454.

Carrasco, L., Giam, X., Papeş, M. & Sheldon, K.S. (2019). Metrics of Lidar-Derived 3D Vegetation Structure Reveal Contrasting Effects of Horizontal and Vertical Forest Heterogeneity on Bird Species Richness. *Remote Sens.*, 11, 743.

Chapman, M. & Underwood, A. (1999). Ecological patterns in multivariate assemblages:information and interpretation of negative values in ANOSIM tests. *Mar. Ecol. Prog. Ser.*, 180, 257–265.

Chow, T.E. & Hodgson, M.E. (2009). Effects of lidar post-spacing and DEM resolution to mean slope estimation. *Int. J. Geogr. Inf. Sci.*, 23, 1277–1295.

Connell, J.H. (1978). Diversity in Tropical Rain Forests and Coral Reefs. *Science*, 199, 1302–1310.

Critchlow, R., Cunningham, C.A., Crick, H.Q.P., Macgregor, N.A., Morecroft, M.D., Pearce-Higgins, J.W., *et al.* (2022). Multi-taxa spatial conservation planning reveals similar priorities between taxa and improved protected area representation with climate change. *Biodivers. Conserv.*

Ehbrecht, M., Seidel, D., Annighöfer, P., Kreft, H., Köhler, M., Zemp, D.C., *et al.* (2021). Global patterns and climatic controls of forest structural complexity. *Nat. Commun.*, 12, 519.

Erasmy, M., Leuschner, C., Balkenhol, N. & Dietz, M. (2021). Shed light in the dark – How do natural canopy gaps influence temperate bat diversity and activity? *For. Ecol. Manag.*, 497, 119509.

FAO & UNEP. (2020). *The State of the World's Forests 2020*. FAO and UNEP.

Fleishman, E., Murphy, D.D. & Brussard, P.F. (2000). A New Method for Selection of Umbrella Species for Conservation Planning. *Ecol. Appl.*, 10, 569–579.

Franklin, J.F. (1981). *Ecological Characteristics of Old-growth Douglas-fir Forests*. U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station.

Gil-Tena, A., Saura, S. & Brotons, L. (2007). Effects of forest composition and structure on bird species richness in a Mediterranean context: Implications for forest ecosystem management. *For. Ecol. Manag.*, 242, 470–476.

Gough, C.M., Atkins, J.W., Fahey, R.T., Hardiman, B.S. & LaRue, E.A. (2020). Community and structural constraints on the complexity of eastern North American forests. *Glob. Ecol. Biogeogr.*, 29, 2107–2118.

Gouveia, S.F., Villalobos, F., Dobrovolski, R., Beltrão-Mendes, R. & Ferrari, S.F. (2014). Forest structure drives global diversity of primates. *J. Anim. Ecol.*, 83, 1523–1530.

Guo, X., Coops, N.C., Tompalski, P., Nielsen, S.E., Bater, C.W. & John Stadt, J. (2017). Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data. *Ecol. Inform.*, 38, 50–61.

Hardiman, B.S., Bohrer, G., Gough, C.M., Vogel, C.S. & Curtis, P.S. (2011). The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. *Ecology*, 92, 1818–1827.

Hardiman, B.S., LaRue, E.A., Atkins, J.W., Fahey, R.T., Wagner, F.W. & Gough, C.M. (2018). Spatial Variation in Canopy Structure across Forest Landscapes. *Forests*, 9, 474.

Houghton, R.A. (2005). Aboveground Forest Biomass and the Global Carbon Balance. *Glob. Change Biol.*, 11, 945–958.

Ishii, H.T., Tanabe, S. & Hiura, T. (2004). Exploring the Relationships Among Canopy Structure, Stand Productivity, and Biodiversity of Temperate Forest Ecosystems, 14.

Kern, C.C., Montgomery, R.A., Reich, P.B. & Strong, T.F. (2014). Harvest-Created Canopy Gaps Increase Species and Functional Trait Diversity of the Forest Ground-Layer Community. *For. Sci.*, 60, 335–344.

LaRue, E.A., Wagner, F.W., Fei, S., Atkins, J.W., Fahey, R.T., Gough, C.M., *et al.* (2020). Compatibility of Aerial and Terrestrial LiDAR for Quantifying Forest Structural Diversity. *Remote Sens.*, 12, 1407.

Lunch, C., Laney, C., Mietkiewicz, N., Sokol, E., Cawley, K., and NEON

(National Ecological Observatory Network) (2021). neonUtilities: Utilities for Working with NEON Data. R package version 2.1.3. https://CRAN.R-project.org/package=neonUtilities

MacArthur, R.H. & MacArthur, J.W. (1961). On Bird Species Diversity. *Ecology*, 42, 594–598.

McElhinny, C., Gibbons, P., Brack, C. & Bauhus, J. (2005). Forest and woodland stand structural complexity: Its definition and measurement. *For. Ecol. Manag.*, 218, 1–24.

Millennium Ecosystem Assessment (Program) (Ed.). (2005). *Ecosystems and human well-being: synthesis*. Island Press, Washington, DC.

Molina, R., Marcot, B.G. & Lesher, R. (2006). Protecting Rare, Old-Growth, Forest-Associated Species under the Survey and Manage Program Guidelines of the Northwest Forest Plan. *Conserv. Biol.*, 20, 306–318.

Müller, J., Brandl, R., Brändle, M., Förster, B., de Araujo, B.C., Gossner, M.M., *et al.* (2018). LiDAR-derived canopy structure supports the more-individuals hypothesis for arthropod diversity in temperate forests. *Oikos*, 127, 814–824.

Neary, D.G., Ice, G.G. & Jackson, C.R. (2009). Linkages between forest soils and water quality and quantity. *For. Ecol. Manag.*, Forest Soil Science: Celebrating 50 Years of Research on Properties, Processes and Management of Forest Soils Forest Soil Science: Celebrating 50 Years of Research on Properties, Processes and Management of Forest Soils Forest Soil Science: Celebrating 50 Years of Research on Properties, Processes and Management of Forest Soils Forest Soil Science: Celebrating 50 Years of Research on Properties, Processes and Management of Forest Soils, 258, 2269–2281.

NEON (National Ecological Observatory Network). Breeding landbird point counts, RELEASE-2021 (DP1.10003.001). https://doi.org/10.48443/88sy-ah40. Dataset accessed from https://data.neonscience.org on September 29, 2021.

NEON (National Ecological Observatory Network). Ground beetles sampled from pitfall traps, RELEASE-2021 (DP1.10022.001). https://doi.org/10.48443/xgea-hw23. Dataset accessed from https://data.neonscience.org on September 29, 2021.

NEON (National Ecological Observatory Network). Plant presence and percent cover, RELEASE-2021 (DP1.10058.001). https://doi.org/10.48443/pr5e-1q60. Dataset accessed from https://data.neonscience.org on September 29, 2021.

NEON (National Ecological Observatory Network). Elevation - LiDAR, RELEASE-2022 (DP3.30024.001). https://doi.org/10.48443/ymmp-fr93. Dataset accessed from https://data.neonscience.org on September 29, 2021.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs E., and Helene Wagner, H. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Parker, G.G., Harmon, M.E., Lefsky, M.A., Chen, J., Pelt, R.V., Weis, S.B., *et al.* (2004). Three-dimensional Structure of an Old-growth Pseudotsuga-Tsuga Canopy and Its Implications for Radiation Balance, Microclimate, and Gas Exchange. *Ecosystems*, 7, 440–453.

Parker, G.G. & Russ, M.E. (2004). The canopy surface and stand development: assessing forest

canopy structure and complexity with near-surface altimetry. *For. Ecol. Manag.*, 189, 307–315.

Pearson, D.L. (1971). Vertical Stratification of Birds in a Tropical Dry Forest. *The Condor*, 73, 46–55.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences, 30: 683-691.

Poiani, K.A., Merrill, M.D. & Chapman, K.A. (2001). Identifying Conservation-Priority Areas in a Fragmented Minnesota Landscape Based on the Umbrella Species Concept and Selection of Large Patches of Natural Vegetation. *Conserv. Biol.*, 15, 513–522.

Pretzsch, H., Schütze, G. & Uhl, E. (2013). Resistance of European tree species to drought stress in mixed versus pure forests: evidence of stress release by inter-specific facilitation. *Plant Biol.*, 15, 483–495.

RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Roussel, J.R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Sánchez Meador, A., Bourdon, J.F., De Boissieu, F., Achim, A. (2020). lidR : An R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment, 251 (August), 112061. <doi:10.1016/j.rse.2020.112061>.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27, 379–423.

Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C.V., Costa, F.R.C., *et al.* (2012). Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecol. Lett.*, 15, 1406–1414.

Suter, W., Graf, R.F. & Hess, R. (2002). Capercaillie (Tetrao urogallus) and Avian Biodiversity: Testing the Umbrella-Species Concept. *Conserv. Biol.*, 16, 778–788.

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Hauffe, H.C., Heym, M., *et al.* (2020). Height variation hypothesis: A new approach for estimating forest species diversity with CHM LiDAR data. *Ecol. Indic.*, 117, 106520.

Vojík, M. & Boublík, K. (2018). Fear of the dark: decline in plant diversity and invasion of alien species due to increased tree canopy density and eutrophication in lowland woodlands. *Plant Ecol.*, 219, 749–758.

Walter, J.A., Stovall, A.E.L. & Atkins, J.W. (2021). Vegetation structural complexity and biodiversity in the Great Smoky Mountains. *Ecosphere*, 12, e03390.

Watts, C.H. & Gibbs, G.W. (2002). Revegetation and its Effect on the Ground-Dwelling Beetle Fauna of Matiu-Somes Island, New Zealand. *Restor. Ecol.*, 10, 96–106.

Wolf, J.A., Fricker, G., Meyer, V., Hubbell, S.P., Gillespie, T.W. & Saatchi, S.S. (2012). Plant Species Richness is Associated with Canopy Height and Topography in a Neotropical Forest. *Remote Sens.*, 4, 4010–4021.

Yahner, R.H. (1982). Avian Use of Vertical Strata and Plantings in Farmstead Shelterbelts. *J. Wildl. Manag.*, 46, 50–60.

Zellweger, F., Braunisch, V., Baltensweiler, A. & Bollmann, K. (2013). Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. *For. Ecol. Manag.*, 307, 303–312.

Zhao, F., Strahler, A.H., Schaaf, C.L., Yao, T., Yang, X., Wang, Z., *et al.* (2012). Measuring gap fraction, element clumping index and LAI in Sierra Forest stands using a full-waveform ground-based lidar. *Remote Sens. Environ.*, 125, 73–79.

Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C. & Gerard, P.D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sens. Environ.*, 87, 171–182.

Supplementary Information

Multi-trophic biodiversity increases with increasing structural complexity of forest canopy

Ayanna St. Rose[1*] and Kusum Naithani[1]

[1]Department of Biology, University of Arkansas, Fayetteville, AR 72701 USA

Table 1: Site names,site ID, domain ID, state, latitude (Lat), longitude (Lon), mean elevation (ME) mean annual temperature (MAT), mean annual precipitation (MAP) measure in millimeters, and forest type (evergreen, deciduous and mixed forests). Forests were classified as mixed if there were more than one forest type listed.

Site	Site ID	Domain	State	Lat (°)	Long (°)	ME (m)	MAT (°C)	MAP (mm)	Forest Type	Year LiDAR Obtained
Abby Road NEON	**ABBY**	16	WA	45.762439	-122.33032	365	10	2451	Evergreen	2021
Bartlett Experimental Forest NEON	**BART**	1	NH	44.063889	-71.287375	274	6.2	1325	Mixed	2019
Blandy Experimental Farm NEON	**BLAN**	2	VA	39.033698	-78.041788	183	12.1	983	Deciduous	2021
Caribou-Poker Creeks Research Watershed	**BONA**	19	AK	65.15401	-147.50258	230	-3	262	Mixed	2018
Lyndon B. Johnson National Grassland NEON	**CLBJ**	11	TX	33.40123	-97.57	272	17.5	926	Deciduous	2021

Delta Junction NEON	**DEJU**	19	AK	63.88112	-145.75136	517	-3	305	Mixed	2019
Great Smoky Mountains National Park NEON	**GRSM**	7	TN	35.68896	-83.50195	575	13.1	1375	Mixed	2018
Guanica Forest NEON	**GUAN**	4	PR	17.96955	-66.8687	125	23	840	Evergreen	2018
Harvard Forest & Quabbin Watershed NEON	**HARV**	1	MA	42.53691	-72.17265	348	7.4	1199	Mixed	2019
Healy NEON	**HEAL**	19	AK	63.875798	-149.21335	677	-1.3	385	Evergreen	2019
The Jones Center At Ichauway NEON	**JERC**	3	GA	31.194839	-84.468623	47	19.2	1308	Mixed	2019
Konza Prairie Biological Station NEON	**KONZ**	6	KS	39.100774	-96.563075	414	12.4	870	Deciduous	2020
Lenoir Landing NEON	**LENO**	8	AL	31.853861	-88.161181	13	18.1	1386	Mixed	2021
Mountain Lake Biological Station NEON	**MLBS**	7	VA	37.378314	-80.524847	1170	8.8	1227	Deciduous	2021

Moab NEON	**MOAB**	13	UT	38.248283	-109.38827	1799	10.1	319	Evergreen	2021
Niwot Ridge NEON	**NIWO**	13	CO	40.05425	-105.58237	3490	0.3	1005	Evergreen	2020
Onaqui NEON	**ONAQ**	15	UT	40.177599	-112.45245	1662	9	288	Evergreen	2021
Ordway-Swisher Biological Station NEON	**OSBS**	3	FL	29.689282	-81.993431	46	20.9	1302	Mixed	2019
Pu'u Maka'ala Natural Area Reserve NEON	**PUUM**	20	HI	19.55309	-155.31731	1685	12.7	2657	Evergreen	2020
Rocky Mountains NEON	**RMNP**	10	CO	40.275903	-105.54596	2742	2.9	731	Evergreen	2018
Smithsonian Conservation Biology Institute NEON	**SCBI**	2	VA	38.892925	-78.139494	352	11.6	1126	Mixed	2021
Smithsonian Environmental Research Center NEON	**SERV**	2	MD	38.890131	-76.560014	33	13.6	1075	Deciduous	2021
San Joaquin Experimental Range NEON	**SJER**	17	CA	37.10878	-119.73228	400	16.4	540	Evergreen	2021

Soaproot Saddle NEON	**SOAP**	17	CA	37.03337	-119.26219	1210	13.4	900	Evergreen	2021
Steigerwaldt-Chequamegon NEON	**STEI**	5	WI	45.50894	-89.58637	476	4.8	797	Mixed	2020
Talladega National Forest NEON	**TALL**	8	AL	32.95047	-87.393259	166	17.2	1383	Mixed	2021
Lower Teakettle NEON	**TEAK**	17	CA	37.00583	-119.00602	2149	8	1223	Evergreen	2021
Treehaven NEON	**TREE**	5	WI	45.49369	-89.58571	467	4.8	797	Mixed	2020
University of Kansas Field Station NEON	**UKFS**	6	KS	39.040431	-95.19215	322	12.7	990	Deciduous	2020
University of Notre Dame Environmental Research Center NEON	**UNDE**	5	MI	46.23391	-89.537254	521	4.3	802	Mixed	2020
Wind River Experimental Forest NEON	**WREF**	16	WA	45.82049	-121.95191	351	9.2	2225	Evergreen	2021
Yellowstone National Park NEON	**YELL**	12	WY	44.95348	-110.53914	2133	3.4	493	Evergreen	2020

Table 2: Model parameters associated with polynomial models ($y = ax^2 + bx + c$) in Figure 2

Figure	Organism	a	b	c	R^2
2a	Plant	-0.0124	0.2106	3.9515	0.29
	Bird	-0.0054	0.1196	2.7843	0.19
	Beetle	-0.0088	0.2044	1.6035	0.26
2b	Plant	-0.0006	0.0557	3.5399	0.20
	Bird	-0.0004	0.0437	2.2199	0.28
	Beetle	-0.0004	0.0472	1.3161	0.13
2c	Plant	-3.1642	3.3155	3.921	0.21
	Bird	-0.5867	1.0681	2.8998	0.10
	Beetle	1.2324	0.1515	1.9864	0.17
2d	Plant	-0.061	0.5428	3.6525	0.20
	Bird	---	---	---	---
	Beetle	---	---	---	---
2e	Plant	-2.0955	2.7889	3.8697	0.19
	Bird	---	---	---	---
	Beetle	---	---	---	---
2f	Plant	-0.095	0.4654	4.3075	0.12

Bird	---	---	---	---
Beetle	0.0263	0.0386	2.1687	0.14

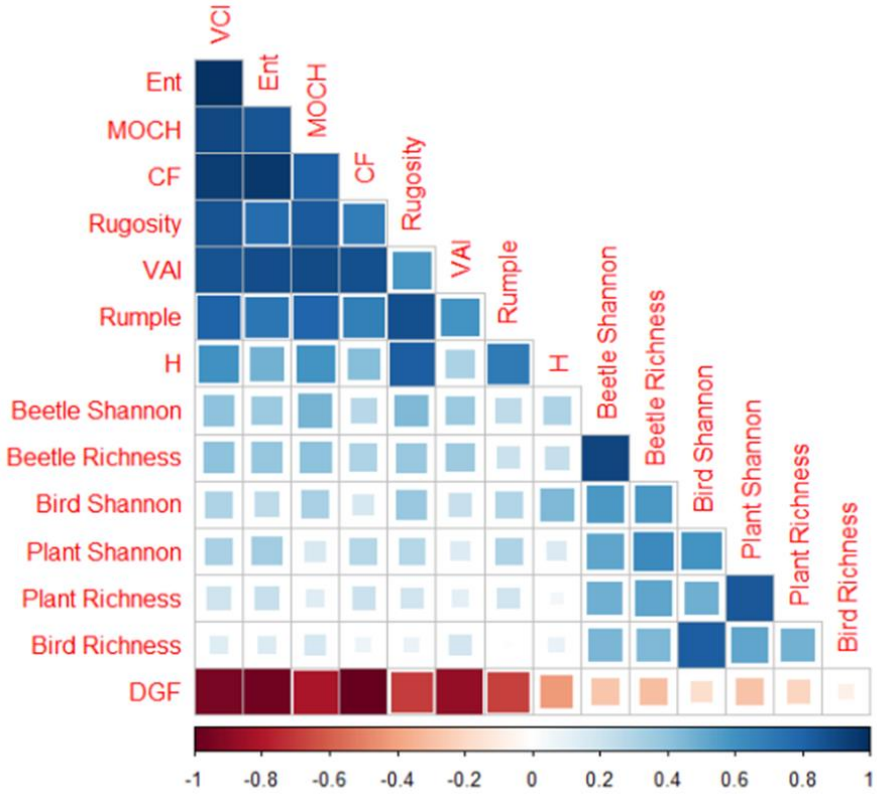

Figure 1: Correlation plots of canopy structural complexity metrics (Vertical Complexity Index (VCI), entropy (Ent), mean maximum canopy height (MOCH), maximum canopy height (H), Vegetative Area Index (VAI), canopy fraction (CF), deep-gap fraction (DGF), and rumple) and beetle, bird and plant diversity and richness. Size and shade indicates level of correlation (Smaller squares and lighter shade represent smaller correlation values and Larger squares and darker shades represent larger) correlation values. Blue represents positive correlation, while red represents negative correlation.

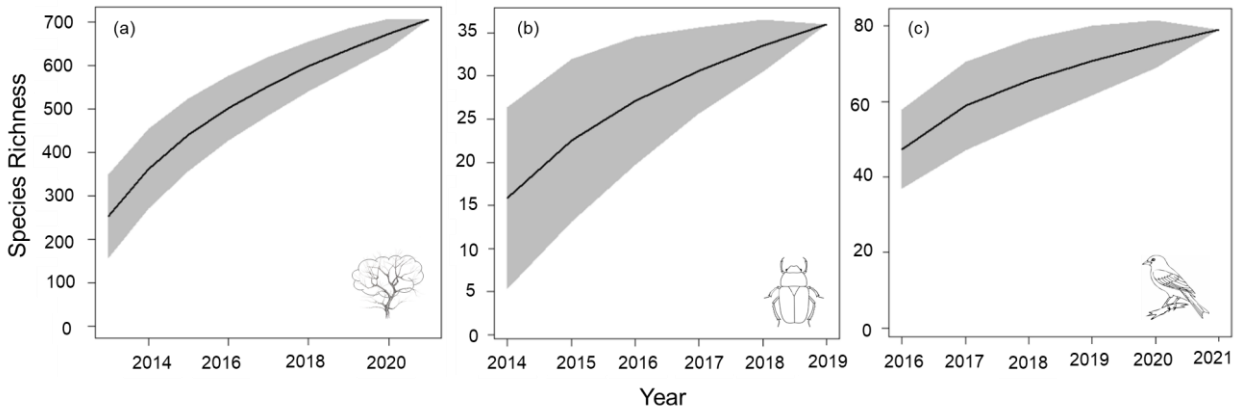

Figure 2: Species accumulation curve for plants (a), beetles (b), and birds (c) in the Ordway-Swisher Biological Station NEON site using all available data.

Data Processing Code

---
title: "Bird Diversity"
author: "Ayanna St. Rose"
date: "8/6/2021"
output: pdf_document

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

#Download and load necessary packages
```{r}
install.pakages("tidyverse")
install.packages("ggpubr")
install.pakages("tidyverse")
install.packages("neonUtilities")
install.packages("vegan")
install.packages("dplyr")
install.packages("janitor")
install.packages("ggExtra")
install.packages("plotly")

library(plotly)
library(ggExtra)
library(janitor)
library(plyr)
library(dplyr)
library(vegan)
library(tidyverse)
library(neonUtilities)
library(ggpubr)
library(colorBlindness)
library(ggplot2)
library(devtools)
library(tidyr)
```

#Get data from neon site
```{r}
bird_data= loadByProduct(dpID = "DP1.10003.001",
            site= c("BONA", "CLBJ", "DEJU", "GUAN", "HARV","KONZ", "LENO",
           "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
           "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
```

```
          "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
          "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
          "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
          "TECR", "TOMB", "TREE", "WALK", "WLOU"),
            startdate = "2013-01", enddate = "2021-09",
            check.size = FALSE)

```

#I think GUAN is an incomplete site, we could remove it
#Extract expert taxonomic ID from data list
#Drop UID because it will be unique for repeated data
```{r}
bird_taxa = bird_data$brd_countdata

#Drop UID
bird_taxa = subset(bird_taxa, select = -c(uid))

bird_fieldData = bird_data$brd_perpoint

#Drop UID
bird_fieldData = subset(bird_fieldData, select = -c(uid))
```


#Merge taxonomic and field data by namedLocation
```{r}
bird_full_data = merge(bird_taxa, bird_fieldData,
                by = c("siteID", "domainID", "plotID", "pointID", "eventID",
                      "plotType", "startDate", "namedLocation", "publicationDate") ,
                all = TRUE, check.duplicates = FALSE)

#Remove rows where taxon ID says NA
#This will have the same contents as the bird taxa dataframe,
#with added columns from field table
bird_full_data = bird_full_data %>% drop_na(taxonID)
```

#Write bird total diversity that includes taxonomic information and field information
#Upload to box
```{r}
write.csv(bird_full_data, "bird_full_data.csv", row.names = FALSE)
```

#There isn't a count for birds, more so a presence and absence data
###Goal:

#####Run species richness for each site
#####Create data frame with different species for each site
#####Pseudoabundance is how many times the species is repeated (run for each site)
#####Count number of times a unique species occurs in each site

```{r}
#bird occurrence only sorted by siteID
#Quantify bird occurrence
#Group by specific epithet and site ID
install.packages("janitor")
library(janitor)
bird_occ = bird_full_data %>%
  tabyl(taxonID, siteID, sort = TRUE)

#Show species distribution for each species for each site
bird_occ = data.frame(t(bird_occ))

names(bird_occ) <- as.matrix(bird_occ[1,])
bird_occ = bird_occ[-1,]
bird_occ[] <- lapply(bird_occ, function(x) type.convert(as.character((x))))

#Count the number of non-zero cells for each row
#This will give you the richness per site
#This will appear in the final column of the dataset
bird_occ$count <- rowSums(bird_occ!=0)

#Rename count as species richness
names(bird_occ)[447] = "species_richness"

#Calculate abundance
#Sum all rows except species_richness column
row_sum = data.frame(rowSums(bird_occ[, -447]))

#rename column
names(row_sum)[1] = "total_organisms"

#rename 0th column
bird_occ<-tibble::rownames_to_column(bird_occ, "siteID")

#rename 0th column
row_sum<-tibble::rownames_to_column(row_sum, "siteID")

#merge row_sum dataframe with bird_occ dataframe by
bird_occ = merge(bird_occ, row_sum,
                 by = "siteID",
                 all = TRUE, check.duplicates = FALSE)
```

#This table contains occurrence and richness values (last 2 columns)
```

#Write Occurrence data to CSV file.
```{r}
write.csv(bird_occ, "bird_occ.csv", row.names = FALSE)
```

```{r}
site_info = NEON_Field_Site_Metadata_20210226_0

# Select interest sites
site_info = site_info[site_info$field_site_id %in% c("BONA", "CLBJ", "DEJU", "GUAN",
        "HARV","KONZ", "LENO", "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
        "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
        "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
        "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
        "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
        "TECR", "TOMB", "TREE", "WALK", "WLOU"), ]

#Merge info for interest site with data
bird_field_data = merge(bird_occ, site_info, by.x = "siteID", by.y = "field_site_id")

#Export previous file
write.csv(bird_field_data, "bird_field_data.csv", row.names = FALSE)


#Import table into GIS and Make layers in GIS with species richness and total
#number of species. This helps to visualize diversity.
#Still have to figure out of to graph it
```


#Visualize species richness
```{r}
p1 = ggplot(bird_occ, aes(x = siteID, y = species_richness)) +
  geom_histogram(stat = "identity") +
  labs (title = "Histogram of Species Richness per Site", x= "site ID", y = "Number of Species") +
     theme_bw() +
      theme(panel.grid = element_blank(),
          axis.text = element_text(size = 12),
          axis.text.x = element_text(angle = 45, hjust = 1),
          axis.title = element_text(size = 12),
          plot.title = element_text(size = 14, hjust = 0.5, face = "bold"))
```
```

```
plot(p1)
```

```
#I would like to have NLCD class on there, Domain, having some trouble
#organizing the data, when I filter to include NLCD class and elevation
#I get repeating organisms, view trial below
```{r}
trial = bird_full_data %>%
  count(taxonID, siteID, nlcdClass, decimalLatitude, decimalLongitude, domainID, sort = TRUE)

#Trial is a shows the specific epithet sorted by siteID,
#nlcdClass, decimalLatitude, decimalLongitude, and domains
#I included the lat and long to plot it on the US map

ggplot(trial, aes(x=domainID, y = n, fill = nlcdClass)) +
  geom_boxplot()

```

```
#Create another dataframe which contains diversity information
#Calculate Diversity indices using the Vegan Package in R
```{r}
bird_sp_div = t(bird_full_data %>%
  tabyl(taxonID, siteID, sort = TRUE))

#Convert to Data frame
bird_sp_div = data.frame(bird_sp_div)


#Place specific epithet as row titles
names(bird_sp_div) <- as.matrix(bird_sp_div[1,])
bird_sp_div = bird_sp_div[-1,]
bird_sp_div[] <- lapply(bird_sp_div, function(x) type.convert(as.character((x))))

#Calculate Diversity Indices
#Simpson Diversity
simpson_diversity = data.frame(diversity(bird_sp_div, "simpson"))
simpson_diversity <-tibble::rownames_to_column(simpson_diversity, "siteID")

#eStar refers to the evenness, dStar = true diversity
evenness = data.frame(eventstar(bird_sp_div))
evenness<-tibble::rownames_to_column(evenness, "siteID")

#Shannon Diversity
```

shannon_diversity = data.frame(diversity(bird_sp_div, "shannon"))
shannon_diversity <-tibble::rownames_to_column(shannon_diversity, "siteID")

#Merge all 3 diversity tables

diversity_index = merge(shannon_diversity, simpson_diversity, by = "siteID")
diversity_index = merge(diversity_index, evenness, by = "siteID")

names(diversity_index)[1] = "siteID"
names(diversity_index)[2] = "shannon"
names(diversity_index)[3] = "simpson"
names(diversity_index)[5] = "evenness"
names(diversity_index)[6] = "true_div"

bird_field_data = merge(diversity_index, bird_field_data, by = "siteID")
write.csv(bird_field_data, "bird_field_data.csv", row.names = FALSE)

#merge richness and diversity indices
richness = data.frame(bird_occ[, -c(2:447)])
bird_diversity_indices = merge(diversity_index, richness, by = "siteID")
write.csv(bird_diversity_indices, "bird_div_indices.csv")
```

---
title: "Plant Diversity"
author: "Ayanna St. Rose"
date: "7/21/2021"
output: pdf_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
#ft stands for forest type
#Guanica Forest (GUAN) is a tropical site in Puerto Rico, ft= evergreen forest
#Mountain Lake Biological Station (MLBS) is a site in Virginia, ft= deciduous forest
#Wind River Experimental Forest (WREF) is a site in Washington, ft= evergreen forest

#Download and load necessary packages
```{r}
install.packages("tidyverse")
install.packages("neonUtilities")
install.packages("multcomp")
install.packages("agricolae")

library(agricolae)
```
```

```
library(multcomp)
library(neondiversity)
library(dbplyr)
library(tidyverse)
library(neonUtilities)
```

#Download plant data using neonutilities
```{r}
plant_data = loadByProduct(dpID = "DP1.10058.001",
            site= c("BONA", "CLBJ", "DEJU", "GUAN", "HARV","KONZ", "LENO",
           "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
           "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
           "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
           "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
           "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
           "TECR", "TOMB", "TREE", "WALK", "WLOU"),
            startdate = "2013-01", enddate = "2021-09",
            check.size = FALSE)
```

#Extract expert taxonomic ID from data list
```{r}
#Get data set for presence of species is observed in six 10m2 subplots
#and four 100m2 subplots per plot
plant_taxa = plant_data$div_10m2Data100m2Data

#Drop UID
plant_taxa = subset(plant_taxa, select = -c(uid))

#Get dataset for  presence and percent cover of plant species and ground
#cover is observed in six 1m2 subplots per plot
plant_taxa1 = plant_data$div_1m2Data

#Drop UID
plant_taxa1 = subset(plant_taxa1, select = -c(uid))

#Combine for a list of species at the 400m2 plot scale.
#For no duplicates, omit "by" command
plant_full_data = merge(plant_taxa, plant_taxa1,
                all = TRUE, no.dups = TRUE)

#Remove rows where taxon ID says NA
plant_full_data = plant_full_data %>% drop_na(taxonID)
```

#Write Plant total diversity including taxonomic information
```{r}
write.csv(plant_full_data, "plant_full_data.csv", row.names = FALSE)
```
#There isn't a count for plants, more so a presence and absence data
###Goal:
#####Run species richness for each site
#####Create data frame with different species for each site
#####Pseudoabundance is how many times the species is repeated (run for each site)
#####Count number of times a unique species occurs in each site

```{r}
#plant occurrence only sorted by siteID
#Quantify plant occurrence
#Group by specific epithet and site ID

plant_occ = plant_full_data %>%
  tabyl(taxonID, siteID, sort = TRUE)

#Show species distribution for each species for each site
plant_occ = data.frame(t(plant_occ))

names(plant_occ) <- as.matrix(plant_occ[1,])
plant_occ = plant_occ[-1,]

#Remove Unknown species (indicated as 2Plant in dataset)
plant_occ = plant_occ[, !(colnames(plant_occ) %in% c("2PLANT","2PLANT-H", "2PLANT-
          S"))]


plant_occ[] <- lapply(plant_occ, function(x) type.convert(as.character((x))))

#Count the number of non-zero cells for each row
#This will give you the richness per site
#This will appear in the final column of the dataset
plant_occ$count <- rowSums(plant_occ!=0)

#Rename count as species richness
names(plant_occ)[5226] = "species_richness"

#Calculate abundance
#Sum all rows except species_richness column
row_sum = data.frame(rowSums(plant_occ[, -5226]))

#rename column
names(row_sum)[1] = "total_organisms"
```

```
#rename 0th column
plant_occ<-tibble::rownames_to_column(plant_occ, "siteID")

#rename 0th column
row_sum<-tibble::rownames_to_column(row_sum, "siteID")

#merge row_sum dataframe with plant_occ dataframe by
plant_occ = merge(plant_occ, row_sum,
                  by = "siteID",
                  all = TRUE, check.duplicates = FALSE)

#This table contains occurrence and richness values (last 2 columns)
```

#Write Occurrence data to CSV file.

```{r}
write.csv(plant_occ, "plant_occ.csv", row.names = FALSE)
```

```{r}
site_info = NEON_Field_Site_Metadata_20210226_0

# Select interest sites
site_info = site_info[site_info$field_site_id %in% c("BONA", "CLBJ", "DEJU", "GUAN",
          "HARV","KONZ", "LENO", "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
          "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
          "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
          "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
          "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
          "TECR", "TOMB", "TREE", "WALK", "WLOU"), ]

#Merge info for interest site with data
plant_field_data = merge(plant_occ, site_info, by.x = "siteID", by.y = "field_site_id")

#Export previous file
write.csv(plant_field_data, "plant_field_data.csv", row.names = FALSE)


#Import table into GIS and Make layers in GIS with species richness and total
#number of species. This helps to visualize diversity.
#Still have to figure out of to graph it
```

#Visualize species richness
```{r}
p1 = ggplot(plant_occ, aes(x = siteID, y = species_richness)) +
  geom_histogram(stat = "identity") +
  labs (title = "Histogram of Species Richness per Site", x= "site ID", y = "Number of Species") +
     theme_bw() +
      theme(panel.grid = element_blank(),
            axis.text = element_text(size = 12),
            axis.text.x = element_text(angle = 45, hjust = 1),
            axis.title = element_text(size = 12),
            plot.title = element_text(size = 14, hjust = 0.5, face = "bold"))

plot(p1)
```

#Create another dataframe which contains diversity information
#Calculate Diversity indices using the Vegan Package in R
```{r}
plant_sp_div = t(plant_full_data %>%
  tabyl(taxonID, siteID, sort = TRUE))

#Convert to Data frame
plant_sp_div = data.frame(plant_sp_div)


#Place specific epithet as row titles
names(plant_sp_div) <- as.matrix(plant_sp_div[1,])
plant_sp_div = plant_sp_div[-1,]
plant_sp_div[] <- lapply(plant_sp_div, function(x) type.convert(as.character((x))))

#Calculate Diversity Indices
simpson_diversity = data.frame(diversity(plant_sp_div, "simpson"))
simpson_diversity <-tibble::rownames_to_column(simpson_diversity, "siteID")

#eStar refers to the evenness, dStar = true diversity
evenness = data.frame(eventstar(plant_sp_div))
evenness<-tibble::rownames_to_column(evenness, "siteID")

#Shannon Diversity
shannon_diversity = data.frame(diversity(plant_sp_div, "shannon"))
shannon_diversity <-tibble::rownames_to_column(shannon_diversity, "siteID")

#Merge all 3 diversity tables

diversity_index = merge(shannon_diversity, simpson_diversity, by = "siteID")
```

diversity_index = merge(diversity_index, evenness, by = "siteID")

names(diversity_index)[1] = "siteID"
names(diversity_index)[2] = "shannon"
names(diversity_index)[3] = "simpson"
names(diversity_index)[5] = "evenness"
names(diversity_index)[6] = "true_div"

plant_field_data = merge(diversity_index, plant_field_data, by = "siteID")
write.csv(plant_field_data, "plant_field_data.csv", row.names = FALSE)

#merge richness and diversity indices
richness = data.frame(plant_occ[, -c(2:5226)])
plant_diversity_indices = merge(diversity_index, richness, by = "siteID")
write.csv(plant_diversity_indices, "plant_div_indices.csv")
```

---
title: "Beetle Diversity"
author: "Ayanna St. Rose"
date: "8/5/2021"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

#Download and load packages
```{r}
#install.pakages("tidyverse")
#install.packages("neonUtilities")
#install.packages("vegan")
#install.packages("vegan")
#install.packages("dplyr")
#install.packages("janitor")
#install.packages("ggExtra")
#install.packages("plotly")
#install.packages("rgdal")


library(readr)
library(ggplot2)
library(magrittr)
library(raster)
install.packages("plotly")

```
library(plotly)
library(ggExtra)
library(janitor)
library(plyr)
library(dplyr)
library(vegan)
library(tidyverse)
library(neonUtilities)
```

#Download data using neonUtilities
```{r, eval = FALSE, echo = FALSE}
beetle_data = loadByProduct(dpID = "DP1.10022.001",
                            site= c("BONA", "CLBJ", "DEJU", "GUAN", "HARV","KONZ", "LENO",
        "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
        "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
        "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
        "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
        "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
        "TECR", "TOMB", "TREE", "WALK", "WLOU"),
                            startdate = "2013-01", enddate = "2021-09",
                            check.size = FALSE)

#first part of the code was too large to run so I broke it up into smaller pieces
beetle_data1 = loadByProduct(dpID = "DP1.10022.001",
                            site= c("BONA", "CLBJ", "DEJU", "GUAN", "HARV","KONZ", "LENO",
        "MLBS","NIWO", "ONAQ"),
                            startdate = "2013-01", enddate = "2021-09",
                            check.size = FALSE)

beetle_data2 = loadByProduct(dpID = "DP1.10022.001",
                            site= c("ORNL", "OSBS", "SCBI",
        "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL"),
                            startdate = "2013-01", enddate = "2021-09",
                            check.size = FALSE)

beetle_data3 = loadByProduct(dpID = "DP1.10022.001",
                            site= c("ABBY","BART", "BIGC", "BLAN", "CUPE", "DELA", "FLNT",
        "GRSM", "GUIL", "HEAL"),
                            startdate = "2013-01", enddate = "2021-09",
                            check.size = FALSE)

beetle_data4 = loadByProduct(dpID = "DP1.10022.001",
                            site= c("HOPB", "JERC", "KING", "LECO", "LEWI", "LIRO", "MART",
        "MAYF", "MCRA", "MOAB", "POSE", "PRIN", "TEAK", "PUUM"),
                            startdate = "2013-01", enddate = "2021-09",


```

```
                    check.size = FALSE)

beetle_data5 = loadByProduct(dpID = "DP1.10022.001",
                    site= c("REDB","SERC", "RMNP","STEI", "SUGG", "TECR", "TOMB",
          "TREE", "WALK", "WLOU"),
                    startdate = "2013-01", enddate = "2021-09",
                    check.size = FALSE)
```

#I think GUAN is an incomplete site, we could remove it
#Extract expert taxonomic ID from data list
#Drop UID because it will be unique for repeated data
```{r}
beetle_taxa1 = beetle_data1$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa1 = subset(beetle_taxa1, select = -c(uid))

beetle_fieldData1 = beetle_data1$bet_fielddata

#Drop UID
beetle_fieldData1 = subset(beetle_fieldData1, select = -c(uid))

#number 2
beetle_taxa2 = beetle_data2$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa2 = subset(beetle_taxa2, select = -c(uid))

beetle_fieldData2 = beetle_data2$bet_fielddata

#Drop UID
beetle_fieldData2 = subset(beetle_fieldData2, select = -c(uid))

#number 3
beetle_taxa3 = beetle_data3$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa3 = subset(beetle_taxa3, select = -c(uid))

beetle_fieldData3 = beetle_data3$bet_fielddata

#Drop UID
beetle_fieldData3 = subset(beetle_fieldData3, select = -c(uid))
```

```
#number 4
beetle_taxa4 = beetle_data4$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa4 = subset(beetle_taxa4, select = -c(uid))

beetle_fieldData4 = beetle_data4$bet_fielddata

#Drop UID
beetle_fieldData4 = subset(beetle_fieldData4, select = -c(uid))

#number 5
beetle_taxa5 = beetle_data5$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa5 = subset(beetle_taxa5, select = -c(uid))

beetle_fieldData5 = beetle_data5$bet_fielddata

#Drop UID
beetle_fieldData5 = subset(beetle_fieldData5, select = -c(uid))
```


#Merge taxonomic and field data by namedLocation
```{r}
beetle_taxa = rbind(beetle_taxa1, beetle_taxa2, beetle_taxa3, beetle_taxa4, beetle_taxa5)
beetle_fieldData = rbind(beetle_fieldData1, beetle_fieldData2, beetle_fieldData3,
        beetle_fieldData4, beetle_fieldData5)

beetle_2018_full_data = merge(beetle_taxa, beetle_fieldData,
              by = c("namedLocation", "siteID", "domainID", "plotID", "setDate", "collectDate") ,
              all = TRUE, check.duplicates = FALSE)

#Remove rows where siteID and taxon ID says NA
#This will have the same contents as the beetle taxa dataframe,
#with added columns from field table
beetle_2018_full_data = beetle_2018_full_data %>% drop_na(taxonID)
beetle_2018_full_data = beetle_2018_full_data %>% drop_na(siteID)

#remove unknown species
beetle_2018_full_data = beetle_2018_full_data
beetle_2018_full_data<-
        beetle_2018_full_data[!(beetle_2018_full_data$specificEpithet=="sp."),]
```

#Write beetle total diversity that includes taxonomic information and field information
#Upload to box
```{r}
write.csv(beetle_2018_full_data, "beetle_full_data.csv", row.names = FALSE)
```

#There isn't a count for beetles, more so a presence and absence data
###Goal:
#####Run species richness for each site
#####Create data frame with different species for each site
#####Pseudoabundance is how many times the species is repeated (run for each site)
#####Count number of times a unique species occurs in each site

```{r}
#Beetle occurrence only sorted by siteID
#Quantify beetle occurrence
#Group by specific epithet and site ID

#remove unknown species

beetle_occ = beetle_2018_full_data %>%
  tabyl(specificEpithet, siteID, sort = TRUE)

#Show species distribution for each species for each site
beetle_occ = data.frame(t(beetle_occ))

names(beetle_occ) <- as.matrix(beetle_occ[1,])
beetle_occ = beetle_occ[-1,]
beetle_occ[] <- lapply(beetle_occ, function(x) type.convert(as.character((x))))

#The last column was empty so I removed it
beetle_occ = beetle_occ[,-447]

#Count the number of non-zero cells for each row
#This will give you the richness per site
#This will appear in the final column of the dataset
beetle_occ$count <- rowSums(beetle_occ!=0)

#Rename count as species richness
names(beetle_occ)[447] = "species_richness"


#Calculate abundance
#Sum all rows except species_richness column
row_sum = data.frame(rowSums(beetle_occ[, -447]))
```

```
#rename column
names(row_sum)[1] = "total_organism"

#rename 0th column
beetle_occ<-tibble::rownames_to_column(beetle_occ, "siteID")

#rename 0th column
row_sum<-tibble::rownames_to_column(row_sum, "siteID")

#merge row_sum dataframe with beetle_occ dataframe by
beetle_occ = merge(beetle_occ, row_sum,
                   by = "siteID",
                   all = TRUE, check.duplicates = FALSE)

#This table contains occurrence and richness values (last 2 columns)
```


#Write Occurrence data to CSV file.
```{r}
write.csv(beetle_occ, "beetle_occ.csv", row.names = FALSE)
```

#Import csv of site data from: https://www.neonscience.org/field-sites/explore-field-sites
```{r}
site_info = NEON_Field_Site_Metadata_20210226_0

# Select interest sites
site_info = site_info[site_info$field_site_id %in% c("BONA", "CLBJ", "DEJU", "GUAN",
          "HARV","KONZ", "LENO", "MLBS","NIWO", "ONAQ","ORNL", "OSBS", "SCBI",
          "SJER","SOAP","TALL","UKFS","UNDE","WREF","YELL", "ABBY","BART",
          "BIGC", "BLAN", "CUPE", "DELA", "FLNT", "GRSM", "GUIL", "HEAL", "HOPB",
          "JERC", "KING", "LECO", "LEWI", "LIRO", "MART", "MAYF", "MCRA", "MOAB",
          "POSE", "PRIN", "TEAK", "PUUM", "REDB","SERC", "RMNP","STEI", "SUGG",
          "TECR", "TOMB", "TREE", "WALK", "WLOU"), ]

#Merge info for interest site with data
beetle_field_data = merge(beetle_occ, site_info, by.x = "siteID", by.y = "field_site_id")

#Export previous file
write.csv(beetle_field_data, "beetle_field_data.csv", row.names = FALSE)


#Import table into GIS and Make layers in GIS with species richness and total
#number of species. This helps to visualize diversity.
#Still have to figure out of to graph it
```

```
#Visualize species richness
```{r}
p1 = ggplot(beetle_occ, aes(x = siteID, y = species_richness)) +
  geom_histogram(stat = "identity") +
  labs (title = "Histogram of Species Richness per Site", x= "site ID", y = "Number of Species") +
     theme_bw() +
      theme(panel.grid = element_blank(),
            axis.text = element_text(size = 12),
            axis.text.x = element_text(angle = 45, hjust = 1),
            axis.title = element_text(size = 12),
            plot.title = element_text(size = 14, hjust = 0.5, face = "bold"))

plot(p1)
```


#Create another dataframe which contains diversity information
#Calculate Diversity indices using the Vegan Package in R
```{r}
beetle_sp_div = t(beetle_2018_full_data %>%
  tabyl(specificEpithet, siteID, sort = TRUE))

#Convert to Data frame
beetle_sp_div = data.frame(beetle_sp_div)


#Place specific epithet as row titles
names(beetle_sp_div) <- as.matrix(beetle_sp_div[1,])
beetle_sp_div = beetle_sp_div[-1,]
beetle_sp_div[] <- lapply(beetle_sp_div, function(x) type.convert(as.character((x))))

#Calculate Diversity Indices
#Simpson Diversity
simpson_diversity = data.frame(diversity(beetle_sp_div, "simpson"))
simpson_diversity <-tibble::rownames_to_column(simpson_diversity, "siteID")

#eStar refers to the evenness, dStar = true diversity
evenness = data.frame(eventstar(beetle_sp_div))
evenness<-tibble::rownames_to_column(evenness, "siteID")

#Shannon Diversity
shannon_diversity = data.frame(diversity(beetle_sp_div, "shannon"))
shannon_diversity <-tibble::rownames_to_column(shannon_diversity, "siteID")
```
```

#Merge all 3 diversity tables

diversity_index = merge(shannon_diversity, simpson_diversity, by = "siteID")
diversity_index = merge(diversity_index, evenness, by = "siteID")

names(diversity_index)[1] = "siteID"
names(diversity_index)[2] = "shannon"
names(diversity_index)[3] = "simpson"
names(diversity_index)[5] = "evenness"
names(diversity_index)[6] = "true_div"

beetle_field_data = merge(diversity_index, beetle_field_data, by = "siteID")
write.csv(beetle_field_data, "beetle_field_data.csv", row.names = FALSE)

richness = data.frame(beetle_occ[, -c(2:447)])

beetle_diversity_indices = merge(diversity_index, richness, by = "siteID")
write.csv(beetle_diversity_indices, "beetle_div_indices.csv")
```

#Get species accumulation by curve adding sites in a random order
```{r}
accurve<-specaccum(beetle_sp_div, method="random", permutations=100)

plot(accurve,  ci.type="poly",col="blue", lwd=2, ci.lty=0, ci.col="lightblue")

plot(accurve$sites, accurve$richness,
    xlab="Number of Sites",
    ylab="Species Richness")

```

#Break up table by site so I can easily varify data is correct for each site
```{r}
#for(i in unique(beetle_2018_full_data$siteID)) {
 #       name = paste(i,"beetle_2018_full_data", sep = "_")
  #      assign(name, beetle_2018_full_data[beetle_2018_full_data$siteID==i,])
   #     str(name)
#}
```
---
title: "Total Species Diversity"
author: "Ayanna St. Rose"
date: "9/27/2021"
output: pdf_document

```

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

#Read CSV of diversity indices created from individual runs RMD
#rename columns
#merge
```{r}
#Remove the first column (X1)
beetle_div_indices = beetle_div_indices[,-1]
bird_div_indices = bird_div_indices[,-1]
plant_div_indices = plant_div_indices[,-1]

#Change column names
##Beetle
colnames(beetle_div_indices) = c("siteID", "beetle_shannon", "beetle_simpson",
                "beetle_qstar", "beetle_even", "beetle_tru_div", "beetle_dstar", "beetle_rich",
        "beetle_tot_org")

##Bird

colnames(bird_div_indices) = c("siteID", "bird_shannon", "bird_simpson",
                "bird_qstar", "bird_even", "bird_tru_div", "bird_dstar", "bird_rich", "bird_tot_org")

##Plant
colnames(plant_div_indices) = c("siteID", "plant_shannon", "plant_simpson",
                "plant_qstar", "plant_even", "plant_tru_div", "plant_dstar", "plant_rich",
        "plant_tot_org")
```

#Merge bird, plant and beetle diversity index data into total species div
```{r}
total_sp_div = merge(bird_div_indices, plant_div_indices, by = "siteID")
total_sp_div = merge(total_sp_div, beetle_div_indices, by = "siteID")

#Write total species richness to excel
write.csv(total_sp_div, "total_sp_div.csv", col.names = T)
```

#Missing 2 sites (DELA, ORNL), but this is the merged data
```{r}
forest_str_sp_div = merge(forest_str_div, total_sp_div, by = "siteID")

write.csv(forest_str_sp_div, "forest_str_sp_div.csv", row.names = T)
```

````
---
title: "Species Accumulation"
author: "Ayanna St. Rose"
date: "9/30/2021"
output: pdf_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```



#Load Libraries
```{r}
library(neonUtilities)
library(ggplot2)
library(vegan)
```

# Download beetle data for OSBS from 2013-2021
```{r, error = TRUE, echo=FALSE}
beetle_data = loadByProduct(dpID = "DP1.10022.001",
            site = "OSBS",
            startdate = "2013-01",
            enddate = "2021-09",
            check.size = FALSE)
```

#Extract Data tables we will need
```{r}
beetle_taxa = beetle_data$bet_expertTaxonomistIDProcessed

#Drop UID
beetle_taxa = subset(beetle_taxa, select = -c(uid))

beetle_fieldData = beetle_data$bet_fielddata

#Drop UID
beetle_fieldData = subset(beetle_fieldData, select = -c(uid))
```

#Merge taxonomic and field data by namedLocation, siteID etc to avoice duplicated columns
```{r}
beetle_2018_full_data = merge(beetle_taxa, beetle_fieldData,
                by = c("namedLocation", "siteID", "domainID", "plotID", "setDate", "collectDate") ,
                all = TRUE, check.duplicates = FALSE)
````
````

```
#Remove rows where taxon ID says NA
#This will have the same contents as the beetle taxa dataframe,
#with added columns from field table
library(tidyr)
beetle_2018_full_data = beetle_2018_full_data %>% drop_na(taxonID)

#remove unknown species
beetle_2018_full_data = beetle_2018_full_data
beetle_2018_full_data<-
        beetle_2018_full_data[!(beetle_2018_full_data$specificEpithet=="sp."),]

#Change collect dates to years
beetle_2018_full_data$collectDate =
        data.frame(format(as.Date(beetle_2018_full_data$collectDate, format = "%Y-%m-%d"),
        "%Y"))

#Rename column name
#Need to change between dataframe and matrix format so R can allow column name change
beetle_2018_full_data = as.matrix(beetle_2018_full_data)
beetle_2018_full_data = data.frame(beetle_2018_full_data)
names(beetle_2018_full_data)[6] = "year"
```

```
#Subset Beetle occurrence data
#only sorted by collect date
```{r, error=FALSE}
#Quantify beetle occurrence
#Using Tably to group, therefore need janitor package
install.packages("janitor")
library(janitor)

#Group by year
beetle_occ = beetle_2018_full_data %>%
  tabyl(year, specificEpithet, sort = TRUE, head = TRUE)


#Make 1st col, 0th column
rownames(beetle_occ) = beetle_occ[,1]

#Remove col 1 which will be a duplicate
beetle_occ = beetle_occ[,-1]

#Plot Species Accumulation
p1 = specaccum(beetle_occ)
```
```

```
plot(p1,
    ylab = "Species Richness Using Exact Method",
    xlab = "Years",
     main = "Beetle Species Accumulation Curve")

#Add confidence interval to plot
plot(p1, ci.type="poly", col="blue", lwd=2, ci.lty=0, ci.col="lightblue",
    xlab = "Years",
    ylab = "Beetle Species Richness",
    main = "Beetle Species Accumulation Curve")
```

Download plant data for OSBS from 2013-2021
```{r, error = TRUE, echo = FALSE}
plant_data = loadByProduct(dpID = "DP1.10058.001",
            site = "OSBS",
            startdate = "2013-01",
            enddate = "2021-09",
            check.size = FALSE)
```

#Extract Data tables we will need
```{r}
plant_taxa = plant_data$div_10m2Data100m2Data

#Drop UID
plant_taxa = subset(plant_taxa, select = -c(uid))

plant_taxa2 = plant_data$div_1m2Data

#Drop UID
plant_taxa2 = subset(plant_taxa2, select = -c(uid))
```

#Merge taxonomic and field data by namedLocation, siteID etc to avoice duplicated columns
```{r}
#Combine for a list of species at the 400m2 plot scale.
#For no duplicates, omit "by" command
plant_2018_full_data = merge(plant_taxa, plant_taxa2,
                all = TRUE, no.dups = TRUE)

#Remove rows where taxon ID says NA
plant_2018_full_data = plant_2018_full_data %>% drop_na(taxonID)

#Change collect dates to years
```

```
plant_2018_full_data$endDate = data.frame(format(as.Date(plant_2018_full_data$endDate,
          format = "%Y-%m-%d"), "%Y"))

#Rename column name
#Need to change between dataframe and matrix format so R can allow column name change
plant_2018_full_data = as.matrix(plant_2018_full_data)
plant_2018_full_data = data.frame(plant_2018_full_data)
names(plant_2018_full_data)[14] = "year"
```


#Subset plant occurrence data
#only sorted by collect date
```{r, error=FALSE}
#Quantify plant occurrence
#Using Tably to group, therefore need janitor package
#Group by year
plant_occ = plant_2018_full_data %>%
  tabyl(year, taxonID, sort = TRUE, head = TRUE)


#Make 1st col, 0th column
rownames(plant_occ) = plant_occ[,1]

#Remove col 1 which will be a duplicate
plant_occ = plant_occ[,-1]

#Plot Species Accumulation
p3 = specaccum(plant_occ)

plot(p3,
     ylab = "Species Richness Using Exact Method",
     xlab = "Years",
      main = "Plant Species Accumulation Curve")

#Add confidence interval to plot
plot(p3, ci.type="poly", col="blue", lwd=2, ci.lty=0, ci.col="lightblue",
     xlab = "Years",
     ylab = "Plant Species Richness",
      main = "Plant Species Accumulation Curve")

```


Download plant data for OSBS from 2013-2021
```{r, error = TRUE, echo=FALSE}
bird_data = loadByProduct(dpID = "DP1.10003.001",
```

```
            site = "OSBS",
            startdate = "2013-01",
            enddate = "2021-09",
            check.size = FALSE)
```

#Extract Data tables we will need
```{r}
bird_taxa = bird_data$brd_countdata

#Drop UID
bird_taxa = subset(bird_taxa, select = -c(uid))

bird_taxa2 = bird_data$brd_perpoint

#Drop UID
bird_taxa2 = subset(bird_taxa2, select = -c(uid))
```

#Merge taxonomic and field data by namedLocation, siteID etc to avoice duplicated columns
```{r}
#Combine for a list of species at the 400m2 plot scale.
#For no duplicates, omit "by" command
bird_2018_full_data = merge(bird_taxa, bird_taxa2,
                        by = c("siteID", "domainID", "plotID", "pointID", "eventID", "plotType",
          "startDate", "namedLocation","publicationDate"),
               all = TRUE, no.dups = TRUE)

#Remove rows where taxon ID says NA
bird_2018_full_data = bird_2018_full_data %>% drop_na(taxonID)

#Change collect dates to years
bird_2018_full_data$startDate = data.frame(format(as.Date(bird_2018_full_data$startDate,
          format = "%Y-%m-%d"), "%Y"))

#Rename column name
#Need to change between dataframe and matrix format so R can allow column name change
bird_2018_full_data = as.matrix(bird_2018_full_data)
bird_2018_full_data = data.frame(bird_2018_full_data)
names(bird_2018_full_data)[7] = "year"
```

#Subset bird occurrence data
#only sorted by collect date
```{r, error=FALSE}
```

```
#Quantify bird occurrence
#Using Tably to group, therefore need janitor package
#Group by year
bird_occ = bird_2018_full_data %>%
  tabyl(year, taxonID, sort = TRUE, head = TRUE)


#Make 1st col, 0th column
rownames(bird_occ) = bird_occ[,1]

#Remove col 1 which will be a duplicate
bird_occ = bird_occ[,-1]

#Plot Species Accumulation
p5 = specaccum(bird_occ)

plot(p5,
     ylab = "Species Richness Using Exact Method",
     xlab = "Years",
      main = "Bird Species Accumulation Curve")

#Add confidence interval to plot
plot(p5, ci.type="poly", col="blue", lwd=2, ci.lty=0, ci.col="lightblue",
     xlab = "Years",
     ylab = "Bird Species Richness",
      main = "Bird Species Accumulation Curve")

```
```{r}

plot(p3, ci.type="poly", col="black", lwd=2, ci.lty=0, ci.col="grey",
     xlab = "Years",
     cex = 0.02,
     ylab = "Plant Species Richness")

plot(p1, ci.type="poly", col="black", lwd=2, ci.lty=0, ci.col="grey",
     xlab = "Years",
     ylab = "Beetle Species Richness")
plot(p5, ci.type="poly", col="black", lwd=2, ci.lty=0, ci.col="grey",
     xlab = "Years",
     ylab = "Bird Species Richness")



```
```{r}
```

```
par(mfrow = c(1, 2), pty = "s")
plot(p5, ci.type = "poly", col="red", lwd=2, ci.lty=0, ci.col="grey") +
lines(p1, ci.type = "poly", col="blue", lwd=2, ci.lty=0, ci.col="grey")

plot(p3, ci.type="poly", col="black", lwd=2, ci.lty=0, ci.col="grey")
```
#Get species richness values and export to excel sheet
```{r}
plant_sp_rich = data.frame(p3$richness)
beetle_sp_rich = data.frame(p1$richness)
bird_sp_rich = data.frame(p5$richness)

write.csv(plant_sp_rich, "plant_sp_acc.csv")
write.csv(bird_sp_rich, "bird_sp_acc.csv")
write.csv(beetle_sp_rich, "beetle_sp_acc.csv")
```



```
---
title: "Lidar-CHM"
author: "Ayanna"
date: "9/15/2021"
output: html_document
---

```{r setup, include=FALSE, error = TRUE}
knitr::opts_chunk$set(echo = TRUE)
```

#Point cloud lidar = DP3.30024.001
#CHM Lidar = DP3.30015.001
#Elevation Lidar = DP3.30024.001
#No data for ORNL

#We will use DP3.30024.001
```{r, error=TRUE}
install.packages("neonUtilities")
library(neonUtilities)
library(raster)
library(gstat)
install.packages("lidR")
library(lidR)
install.packages("rgdal")
library(rgdal)
```
```{r}
```
```

```
#Create function to run CHM metrics
chm_metrics = function(chm) {
    mean.max.canopy.ht <- mean(chm@data@values, na.rm = TRUE)
    max.canopy.ht <- max(chm@data@values, na.rm=TRUE)
    rumple <- rumple_index(chm)
    top.rugosity <- sd(chm@data@values, na.rm = TRUE)
    cells <- length(chm@data@values)
    chm.0 <- chm
    chm.0[is.na(chm.0)] <- 0
    zeros <- which(chm.0@data@values == 0)
    deepgaps <- length(zeros)
    deepgap.fraction <- deepgaps/cells
    cover.fraction <- 1 - deepgap.fraction
    Zs <- chm@data@values
    Zs <- Zs[!is.na(Zs)]
    entro <- entropy(Zs, by = 1)
    gap_frac <- gap_fraction_profile(Zs, dz = 1, z0=3)
    GFP.AOP <- mean(gap_frac$gf)
    LADen<-LAD(Zs, dz = 1, k=0.5, z0=3)
    VAI.AOP <- sum(LADen$lad, na.rm=TRUE)
    VCI.AOP <- VCI(Zs, by = 1, zmax=100)
    out.plot <- data.frame(
      matrix(c(mean.max.canopy.ht, max.canopy.ht,
           rumple, top.rugosity, deepgaps, deepgap.fraction,
           cover.fraction,  entro, GFP.AOP, VAI.AOP, VCI.AOP),
         ncol = 11))
      colnames(out.plot) <-
    c("mean_max_canopy_ht",
      "max_canopy_ht", "rumple", "rugosity", "deepgaps","deepgap_fraction", "cover_fraction",
        'entropy',
      "gap_frac_per_cell",
      "veg_area_index", "vert_complexity")
  print(out.plot)
}
```

#Download Data and calculate canopy metrices

```{r, error = TRUE}
#Download data using neonutilities package
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "ABBY",
        easting = "552075",
        northing = "5067870",
        check.size = F)
```

```
#Rasterize the DTM and plot
abby_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D16/2021_ABBY_4/L3/DiscreteLid
        ar/DTMGtif/NEON_D16_ABBY_DP3_552000_5067000_DTM.tif")

plot(abby_dtm)

#Rasterize the DSM and plot
abby_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D16/2021_ABBY_4/L3/DiscreteLid
        ar/DSMGtif/NEON_D16_ABBY_DP3_552000_5067000_DSM.tif")

chm = abby_dsm - abby_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

abby_chm_metrics = chm_metrics(chm)
```

```{r, error=TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "BART",
        easting = "316812",
        northing = "4881511",
        check.size = F)

#Rasterize the DTM and plot
bart_dtm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D01/2019_BART_5/L3/DiscreteLid
        ar/DTMGtif/NEON_D01_BART_DP3_316000_4881000_DTM.tif")

plot(bart_dtm)

#Rasterize the DSM and plot
bart_dsm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D01/2019_BART_5/L3/DiscreteLid
        ar/DSMGtif/NEON_D01_BART_DP3_316000_4881000_DSM.tif")

chm = bart_dsm - bart_dtm
```

```
#View out the details
#Plot LAS file
summary(chm)
plot(chm)

bart_chm_metrics = chm_metrics(chm)
```

```{r, error=TRUE}
byTileAOP(dpID = "DP3.30024.001",
          year = "2021",
          site = "BLAN",
          easting = "753379",
          northing = "4327545",
          check.size = F)

#Rasterize the DTM and plot
blan_dtm =
           raster("DP3.30024.001/2021/FullSite/D02/2021_BLAN_4/L3/DiscreteLidar/DTMGtif/N
           EON_D02_BLAN_DP3_753000_4327000_DTM.tif")

plot(blan_dtm)

#Rasterize the DSM and plot
blan_dsm =
           raster("DP3.30024.001/2021/FullSite/D02/2021_BLAN_4/L3/DiscreteLidar/DSMGtif/N
           EON_D02_BLAN_DP3_753000_4327000_DSM.tif")

chm = blan_dsm - blan_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

blan_chm_metrics = chm_metrics(chm)
```

```{r, error=TRUE}
#does not work
#byTileAOP(dpID = "DP3.30024.001",
 #         year = "2021",
  #        site = "DELA",
   #       easting = "389258",
    #      northing = "3524154",
          check.size = F)
```

```

```{r, error= TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2018",
        site = "GRSM",
        easting = "273599",
        northing = "3952335",
        check.size = F)

#Rasterize the DTM and plot
grsm_dtm =
         raster("/home/rstudio/DP3.30024.001/2018/FullSite/D07/2018_GRSM_4/L3/DiscreteLid
         ar/DTMGtif/NEON_D07_GRSM_DP3_273000_3952000_DTM.tif")

plot(grsm_dtm)

#Rasterize the DSM and plot
grsm_dsm =
         raster("/home/rstudio/DP3.30024.001/2018/FullSite/D07/2018_GRSM_4/L3/DiscreteLid
         ar/DSMGtif/NEON_D07_GRSM_DP3_273000_3952000_DSM.tif")

chm = grsm_dsm - grsm_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

grsm_chm_metrics = chm_metrics(chm)
```



```{r, error=TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "HEAL",
        easting = "391276",
        northing = "7085047",
        check.size = F)

#Rasterize the DTM and plot
heal_dtm =
         raster("/home/rstudio/DP3.30024.001/2019/FullSite/D19/2019_HEAL_3/L3/DiscreteLid
         ar/DTMGtif/NEON_D19_HEAL_DP3_391000_7085000_DTM.tif")
```
```

```
plot(heal_dtm)

#Rasterize the DSM and plot
heal_dsm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D19/2019_HEAL_3/L3/DiscreteLid
        ar/DSMGtif/NEON_D19_HEAL_DP3_391000_7085000_DSM.tif")

chm = heal_dsm - heal_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

heal_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "JERC",
        easting = "741205",
        northing = "3453956",
        check.size = F)

#Rasterize the DTM and plot
jerc_dtm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D03/2019_JERC_5/L3/DiscreteLida
        r/DTMGtif/NEON_D03_JERC_DP3_741000_3453000_DTM.tif")

plot(jerc_dtm)

#Rasterize the DSM and plot
jerc_dsm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D03/2019_JERC_5/L3/DiscreteLida
        r/DSMGtif/NEON_D03_JERC_DP3_741000_3453000_DSM.tif")

chm = jerc_dsm - jerc_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

jerc_chm_metrics = chm_metrics(chm)
```

```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "MOAB",
        easting = "641031",
        northing = "4234596",
        check.size = F)

#Rasterize the DTM and plot
moab_dtm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D13/2021_MOAB_5/L3/DiscreteLid
         ar/DTMGtif/NEON_D13_MOAB_DP3_641000_4234000_DTM.tif")

plot(moab_dtm)

#Rasterize the DSM and plot
moab_dsm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D13/2021_MOAB_5/L3/DiscreteLid
         ar/DSMGtif/NEON_D13_MOAB_DP3_641000_4234000_DSM.tif")

chm = moab_dsm - moab_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

moab_chm_metrics = chm_metrics(chm)
```

```{r, error=TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "PUUM",
        easting = "256868",
        northing = "2163673",
        check.size = F)

#Rasterize the DTM and plot
puum_dtm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D20/2020_PUUM_2/L3/DiscreteLid
         ar/DTMGtif/NEON_D20_PUUM_DP3_256000_2163000_DTM.tif")

plot(puum_dtm)
```
```

```
#Rasterize the DSM and plot
puum_dsm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D20/2020_PUUM_2/L3/DiscreteLid
        ar/DSMGtif/NEON_D20_PUUM_DP3_256000_2163000_DSM.tif")

chm = puum_dsm - puum_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

puum_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2018",
        site = "RMNP",
        easting = "453588",
        northing = "4458524",
        check.size = F)

#Rasterize the DTM and plot
rmnp_dtm =
        raster("/home/rstudio/DP3.30024.001/2018/FullSite/D10/2018_RMNP_2/L3/DiscreteLid
        ar/DTMGtif/NEON_D10_RMNP_DP3_453000_4458000_DTM.tif")

plot(rmnp_dtm)

#Rasterize the DSM and plot
rmnp_dsm =
        raster("/home/rstudio/DP3.30024.001/2018/FullSite/D10/2018_RMNP_2/L3/DiscreteLid
        ar/DSMGtif/NEON_D10_RMNP_DP3_453000_4458000_DSM.tif")

chm = rmnp_dsm - rmnp_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

rmnp_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "SERC",
        easting = "364703",
        northing = "4305735",
        check.size = F)

#Rasterize the DTM and plot
serc_dtm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D02/2021_SERC_5/L3/DiscreteLida
         r/DTMGtif/NEON_D02_SERC_DP3_364000_4305000_DTM.tif")

plot(serc_dtm)

#Rasterize the DSM and plot
serc_dsm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D02/2021_SERC_5/L3/DiscreteLida
         r/DSMGtif/NEON_D02_SERC_DP3_364000_4305000_DSM.tif")

chm = serc_dsm - serc_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

serc_chm_metrics = chm_metrics(chm)
```

```{r}
#used 297000, 5042000
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "STEI",
        easting = "297968",
        northing = "5042743",
        check.size = F)

#Rasterize the DTM and plot
stei_dtm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_STEI_4/L3/DiscreteLidar
         /DTMGtif/NEON_D05_STEI_DP3_297000_5042000_DTM.tif")

plot(stei_dtm)
```

```
#Rasterize the DSM and plot
stei_dsm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_STEI_4/L3/DiscreteLidar
        /DSMGtif/NEON_D05_STEI_DP3_297000_5042000_DSM.tif")

chm = stei_dsm - stei_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

stei_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "TEAK",
        easting = "321515",
        northing = "4097400",
        check.size = F)
#Rasterize the DTM and plot
teak_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_TEAK_5/L3/DiscreteLid
        ar/DTMGtif/NEON_D17_TEAK_DP3_321000_4097000_DTM.tif")

plot(teak_dtm)

#Rasterize the DSM and plot
teak_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_TEAK_5/L3/DiscreteLid
        ar/DSMGtif/NEON_D17_TEAK_DP3_321000_4097000_DSM.tif")

chm = teak_dsm - teak_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

teak_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}


```

```
#Warning: TREE is a part of the flight boc of STEI site, downloaded data from STEI 297000,
        5041000
#NEON_D05_STEI_DP3_297000_5041000_DSM and DSM
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "TREE",
        easting = "297965",
        northing = "5041047",
        check.size = F)

#Rasterize the DTM and plot
tree_dtm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_STEI_4/L3/DiscreteLidar
         /DTMGtif/NEON_D05_STEI_DP3_297000_5041000_DTM.tif")

plot(tree_dtm)

#Rasterize the DSM and plot
tree_dsm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_STEI_4/L3/DiscreteLidar
         /DSMGtif/NEON_D05_STEI_DP3_297000_5041000_DSM.tif")

chm = tree_dsm - tree_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

tree_chm_metrics = chm_metrics(chm)
```

```{r}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "HARV",
        easting = "732183",
        northing = "4713265",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
harv_dtm =
         raster("/home/rstudio/DP3.30024.001/2019/FullSite/D01/2019_HARV_6/L3/DiscreteLid
         ar/DTMGtif/NEON_D01_HARV_DP3_732000_4713000_DTM.tif")
```

```
plot(harv_dtm)

#Rasterize the DSM and plot
harv_dsm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D01/2019_HARV_6/L3/DiscreteLid
        ar/DSMGtif/NEON_D01_HARV_DP3_732000_4713000_DSM.tif")

chm = harv_dsm - harv_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

harv_chm_metrics = chm_metrics(chm)
```

```{r}
#used 633000, 3696000
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "CLBJ",
        easting = "632982",
        northing = "3696682",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
clbj_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D11/2021_CLBJ_5/L3/DiscreteLida
        r/DTMGtif/NEON_D11_CLBJ_DP3_633000_3696000_DTM.tif")

plot(clbj_dtm)

#Rasterize the DSM and plot
clbj_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D11/2021_CLBJ_5/L3/DiscreteLida
        r/DSMGtif/NEON_D11_CLBJ_DP3_633000_3696000_DSM.tif")

chm = clbj_dsm - clbj_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)
```

```
clbj_chm_metrics = chm_metrics(chm)
```

```{r, error=TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "DEJU",
        easting = "561330",
        northing = "7084367",
        buffer = 20,
        check.size = F)
#Rasterize the DTM and plot
deju_dtm =
         raster("/home/rstudio/DP3.30024.001/2019/FullSite/D19/2019_DEJU_3/L3/DiscreteLida
         r/DTMGtif/NEON_D19_DEJU_DP3_561000_7084000_DTM.tif")

plot(deju_dtm)

#Rasterize the DSM and plot
deju_dsm =
         raster("/home/rstudio/DP3.30024.001/2019/FullSite/D19/2019_DEJU_3/L3/DiscreteLida
         r/DSMGtif/NEON_D19_DEJU_DP3_561000_7084000_DSM.tif")

chm = deju_dsm - deju_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

deju_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "KONZ",
        easting = "710729",
        northing = "4330786",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
konz_dtm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D06/2020_KONZ_6/L3/DiscreteLid
         ar/DTMGtif/NEON_D06_KONZ_DP3_710000_4330000_DTM.tif")
```

```
plot(konz_dtm)

#Rasterize the DSM and plot
konz_dsm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D06/2020_KONZ_6/L3/DiscreteLid
        ar/DSMGtif/NEON_D06_KONZ_DP3_710000_4330000_DSM.tif")

chm = konz_dsm - konz_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

konz_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "LENO",
        easting = "390139",
        northing = "3524827",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
leno_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D08/2021_LENO_6/L3/DiscreteLid
        ar/DTMGtif/NEON_D08_LENO_DP3_390000_3524000_DTM.tif")

plot(leno_dtm)

#Rasterize the DSM and plot
leno_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D08/2021_LENO_6/L3/DiscreteLid
        ar/DSMGtif/NEON_D08_LENO_DP3_390000_3524000_DSM.tif")

chm = leno_dsm - leno_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)
```

```
leno_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
 byTileAOP(dpID = "DP3.30024.001",
         year = "2021",
         site = "MLBS",
         easting = "542067",
         northing = "4136943",
         buffer = 20,
         check.size = F)

#Rasterize the DTM and plot
mlbs_dtm =
          raster("/home/rstudio/DP3.30024.001/2021/FullSite/D07/2021_MLBS_4/L3/DiscreteLid
          ar/DTMGtif/NEON_D07_MLBS_DP3_542000_4136000_DTM.tif")

plot(mlbs_dtm)

#Rasterize the DSM and plot
mlbs_dsm =
          raster("/home/rstudio/DP3.30024.001/2021/FullSite/D07/2021_MLBS_4/L3/DiscreteLid
          ar/DSMGtif/NEON_D07_MLBS_DP3_542000_4136000_DSM.tif")

chm = mlbs_dsm - mlbs_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

mlbs_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
         year = "2020",
         site = "NIWO",
         easting = "450328",
         northing = "4433940",
         buffer = 20,
         check.size = F)

#Rasterize the DTM and plot
niwo_dtm =
```

```
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D13/2020_NIWO_4/L3/DiscreteLid
        ar/DTMGtif/NEON_D13_NIWO_DP3_450000_4433000_DTM.tif")

plot(niwo_dtm)

#Rasterize the DSM and plot
niwo_dsm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D13/2020_NIWO_4/L3/DiscreteLid
        ar/DSMGtif/NEON_D13_NIWO_DP3_450000_4433000_DSM.tif")

chm = niwo_dsm - niwo_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

niwo_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2019",
        site = "OSBS",
        easting = "403886",
        northing = "3284767",
        buffer = 20,
        check.size = F)

 #Rasterize the DTM and plot
osbs_dtm =
         raster("/home/rstudio/DP3.30024.001/2019/FullSite/D03/2019_OSBS_5/L3/DiscreteLida
         r/DTMGtif/NEON_D03_OSBS_DP3_403000_3284000_DTM.tif")

plot(osbs_dtm)

#Rasterize the DSM and plot
osbs_dsm =
        raster("/home/rstudio/DP3.30024.001/2019/FullSite/D03/2019_OSBS_5/L3/DiscreteLida
        r/DSMGtif/NEON_D03_OSBS_DP3_403000_3284000_DSM.tif")

chm = osbs_dsm - osbs_dtm

#View out the details
#Plot LAS file
summary(chm)
```

```
plot(chm)

osbs_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "SJER",
        easting = "257213",
        northing = "4110433",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
sjer_dtm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_SJER_5/L3/DiscreteLidar
         /DTMGtif/NEON_D17_SJER_DP3_257000_4110000_DTM.tif")

plot(sjer_dtm)

#Rasterize the DSM and plot
sjer_dsm =
         raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_SJER_5/L3/DiscreteLidar
         /DSMGtif/NEON_D17_SJER_DP3_257000_4110000_DSM.tif")

chm = sjer_dsm - sjer_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

sjer_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "SOAP",
        easting = "298792",
        northing = "4100967",
        buffer = 20,
        check.size = F)
```

```
#Rasterize the DTM and plot
soap_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_SOAP_5/L3/DiscreteLida
        r/DTMGtif/NEON_D17_SOAP_DP3_298000_4100000_DTM.tif")

plot(soap_dtm)

#Rasterize the DSM and plot
soap_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D17/2021_SOAP_5/L3/DiscreteLida
        r/DSMGtif/NEON_D17_SOAP_DP3_298000_4100000_DSM.tif")

chm = soap_dsm - soap_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

soap_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "TALL",
        easting = "463241",
        northing = "3645863",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
tall_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D08/2021_TALL_6/L3/DiscreteLida
        r/DTMGtif/NEON_D08_TALL_DP3_463000_3645000_DTM.tif")

plot(tall_dtm)

#Rasterize the DSM and plot
tall_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D08/2021_TALL_6/L3/DiscreteLida
        r/DSMGtif/NEON_D08_TALL_DP3_463000_3645000_DSM.tif")

chm = tall_dsm - tall_dtm

#View out the details
```

```
#Plot LAS file
summary(chm)
plot(chm)

tall_chm_metrics = chm_metrics(chm)
```


```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "UKFS",
        easting = "310276",
        northing = "4323549",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
ukfs_dtm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D06/2020_UKFS_5/L3/DiscreteLida
         r/DTMGtif/NEON_D06_UKFS_DP3_310000_4323000_DTM.tif")

plot(ukfs_dtm)

#Rasterize the DSM and plot
ukfs_dsm =
         raster("/home/rstudio/DP3.30024.001/2020/FullSite/D06/2020_UKFS_5/L3/DiscreteLida
         r/DSMGtif/NEON_D06_UKFS_DP3_310000_4323000_DSM.tif")

chm = ukfs_dsm - ukfs_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

ukfs_chm_metrics = chm_metrics(chm)
```


```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "WREF",
        easting = "581417",
        northing = "5074636",
        buffer = 20,
```

```
        check.size = F)

#Rasterize the DTM and plot
wref_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D16/2021_WREF_4/L3/DiscreteLid
        ar/DTMGtif/NEON_D16_WREF_DP3_581000_5074000_DTM.tif")

plot(wref_dtm)

#Rasterize the DSM and plot
wref_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D16/2021_WREF_4/L3/DiscreteLid
        ar/DSMGtif/NEON_D16_WREF_DP3_581000_5074000_DSM.tif")

chm = wref_dsm - wref_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

wref_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "YELL",
        easting = "536352",
        northing = "4977885",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
yell_dtm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D12/2020_YELL_3/L3/DiscreteLida
        r/DTMGtif/NEON_D12_YELL_DP3_536000_4977000_DTM.tif")

plot(yell_dtm)

#Rasterize the DSM and plot
yell_dsm =
        raster("/home/rstudio/DP3.30024.001/2020/FullSite/D12/2020_YELL_3/L3/DiscreteLida
        r/DSMGtif/NEON_D12_YELL_DP3_536000_4977000_DSM.tif")

chm = yell_dsm - yell_dtm
```

```
#View out the details
#Plot LAS file
summary(chm)
plot(chm)

yell_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2020",
        site = "UNDE",
        easting = "304366",
        northing = "5123162",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
unde_dtm =
          raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_UNDE_4/L3/DiscreteLid
          ar/DTMGtif/NEON_D05_UNDE_DP3_304000_5123000_DTM.tif")

plot(unde_dtm)

#Rasterize the DSM and plot
unde_dsm =
          raster("/home/rstudio/DP3.30024.001/2020/FullSite/D05/2020_UNDE_4/L3/DiscreteLid
          ar/DSMGtif/NEON_D05_UNDE_DP3_304000_5123000_DSM.tif")

chm = unde_dsm - unde_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

unde_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2018",
        site = "GUAN",
        easting = "725706",
        northing = "1988112",
```

```
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
guan_dtm =
          raster("/home/rstudio/DP3.30024.001/2018/FullSite/D04/2018_GUAN_1/L3/DiscreteLid
          ar/DTMGtif/NEON_D04_GUAN_DP3_725000_1988000_DTM.tif")

plot(guan_dtm)

#Rasterize the DSM and plot
guan_dsm =
          raster("/home/rstudio/DP3.30024.001/2018/FullSite/D04/2018_GUAN_1/L3/DiscreteLid
          ar/DSMGtif/NEON_D04_GUAN_DP3_725000_1988000_DSM.tif")

chm = guan_dsm - guan_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

guan_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
        site = "ONAQ",
        easting = "376339",
        northing = "4448479",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
onaq_dtm =
          raster("/home/rstudio/DP3.30024.001/2021/FullSite/D15/2021_ONAQ_3/L3/DiscreteLid
          ar/DTMGtif/NEON_D15_ONAQ_DP3_376000_4448000_DTM.tif")

plot(onaq_dtm)

#Rasterize the DSM and plot
onaq_dsm =
          raster("/home/rstudio/DP3.30024.001/2021/FullSite/D15/2021_ONAQ_3/L3/DiscreteLid
          ar/DSMGtif/NEON_D15_ONAQ_DP3_376000_4448000_DSM.tif")
```

```
chm = onaq_dsm - onaq_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

onaq_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
#2019 data gave an error for rumple, resorted to 2018 data

byTileAOP(dpID = "DP3.30024.001",
        year = "2018",
        site = "BONA",
        easting = "476436",
        northing = "7225712",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
bona_dtm =
          raster("/home/rstudio/DP3.30024.001/2018/FullSite/D19/2018_BONA_2/L3/DiscreteLid
          ar/DTMGtif/NEON_D19_BONA_DP3_476000_7225000_DTM.tif")

plot(bona_dtm)

#Rasterize the DSM and plot
bona_dsm =
          raster("/home/rstudio/DP3.30024.001/2018/FullSite/D19/2018_BONA_2/L3/DiscreteLid
          ar/DSMGtif/NEON_D19_BONA_DP3_476000_7225000_DSM.tif")

chm = bona_dsm - bona_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

bona_chm_metrics = chm_metrics(chm)
```

```{r, error = TRUE}
byTileAOP(dpID = "DP3.30024.001",
        year = "2021",
```

```
        site = "SCBI",
        easting = "748090",
        northing = "4308784",
        buffer = 20,
        check.size = F)

#Rasterize the DTM and plot
scbi_dtm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D02/2021_SCBI_4/L3/DiscreteLidar
        /DTMGtif/NEON_D02_SCBI_DP3_748000_4308000_DTM.tif")

plot(scbi_dtm)

#Rasterize the DSM and plot
scbi_dsm =
        raster("/home/rstudio/DP3.30024.001/2021/FullSite/D02/2021_SCBI_4/L3/DiscreteLidar
        /DSMGtif/NEON_D02_SCBI_DP3_748000_4308000_DSM.tif")

chm = scbi_dsm - scbi_dtm

#View out the details
#Plot LAS file
summary(chm)
plot(chm)

scbi_chm_metrics = chm_metrics(chm)


#ORNL no data, DELA no data
```

```{r}
#Join tables
forest_str_div = rbind(abby_chm_metrics, bart_chm_metrics, blan_chm_metrics,
        bona_chm_metrics, clbj_chm_metrics, deju_chm_metrics, grsm_chm_metrics,
        guan_chm_metrics, harv_chm_metrics, heal_chm_metrics, jerc_chm_metrics,
        konz_chm_metrics, leno_chm_metrics, mlbs_chm_metrics, moab_chm_metrics,
        onaq_chm_metrics, osbs_chm_metrics, puum_chm_metrics, rmnp_chm_metrics,
        scbi_chm_metrics, serc_chm_metrics, sjer_chm_metrics, soap_chm_metrics,
        stei_chm_metrics, tall_chm_metrics, teak_chm_metrics, tree_chm_metrics,
        ukfs_chm_metrics, unde_chm_metrics, wref_chm_metrics, yell_chm_metrics)

row.names(forest_str_div) = c("ABBY", "BART", "BLAN" , "BONA", "CLBJ", "DEJU",
        "GRSM", "GUAN", "HARV", "HEAL", "JERC", "KONZ", "LENO", "MLBS",
```

```
        "MOAB", "ONAQ", "OSBS", "PUUM", "RMNP", "SCBI", "SERC", "SJER", "SOAP",
        "STEI", "TALL", "TEAK", "TREE", "UKFS", "UNDE", "WREF", "YELL")

write.csv(forest_str_div, "forest_str_div.csv", row.names = TRUE)
```

 # Statistical Analysis
 # Construct and Visualize Corrplot
```{r, error = T}

#Looking at Corrplot for structural metrics
sp_corr = data.frame(all_sp[,c(3:17)])

colnames(sp_corr)[1]<- "Ent"
colnames(sp_corr)[2]<- "VCI"
colnames(sp_corr)[3]<- "VAI"
colnames(sp_corr)[4]<- "MOCH"
colnames(sp_corr)[5]<- "H"
colnames(sp_corr)[6]<- "Rumple"
colnames(sp_corr)[7]<- "Rugosity"
colnames(sp_corr)[8]<- "CF"
colnames(sp_corr)[9]<- "DGF"
colnames(sp_corr)[10]<- "Bird Shannon"
colnames(sp_corr)[11]<- "Bird Richness"
colnames(sp_corr)[12]<- "Beetle Shannon"
colnames(sp_corr)[13]<- "Beetle Richness"
colnames(sp_corr)[14]<- "Plant Shannon"
colnames(sp_corr)[15]<- "Plant Richness"

corr_all = cor(sp_corr)
all_cor = corrplot(corr_all, method = 'square', order = 'FPC', type = 'lower', diag =
FALSE)
```