# Estimating Climate Resilience in Boreal Forests Through Functional Diversity

ESP Capstone – SP 2021

**Michael Reitz - Johns Hopkins** 

Mark White - The Nature Conservancy

# **Table of Contents**

Executive Summa	r <b>y</b>	••••	• • • • •	••••	• • • •	•••	• • • •	• • • •	•••	• • • •	•••	•••	•••	•••	• • • •	•••	••••	pg. 2
Introduction		••••	• • • • •	••••	••••	•••	• • • •	••••	•••	• • • •	•••	•••	•••	•••	••••	•••		pg. 3
Literature Review	•••••		••••	••••		•••	••••	••••	•••	• • • •	•••	•••	•••	• • • •	•••	•••	••••	. pg. 5
Methods		••••	••••	••••	••••	•••	• • • •	••••	•••	• • • •	•••	•••	•••	•••	••••	•••	••••	pg. 8
Results		••••	••••	• • • •	••••	• • • •	•••	•••	• • • •	•••	••••	•••	•••	••••	•••	•••	••••	pg. 13
Discussion	•••••		••••	••••	••••	••••	••••	•••	••••	••••	•••	•••	•••	•••	•••	•••	••••	.pg. 19
Conclusion		••••	• • • • •	••••	••••	••••	•••	• • • •	••••	•••	•••	•••	• • •	• • • •	•••		••••	pg. 22
References															•••			pg. 23

#### **Executive Summary**

The largest threat to northern forests ecosystems is the poleward migration of southern species, which are outcompeting and displacing the native vegetation. To mitigate this shift, land managers must optimize the growth rate of their forests and identify high priority areas for management or conservation. This study addressed these needs by analyzing which factors correlated to functional diversity in northeast Minnesota, which has been shown to increase overall productivity in southern boreal forests. We hypothesized that climatic and biological factors would have a statistically significant influence on functional diversity.

Four sites were selected across a range of forest and landscape types within the southern boreal region. Averaged trait measurements by species were put into novel functional diversity indices to estimate ecological function. Data collected from climate models, site visits, and LiDAR were put into models to explain variation of functional diversity.

We found that our hypothesis was supported. Both biological and climate variables were statistically significant in nearly all of the strongest models. We also noted that scope of our models significantly impacted their explanatory power. Models run with all plots together tended to be the weakest, models divided by broad plant community classifications were slightly stronger, and models grouped by a specific site or plant community both tended to be the strongest. These targeted models also leaned more heavily either towards biological or climate models, but did not completely abandon either category. Our results suggest that both site and native plant community play an important role in how tree species relate to one another. Within the scope of northeast Minnesota, these models will begin guiding conservation management and functional diversity research.

### Introduction

Boreal forests are under threat of displacement through competition as warming climates allow plants to migrate north. This is evident in northern Minnesota, along the southern extent of boreal forest habitat, where mild winters have allowed southern deciduous species to thrive in habitats they could not previously survive. Therefore, the Nature Conservancy in Duluth, Minnesota has started a "Conifer Strongholds" research program to understand where and how conservation money can be spent to maximize efficiency in combating the effects of climate change.

# Figure 1. Site Locations



Against invading species that are predisposed to warmer temperatures and longer growing seasons, native vegetation resilience will equate to the annual incremental growth rate. Studies have shown that overall forest productivity intuitively correlates to climatic factors plants generally rely on; photoperiod, precipitation, temperature, and soil fertility all being broadly important. Within temperate forests, for example, the next most important factor in predicting productivity is the number of individuals in a given area, but this isn't true everywhere. In nutrient-poor areas like boreal forests, population density is limited by resource availability. Therefore, a boreal forest community with high diversity will more efficiently divide limited resources, suggesting that they will have higher overall productivity and higher resiliency to encroachment (Paquette & Messier 2011).

This study will utilize data collected from four research natural areas (RNA's) within Superior National Forest in northern Minnesota to see what factors contribute to a boreal forest's diversity in our study area (**Figure 1**). Our study sites, Cabin Creek, Blueberry Lake, Southwest Greenwood Creek, and Keeley Creek, represent relatively undisturbed forests with unique species pools, geology, climates, etc. that still fall under the southern boreal mesoregion. In each RNA, randomized plots were established and measured in the summer of 2018. Plots will be periodically re-measured to track productivity and assess the models set forth in this study.

We will create novel functional diversity indices, using available trait data, to serve as a measure of the communities' ability to partition resources and resist climate change. We will then build models consisting of climatic factors and biological factors, independently and combined, to select a subset of variables shown to account for the variability in our functional diversity indices. Further models will combine the selected variables into quantifiable, predictive equations. Plots will be subdivided by plot, native plant community (NPC), and a broader NPC class to estimate the isolated role of each. We predict that both climatic and biological factors will have significant influence over the functional diversity in these forests.

Across our four sites, we also estimated their relative climate resilience using our diversity indices and variables shown to increase resiliency in other research. We compared each site's functional diversity rank to its resilience rank to estimate the relative influence of functional diversity to site-specific factors.

The end goal is to use this research to help guide management decisions: increasing functional diversity and, by extension, resilience in boreal forests that are at high risk from climate change. The models we intend to put forth will use data that can be collected from readily available databases like the Forest Inventory Analysis (FIA), National Map Digital Elevation Models (DEM), and species trait databases that will allow land managers to prioritize research, funding, and conservation effort. Furthermore, our plots in the RNA's will continue to be studied into the future, giving us a better understanding to what degree diversity improves climate change resilience.

#### **Literature Review**

Fei et al. (2017) demonstrate how tree species ranges have shifted as a response to climatic changes in precipitation and temperature. In transitional areas between ecoregions, species makeup can shift dramatically based on species-specific migration patterns. Specifically, 15 of 21 gymnosperm species typically found in boreal or sub-boreal northern forests have already displayed significant poleward migrations. Angiosperms, not typically associated with boreal habitats, showed a less focused migration path, but generally replaced gymnosperm-dominated habitats at the southern extent of their range.

Paquette et al. (2011) demonstrate the effect of biodiversity and climate on overall forest productivity across a temperature gradient. Both climate and diversity have larger roles in a boreal forest's productivity when compared to a temperate forest because harsher climate, scarcer resources, and forest dynamics through stand-replacing disturbances emphasizes the importance of beneficial, complimentary interspecies interaction. To account for different methods of measuring diversity, Tobner et al. (2014) utilized the International Diversity Experiment Network (IDENT) to study forests along diversity gradients. They compare functional diversity, measuring species trait data to serve as a proxy for plant history strategies, to traditional species diversity measures (Adler et al. 2014). Species richness and functional diversity will have a linear relationship as long as new species add entirely unique functions to the ecosystem. However, in actual forests, species will have functional redundancies, which are counter to positive, complimentary effects (Tobner et al. 2014; Paquette et al. 2011).

Therefore, to best estimate diversity's role in a boreal forest's productivity, complementarity, and by extension, climate resilience, the authors of Paquette et al. (2011) and Tobner et al. (2014) suggest either phylogenetic diversity (measuring the degree of relatedness between species on an evolutionary tree) or functional diversity.

Paquette et al. (2015) compare and contrast the influence of phylogenetic diversity and functional diversity on forest productivity. The two measures together, without any environmental factors included, explained 40% of total variance. Half of the explained variance (20%) was shared by both diversities. Of the remaining explained variation, 18% was solely explained by functional diversity and 2% was solely explained by phylogenetic diversity. Combining all measures of diversity and environmental factors could explain up to 72% of variance in forest productivity. Paquette et al. (2011) suggest that functional diversity is a better predictor of boreal forest productivity because it identifies functional redundancy, which is indicative of competitive exclusion and maladaptive for boreal ecosystems.

Both measures have their faults: phylogenetic diversity is sensitive to convergent evolution among distantly related species, divergent evolution among closely related species, and inconsistent rates of evolution. Functional diversity best explained variance in forest productivity, but has practical limitations

(Paquette et al. 2015). A massive amount of trait data is required for each species present and such data can be scant, especially for below-ground traits. Paquette et al. (2011) found that seed mass, wood density, and maximum height explained the most variation in productivity for temperate and boreal forests of Quebec province. Wright et al. (2006) found similar results, but also included specific leaf area and leaf size for their tropical research program. Although some traits are commonly included in functional diversity indices, a meta-analysis concludes that the traits used in functional diversity metrics must account for community- and site-specific functional space, niche realization, and trait-function relationships (Villeger et al. 2008). Redundant traits must be removed and the remaining traits should be weighted equally without a priori knowledge of relative importance.

As with other measures of diversity, functional diversity indices are an imperfect trade-off between richness, evenness, and dispersion. For purposes of this research, functional richness will be defined as the "amount of functional space filled by the community"; functional evenness is the evenness of "abundance distribution in a functional trait space"; and functional dispersion is how "abundance is spread along a functional trait axis, within the range occupied by the community" (Villeger et al. 2008). There are many ways to estimate diversity, but Shannon's and Gini-Simpson's Diversity indices are widely accepted in ecological research and can be applied to traits to estimate functional diversity (Jost 2006).

Amatangelo et al. (2014) have shown that, although vegetation community assemblages have changed in make-up and diversity, trait richness and diversity continue to be strongly correlated with environmental factors despite increasing pressure from climate change. The ecosystems that had a decline in trait-environment relationships tended to be in temperate forests with high fragmentation. Therefore, the authors conclude that remote northern forests with minimal disturbance are more likely to display climate resilience if trends continue. Additionally, sites that retain cooler winter microclimates

will be more likely to kill invading pioneers. Cooler summers with high levels of precipitation will put native vegetation under less stress, also increasing resilience.

## Methods

As part of the Nature Conservancy's 2018 Conifer Strongholds study program, 52 plots were randomly placed within 4 research natural areas (RNA's) in northeast Minnesota. These sites represent a diverse set of plant communities and environmental factors within the southern edge of the boreal range. All independent variables, collected either through fieldwork or GIS, were gathered from these newly established plots..

At each point, tree diameter and species were recorded within a 10 basal area factor (BAF) variable area plot. Plots with no recorded trees were disregarded. Herbaceous plant species were also recorded within a 1.128m fixed radius plot (1/1000<sup>th</sup> of an acre) to categorize each plot within a native plant community (NPC) using the Minnesota Department of Natural Resources field guide (Minnesota DNR 2003). Landscape variables such as surface soil texture, worm damage, slope, aspect, and topographic position were recorded for each plot. Aspect and slope data collected from each data point and input into a topographic moisture index (TMI) developed by Dougherty & Vankat (1983) to estimate relative soil moisture for each plot. An adjusted TMI was also included to account for thin soils, particularly in Keeley Creek RNA, where trees primarily grow on exposed bedrock or extremely young soil.

Temperature and precipitation data were gathered from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) Database (PRISM Data Group 2017) across several intervals to estimate seasonal temperature ranges and overall climate. Evapotranspiration, a product of temperature, precipitation, and plant vigor, was also included as a climate statistic (Velpuri et al. 2013).

Light Detection and Ranging (LiDAR) Digital Elevation Model (DEM) data was gathered from the US Geological Survey's publically available National Map Database (USGS 2021) to estimate the canopy complexity, landscape diversity, and patchiness of each of each site. Height values from point cloud data were used to calculate canopy coverages for heights greater than 2 meters and 6 meters, as well as total canopy height. Rumple, a measurement of canopy texture, was calculated by comparing outer canopy surface area to ground level surface area.

Surface soil texture was keyed out for the top mineral layer at each plot and supplemented by the Natural Resources Conservation Service (NRCS) Soil Survey Geographic Database (SSURGO). The Superior National Forest is unmapped, so field observations were compared to nearby SSURGO data points with similar soil texture to estimate root zone depth, organic matter, and available water supply for each plot.

Species trait data, the dependent variables, were sourced from the Paquette and Messier (2011) FDis database that averaged measurements of individuals growing in northeast Minnesota. Traits selected from the database included those that were shown to correlate with productivity in previous studies, such as maximum height, seed mass, and wood density, and traits with weaker or less understood correlations, like leaf mass, leaf nitrogen content, leaf longevity, and growth rate (Paquette and Messier 2011; Wright et al. 2006). Species midrange, while not a direct functional trait, was included as a measure of geographic dispersion. Species range midpoint diversity and the seven trait diversities were calculated for each plot using Shannon's diversity index and Gini-Simpson's index and converted from entropies to diversities using Jost's (2006) formulas.

Two principal component analyses (PCA) were performed to select traits novel functional diversity indices for Shannon's and Gini-Simpson's diversities. Diversities showing redundancies and lesser contributions to primary components were excluded from the functional diversity indices to prevent

overfitting. Each PCA split into distinct clusters, from which the strongest variable was selected (**Figures 2-3**). Species diversity, while not a part of functional diversity and not selected for the functional diversity indices, was included to compare against functional traits. In both PCA's, species diversity correlated well with average max height diversity, with only marginally stronger values.

In the Shannon's diversity PCA, traits split into five distinct clusters. From the five groups, the strongest were selected for the functional diversity index, including species range midpoint, leaf longevity, average maximum height, leaf nitrogen content, and seed mass (**Figure 2**). Leaf mass was cut in favor of leaf longevity, due to large overlap and weaker explanatory strength. Similarly, wood density was removed in favor of average maximum height, and growth rate was removed in favor of leaf nitrogen content.





The Gini-Simpson's PCA selected all of the same traits with the notable exception of seed mass, which was outperformed by leaf longevity (**Figure 3**). Clusters within the Gini-Simpson's PCA were otherwise similar in dispersion, with the exception of species range midpoint, which had a notable influence on a stronger component 2 (10.4%).

With no a priori knowledge of relative trait importance, traits were weighted equally in accordance with Villeger et al. (2008) and combined via averaging and summing to create four total functional diversity indices: Shannon's Sum (FDSS), Shannon's Mean (FDSM), Gini-Simpson's Sum (FDGSS), and Gini-Simpson's Mean (FDGSM).





For our independent variables, 35 measurements were taken and placed in a pairwise model with all 7 trait diversities, species midrange diversity, and the 4 functional diversity indices. Any variables that could not display significant correlation to any of the diversity metrics were thrown out. In the end, 27 independent variables showed some level of correlation to a trait diversity statistic in pairwise correlation tests, including basal area, average diameter at breast height (DBH), available water supply (at several depth ranges), soil organic matter content, root zone depth, TMI, adjusted TMI, evapotranspiration 2014-2018, evapotranspiration 2000-2018 (Velpurri et al. 2013), temperature minimum 2014-2018, temperature minimum 2000-2018, average summer temperature 2014-2018, average temperature 2014-2018, overall average temperature 2014-2018, average summer precipitation 2014-2018, average summer precipitation 2014-2018, average precipitation 2014-2018, average summer precipitation 2014-2018, average diversity, rumple, cover over 2 meters high, cover over 6 meters high, and canopy height over 6 meters high.

Similar to the trait variables, all of the independent variables were put into a PCA to identify redundancies and avoid overfitting. Compared to the trait diversities, the independent variables showed a much more scattered pattern across the PCA spectrum and the influence of each successive component had a significantly more gradual dropoff. With 27 variables, there was also much more overlap. No variables were dismissed at this point, but rather they were assessed on a case-by-case basis depending on which variables the models selected.

The independent variables were put into a stepwise model to filter out unnecessary components for each individual functional diversity model. Maximum K-folds were used to cross-validate the selected variables, which were put into a partial least squares model utilizing effect leveraging. Any variables in the model with P>0.05 were removed and the model was run again. The same process was run twice

more: once including only climate variables and once only using biological variables, to assess the individual strength of each component.

Models were initially run with all plots included, but subsequent models divided plots by site and NPC. The types of models used remained the same, however separate PCA's were run for each plot grouping. Variables within each plot grouping overlapped in unique ways, meaning that variables thrown out due to redundancy in some models would be considered in others. To account for this, each model was cross-referenced to it's particular PCAs.

# Results

The predictive power of each model varied greatly depending on the Functional Diversity Index, the set of independent variables included (climatic, biological, or all together), and the plot groupings (by site, by NPC, or all together). The purposefully diverse set of plots we chose yielded layered patterns, starting with the basic field-gathered data.

Table 1. Site Ba	ckground	l Informatic	)n
------------------	----------	--------------	----

Site	NPC Code(s)	NPC Description(s)	Basal Area (ft <sup>2</sup> /ac)	Average DBH (cm)	
Blueberry Lake	FDn43	Northern Mesic Mixed Forest	112.5	28.5	
Cabin Crook	FDn43	Northern Mesic Mixed Forest	0E E	25.4	
Cabin Creek	MHn45	Northern Mesic Hardwood (Cedar) Forest	0.0	23.4	
SW Groopwood Grook	MHn44	Northern Wet-Mesic Boreal Hardwood-Conifer Forest	109.2	20.2	
SW Greenwood Creek	MHn45	Northern Mesic Hardwood (Cedar) Fores	106.5	20.5	
Kaalaw Caala	50-22-2	Northern Poor Dry-Mesic Mixed Forest	50.0	26.2	
кеетеу стеек	FDn32c3	(Jack Pine - Black Spruce Subtype)	58.9	20.2	

The sites comprised of four total NPC's that under two broad categories: Fire Dependent (FD) and Mesic Hardwood (MH) communities within the northern region (**Table 1**). The defining characteristic splitting these two is the fire regime and the success of fire-evolved plant species. Therefore, fire dependent communities tend to have patchier canopies, sandier soils, and pioneering species. Mesic Hardwood communities are wetter and more stable, allowing for the forest to denser and older growth. As expected, we found our fire dependent sites to be sparser and younger than mesic hardwood sites, but this dichotomy doesn't necessarily indicate resilience or diversity.

Cito	Latituda	Longitudo	<b>Min Winter Temp</b>	Ave Summer Temp	<b>Ave Annual Precipitation</b>
Site	Latitude	Longitude	(Celsius)	(Celsius)	(cm)
Blueberry Lake	47.794	-90.478	-14.32	15.42	69.56
Cabin Creek	47.575	-91.106	-13.53	16.84	69.85
SW Greenwood Creek	47.466	-91.734	-14.37	16.47	61.11
Keeley Creek	47.772	-91.696	-12.65	16.83	59.24

#### Table 2. Site Climate Information

Resilience will be in part determined by the functional diversity of our sites, but each site's climate will have its own rate of change towards warmer and drier conditions. Sites that already have low minimum winter temperatures are more likely to reach the lethal temperature of potential invading species. Sites that have low summer temperatures and high annual precipitation will subject native species to less stress. Deep shade from a closed canopy will potentially prevent pioneering invaders from establishing themselves. Sites with high landscape diversity will also provide more functional space, increasing the diversity potential. Lastly, sites farther north from the temperate-boreal frontline transition zone will be buffered from seed dispersal until invaders are established within their proximity. These characteristics, along with functional diversity, were combined to create a relative resilience index for each of the four sites.

Blueberry Lake is by far the most resilient by our estimation, with an average resilience index of 4.46. It generally was the most diverse site by any measure of diversity, which may be owed to it's northernness, abundance of water resources, cooler temperatures, proximity to Lake Superior, or its resistance to warming temperatures due to its proximity to Lake Superior.

Cabin Creek, on the other hand, was the least resilient by a decent margin (2.63). The site is split into two distinct NPC's. The FDn43 area generally followed a dry ridgetop with a history of fires. The east-

facing slope was wetter and cooler than the ridgetop, allowing for older, more densely populated hardwoods to establish. Although each portion consisted of a unique species pool, the trait diversities remained somewhat similar between the two. Instead, soil properties and canopy structural differences resulting from fire history were the biggest split between the Cabin Creek communities.

The final two sites, Southwest Greenwood Creek and Keeley Creek, had very similar resiliency scores, averaging 3.01 and 3.08, respectively.

Southwest Greenwood Creek also comprises two separate native plant communities, however the difference is much more subtle than with Cabin Creek. MHn45 is slightly wetter than MHn44, noted by the presence of northern white cedar (*Thuja occidentalis*). As expected, this slight variance in species composition did not seem to affect trait diversity or any of the other variables collected. Southwest Greenwood Creek and Cabin Creek both had similar, intermediate diversity statistics, and both were consistently outperformed by Blueberry Lake and trailed by Keeley Creek. Greenwood Creek is considered more resilient than Cabin Creek because it's wetter climate and mature, dense canopy.

Keeley Creek is exceptionally nutrient poor, restricting the canopy to three tolerant species (Jack Pine [*Pinus banksiana*], Tamarack (*Larix laricina*), and Black Spruce [*Picea mariana*]). The fire history and undeveloped soils in this landscape create fragmented, patchy forests that are low in biological diversity, but high in landscape and canopy diversity. The far northern location and extreme soil conditions under which the native species have adapted have made this site resilient despite lower functional diversity.

The seven trait diversity means were analyzed by plot and NPC (**Figure 4**). Blueberry Creek generally was the most diverse and Keeley Creek was generally the least diverse. Cabin Creek and SW Greenwood Creek, which were both split between two NPC's, had similar, intermediate diversity values. Both Blueberry Lake and Cabin Creek contained FDn43 components, however, the diversities were notably higher in Blueberry Lake. The MHn44 and MHn45 components within SW Greenwood Creek are

functionally very similar, so the lack of separation between these two was expected, whereas the MHn44 and FDn43 portions within Cabin Creek showed significant differences. These results suggest that NPC and site may play a key role in how forests diversify.



## Figure 4. Trait Diversity Comparison

The four final functional diversity indices, comprising trait diversities selected by principal component analysis, were used to create 132 models (**Figure 5**). Independent variables were plotted against functional diversity indices, grouped together, with biological variables independently, and with climatic variables independently. Each bar in **Figure 5** represents the strongest model among the four models in each category.

#### Figure 5. Model Strength



With all plots included, models created from Shannon's diversities consistently explained about half of the variation in our functional diversity indices (R squared = 0.48), while Gini-Simpson's-based models had modest (R squared = 0.23) to negligible (R squared = 0.07) explanatory power. Models based solely on climatic factors nearly performed as well as models with all variables included (R squared = 0.46) and outperformed models including only biological factors significantly (R squared = 0.19). The strongest model incorporating all variable types selected one climate-based variable (available water supply [0-100cm]), and two biological-based variables (cover percentage above 6m height and rumple). Biological-only models just included cover percentages above 6m height, but climate-only models utilized evapotranspiration 2014-2018, adjusted TMI, and overall precipitation 2000-2018 in addition to available water supply (0-100cm).

Plot Grouping		All	By Site													
Subset		All		BI	ueberry L	.ake	0	abin Cre	ek	Gre	enwood Ci	Keely Creek				
Independent Variables	All Climate Bio A		All	Climate	Bio	All	Climate	Bio	All	Climate	Bio	All	Climate	Bio		
Diversity metric					66		~	ŝ		ŝ		<u> </u>	<u> </u>			
(Shannon's / Gini-Simpson's)	5	5	5	5	65	-	65	65	-	63	-	65	5	-	5	
Diversity metric	c/M	C/M	c/14				ç			•		c/M	c / M			
(Traits Sum / Traits Mean)	5/101	5/101	S/M	-	-	-	3	IVI	-	2	-	5/101	5/101	-	5	
# of Sites	47			15			11				12	9				

When plots were divided into subgroupings, our models were able to explain larger percentages of the variation in functional diversity. Using the same modelling methods, site-specific models had R squared values as high as 0.80 in Blueberry Lake, 0.79 in Cabin Creek, 0.69 in SW Greenwood Creek, and 0.74 in Keeley Creek. Although these models can predict functional diversity quite well, their predictive powers are confined to these particular sites or similar, neighboring areas. Additionally, the sample sizes for these models are substantially smaller in some cases, so Blueberry Lake (15 plots) is the most robust, followed by SW Greenwood Creek (12 plots), Cabin Creek (11 plots), and Keeley Creek (10 plots).

Table 4. Model Meta-data: Plots by NPC and NPC Class

Plot Grouping		By NPC													By NPC Class						
Subset		FDn32c3			FDn43			MHn44			MHn45			e Depend	ent	Mesic Hardwood					
Independent Variables	All	Climate	Bio	All	Climate	Bio	All	Climate	Bio	All	Climate	Bio	All	Climate	Bio	All	Climate	Bio			
Diversity metric										6											
(Shannon's / Gini-Simpson's)	5	5	-	GS	-	5	-	-	-	GS	-	65	5	5	5	5	-	5			
Diversity metric	c / 4	c/h4		_						c/14			c / 14	c/h4	c /h 4			c /			
(Traits Sum / Traits Mean)	S/IVI	S/M S/M	-	S	-	5	-	-	-	S/IVI	-	IVI	5/ M	S/IVI	S/M	IVI	-	S/IVI			
# of Sites	9			19			6			13			28			19					

We also divided our plots by NPC to understand the importance of location relative to community. Plots were again unevenly distributed. In this plot grouping, FDn43 (19 plots) was the most numerous, followed by MHn45 (13 plots), FDn32c3 (9 plots), and MHn44 (6 plots). All NPC categories, with the exception of MHn44, had models that explained a strong majority of the variability in the functional diversity indices. R values were recorded as high as 0.99 for MHn45, 0.85 for FDn43, and 0.74 for

FDn32c3. No models for the MHn44 subgroup passed the P<0.05 threshold to be accepted, potentially in part due to the low sample size. Although the model for MHn45 has a remarkably high R squared value, with a decent sample size across two sites and no PCA redundancies, we are hesitant to accept this number until further data is collected. Regardless, this may suggest the role of NPC is more important than originally thought.

To offset small sample sizes within certain NPC's, plots were also filtered into broader community categories: Fire Dependent (28 plots) and Mesic Hardwood (19 plots). Within these two groups, there is larger variation from plot to plot, but each is unmistakably unique due to the cascading effect of fire or lack thereof. Fire Dependent models explained an intermediate level of variation between the combined model and two subgroups. Notably, this was also the only other model group, beyond the combined model, in which climatic and biological variables could explain significant functional diversity variation in isolation.

In general, our results demonstrate that the broad approach of classifying the "southern boreal forest" as a cohesive unit is liable to overlook trends within subgroups. Model strength correlated well with specificity of plot groupings. Models with plots filtered by site or NPC type performed comparably well, followed by a broader NPC classification and models with no plot filtering.

In nearly all of the models they were included in, climatic and biological diversity variables contributed in some way, supporting our hypothesis. Models run with climatic or biological variables in isolation were generally weaker and inconsistent across NPC types and sites.

# Discussion

As we hypothesized, both climate data and biological data contributed to functional diversity, however, this relationship was more imbalanced and scale-dependent than we originally thought. Climate will

always be a driving factor in any forest's metrics, so the steadier, stronger relationship between climate and diversity is quite intuitive. Still, our climate data has macro-scale resolution, meaning that our data was essentially aggregated into four sites. To better understand the role of climate, future studies should include more sites with unique climate conditions and/or capture microclimate conditions for each plot.

The connection between biological data and diversity is also strong, but somewhat muddier. Fire dependent plant communities might be expected to have stronger canopy structure-diversity relationships, but these are all very dependent on fire severity, intensity, patchiness, and interval: statistics that would be of high priority in a more site- or NPC-specific study. Still, our limited approach saw that fire dependent NPC's correlated with LiDAR data more so than their mesic hardwood counterparts.

Other biological variables are loosely defined or poorly quantified. Our study also included basal area and average DBH, which are fundamental to any forestry study, but indicate stem density and size, not distribution. Beyond DBH, basal area, and LiDAR-derived canopy statistics, the "biological" set of variables starts to run into feasibility issues. Species distribution could be captured using GPS during ground surveys, but this would be very labor intensive and greatly reduce the applicability of these models if they were to be used as a desktop planning resource. Invasive species are also a major complication for diversity studies. Our sites were purposefully remote enough to avoid any invasive vegetation within our plots, but that doesn't mean invasive species didn't have an impact.

Earthworms are not native to the study area - a relic from the last glaciation. Recently invading worms have eaten the duff layer, removing winter insulation and harming shallow-rooted species in particular. This effect was potentially observed in the mesic hardwood portion of Cabin Creek, where sugar maple (*Acer saccharum*), a shallow-rooting species that evolved with developed duff layers, showed consistent

die-back on established canopy trees. The sugar maple individuals were old enough and the earthworm invasion was recent enough where the current canopy composition wasn't likely impacted significantly. However, this irreversible trend will almost certainly impact the fertility of the current canopy and the composition of the future canopy. Unfortunately, the science surrounding invasive earthworms in forests is somewhat juvenile. The tool for quantifying earthworm impact, the Invasive Earthworm Rapid Assessment Tool (IERAT), is coarsely discrete, designed for pure deciduous stands, and showed no discernable correlations in our pairwise testing (Alexander 2018).

White-tailed deer populations have also reached a high water mark since the turn of the century. Although hunting has created a downward trend recently, deer overpopulation continues to put pressure on seedling survival and species composition. Additionally, the Minnesota deer herd's preference toward northern white cedar (*Thuja occidentalis*) and oaks (*Quercus spp.*), along with their distaste for balsam fir (*Abies balsamea*) has selected for understories that can hardly resemble the mature canopies above them (Norton & Giudice 2017). Browsing pressure from white-tailed deer and other species (particularly snowshoe hare) can be hard to quantify and analyze, again limiting the applicability of such a study.

There are also further measurements we would have liked to include, but have not yet been able. Measuring the compartmentalization of rooting zones has been a key component for functional diversity studies in grasslands, but extends far beyond the current practicality of forest ecology (Clark et al. 2012).

Future models must consider scope early in the study design to capture all relevant details within reasonable limits. As biological variables become more available and climate variables become more precise, we believe the connection between these variables and functional diversity will be stronger than we observed in this study.

## Conclusion

It is important to note that the plots established for this research will continue to be re-measured in the future, for this ongoing study and others. The species composition and trait data discussed here will steer research toward blindspots or weaker or statistical relationships. Re-measurements will also help affirm or modify our understanding of diversity-productivity correlation in boreal forests that were set forth by previous research (Paquette & Messier 2011; Paquette et al. 2015; Zhang et al. 2012).

The data collected and analyzed in this study will begin guiding the Nature Conservancy's management decisions in a limited capacity. Resilient sites, such as Blueberry Lake, will be prioritized for conservation. Sites with lower functional diversity may be targeted for plantings to increase diversity, based on regeneration data. Topographic, soil, and canopy structure data can further pinpoint intra-site diversity strongholds.

The Nature Conservancy has already begun working with the US Forest Service to plant and monitor the growth of hundreds of thousands of seedlings planted across many acres of recently harvested forests within the southern boreal region. This project alone provides an ongoing opportunity to apply our best management principles established in this paper and observe resilience in real-time.

### References

Adler, P. B., Salguero-Gómez, R., Compagnoni, A., Hsu, J. S., Ray-Mukherjee, J., Mbeau-Ache, C., & Franco, M. (2014). Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences*, *111*(2), 740-745.

Alexander, G. (2018). The effects of invasive earthworms on understory plant community composition in sugar maple-basswood forests of the Chippewa National Forest, MN over two decades.

Allan, E., Weisser, W., Weigelt, A., Roscher, C., Fischer, M., & Hillebrand, H. (2011). More diverse plant communities have higher functioning over time due to turnover in complementary dominant species. *Proceedings of the National Academy of Sciences, 108*(41), 17034-17039.

Amatangelo, K. L., Johnson, S. E., Rogers, D. A., & Waller, D. M. (2014). Trait–environment relationships remain strong despite 50 years of trait compositional change in temperate forests. *Ecology*, *95*(7), 1780-1791.

Anderson, M. G., Clark, M., & Sheldon, A. O. (2014). Estimating climate resilience for conservation across geophysical settings. *Conservation Biology*, *28*(4), 959-970.

Clark, C. M., Flynn, D. F., Butterfield, B. J., & Reich, P. B. (2012). Testing the link between functional diversity and ecosystem functioning in a Minnesota grassland experiment. *PloS one*, 7(12), e52821.

Dougherty, K., & Vankat J.L., (1983). Soil Properties and Community Distributions on Serpentine Areas of Pennsylvania and Maryland. *Ohio Journal of Science* (Vol. 83, No. 2, pp. 9-9).

Duveneck, M. J., & Scheller, R. M. (2015). Climate-suitable planting as a strategy for maintaining forest productivity and functional diversity. *Ecological Applications*, *25*(6), 1653-1668.

Fei, S., Desprez, J. M., Potter, K. M., Jo, I., Knott, J. A., & Oswalt, C. M. (2017). Divergence of species responses to climate change. *Science advances*, *3*(5), e1603055.

Jost, L. (2006). Entropy and diversity. Oikos, 113(2), 363-375.

Kröber, W., Li, Y., Härdtle, W., Ma, K., Schmid, B., Schmidt, K., ... & Wirth, C. (2015). Early subtropical forest growth is driven by community mean trait values and functional diversity rather than the abiotic environment. *Ecology and Evolution*, *5*(17), 3541-3556.

Minnesota DNR. (2003). Field guide to the native plant communities of Minnesota: the Laurentian Mixed Forest Province. *Ecological Land Classification Program, Minnesota County Biological Survey, and Natural Heritage and Nongame Research Program. Minnesota Department of Natural Resources, St. Paul, Minnesota, USA*.

Norton, A., Giudice, J.H. (2017). Monitoring Population Trends of White-Tailed Deer in Minnesota - 2017. *Minnesota Department of Natural Resources, St. Paul, Minnesota, USA*.

Paquette, A., & Messier, C. (2011). The effect of biodiversity on tree productivity: from temperate to boreal forests. *Global Ecology and Biogeography*, *20*(1), 170-180.

Paquette, A., Joly, S., & Messier, C. (2015). Explaining forest productivity using tree functional traits and phylogenetic information: two sides of the same coin over evolutionary scale?. *Ecology and Evolution*, *5*(9), 1774-1783.

PRISM Climate Group. (2017). PRISM Climate Data. Oregon State University, Portland, Oregon.

#### http://prism.oregonstate.edu

Reich, P. B., Sendall, K. M., Rice, K., Rich, R. L., Stefanski, A., Hobbie, S. E., & Montgomery, R. A. (2015). Geographic range predicts photosynthetic and growth response to warming in co-occurring tree species. *Nature Climate Change*, *5*(2), 148-152.

Tobner, C. M., Paquette, A., Reich, P. B., Gravel, D., & Messier, C. (2014). Advancing biodiversity–ecosystem functioning science using high-density tree-based experiments over functional diversity gradients. *Oecologia*, *174*(3), 609-621.

USGS (2021). The National Map. *National Geospatial Program*. Retrieved from: <u>https://www.usgs.gov/core-</u> science-systems/national-geospatial-program/national-map Velpuri, N. M., G. B. Senay, R. K. Singh, S. Bohms, and J. P. Verdin. (2013). A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. Remote Sensing of Environment 139:35-49.

Villéger, S., Mason, N. W., & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89(8), 2290-2301.

Wright, S. J., Bunker, D., Dalling, J., Davies, S., Díaz, S., Engelbrecht, B., ... & Reich10, P. (2006). Towards a functional trait based research program within the Center for Tropical Forest Science. In *Unpublished report synthesizing the recommendations of a CTFS workshop on the priorities of a functional trait research program* (Vol. 15).

Zhang, Y., Chen, H. Y., & Reich, P. B. (2012). Forest productivity increases with evenness, species richness and trait variation: a global meta-analysis. *Journal of ecology*, *100*(3), 742-749.