ABSTRACT

Dahal, Bini. Spatial Distribution of Overstory and Understory Pine in Mixed Pine/Hardwood Forests of the Southern Appalachians. (Under the direction of Dr. Jodi Forrester).

Pine-oak forests of southern Appalachian Mountains have faced rapid decline due to fire suppression, southern pine beetle infestation, land use change, deforestation, and drought changing the forest composition from pine and oak dominated stands to closed-canopy stands dominated by shade-tolerant, fire-sensitive species. The lack of pine regeneration and establishment in the understory of the contemporary forest poses a major risk in the sustainability of these ecosystems. This has led to increased focus from both public and private landowners resulting in intensified restoration efforts that include prescribed burning and planting efforts. However, successful restoration of this ecosystem requires understanding the historical spatial arrangement of dry-site yellow pine to know if the current planting practice will restore the composition, but not the structure and function of these vulnerable yellow pine communities.

This dissertation aims to address these knowledge gaps by investigating the overstory and understory conditions of pine in mixed pine-hardwood forests of the southern Appalachians with a focus on DuPont State Recreational Forest, Cedar Mountain, NC. The study site was a mixed pine/hardwood forest consisting of pine (*Pinus* spp.), oaks (*Quercus* spp.), red maple (*Acer rubrum*), and yellow poplar (*Liriodendron tulipifera*). The study integrates aerial, terrestrial, and mobile LiDAR technologies with field based ecological assessment to classify pine species using machine learning algorithms, determine the spatial pattern of pine species, and analyze the environmental drivers influencing pine occurrence in the overstory and understory. While aerial LiDAR provides detailed canopy information, it lacks in capturing detailed understory information or in measuring important forest attribute such as diameter at breast height (DBH). Therefore, ground based LiDAR was used to supplement understory analysis.

Our results showed that convolutional neural network could classify pine and non-pine species with 85% accuracy using low-density aerial LiDAR data. Ground based laser scanning revealed that understory vegetation density significantly affected the tree detection ability of the LiDAR with mobile LiDAR outperforming terrestrial LiDAR in dense conditions. While understory density significantly affected the accuracy of DBH estimation by mobile LiDAR, such effect was not significant for the terrestrial LiDAR.

The derived spatial distribution of pine trees revealed a strong influence of environmental factors such as elevation, slope, aspect and topographic position index (TPI) in the pine distribution. Even after accounting for these environmental effects, clustering still existed at the smallest scale indicating tree-tree interactions. Temporal analysis of tree establishment and disturbance history highlights the decline of pine recruitment following mid-20th century. Additionally, while the overstory tree composition was influenced by the environmental variables, the understory plant composition was strongly influenced by canopy and midstory openness. The near absence of pine saplings and limited presence of pine seedlings further highlight regeneration challenges, likely driven by a dense understory dominated by mountain laurel and high stand density creating minimal canopy opening between 3.87% to 22.68%.

Yellow pine species exhibited a fine-scale clustering at shorter distance and a transition towards randomness at broader scales, indicating localized aggregation. The findings provide science-based recommendations for restoration and adaptive management of declining pine-oak ecosystems, emphasizing the need for disturbance regimes that promote pine regeneration and maintain ecosystem resilience in the southern Appalachians.

Spatial Distribution of Overstory and Understory Pine in Mixed Pine/Hardwood Forests of the Southern Appalachians

by
Bini Dahal

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Forestry and Environmental Resources

Raleigh, North Carolina
2025

APPROVED BY:

Dr. Jodi Forrester Committee Chair	Dr. Leah Rathbun
Dr. Tara Keyser	Dr. Brian Reich
Dr. Sathish Samiappan	

DEDICATION

To Bikram, for your unwavering support, endless patience, and belief in me.

To my parents, Kedar and Kamala, for your sacrifices, love, and the values you instilled in me that shaped who I am today.

To my brother Bibek and my sister-in-law Akreeti, for being my constant cheerleader.

To my grandmothers, Bishnu Lamichhane and Dirgha Maya Dahal, for your prayers, strength, and gentle encouragement from afar.

And finally, I dedicate this dissertation and PhD to my late grandfather, Krishna Lamichhane, whose wisdom and memory continue to inspire me every single day.

BIOGRAPHY

Bini Dahal was born and raised in Hetauda, Nepal a city nestled in the southern part of Bagmati Province. Growing up in a middle class, forest-dependent household, she gained an early understanding of how essential forest resources were to her family's livelihood. Although she initially considered a career in nursing, a conversation with a friend introduced her to the field of forestry, a path that immediately resonated with her values and aspirations. Bini started her academic journey in Forestry 2011 and graduated with an Associate degree from the Forestry program at Institute of Forestry, Tribhuvan University, Hetauda, in 2013. She went on to complete her B.Sc. in Forestry from Agriculture and Forestry University in 2018. Motivated to expand her knowledge and pursue research, she took a leap of faith and moved to the United States in 2019 to pursue a master's degree in Forestry from Mississippi State University. There her thesis focused on understanding the biomass production potential of short rotation woody crops under different treatment scenarios. In 2021, Bini began her Ph.D. in Forestry and Environmental Resources at North Carolina State, where she explored the pines of Southern Appalachians. During her doctoral studies, she also earned a master's degree in Statistics, further strengthening her analytical foundation.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Jodi Forrester, for taking me under her wings and guiding and supporting me over the past two and a half years. I am also deeply thankful to Dr. Leah Rathbun for her mentorship during my early years at NC State and for continuing to support my work as a member of my committee. Thank you to my other committee members – Dr. Brian Reich, Dr. Tara Keyser, and Dr. Sathish Samiappan for their support and valuable feedback throughout this journey. I am grateful to my present and former lab mates and friends, including Thomas Halting, Jacks Hausle, Casey Wofford, Ivan Raigosa Garcia, and Titilayo Tajudeen, for their invaluable assistance in the field and lab. This research would not have been possible without their contributions. I am also thankful to my dear friend, Gaurav Dhungel, whose friendship and support have meant a great deal to me. Finally, I extend my thanks to Mike and Jordan at Dupont State Recreational Forest for answering my questions and providing helpful information for this research.

TABLE OF CONTENTS

TABLE OF CONTENTS v
LIST OF TABLES viii
LIST OF FIGURES ix
CHAPTER 1 - Introduction: Pine-oak forests of southern Appalachian region 1
CHAPTER 2 - Leveraging deep learning and low-density aerial LiDAR to segment individual trees and distinguish pines in a mixed pine-hardwood forest 11
2.1 Abstract 11
2.2 Introduction 12
2.3 Methodology 15
2.3.1 Study Area 15
2.3.2 Field Data Collection 16
2.3.3 LiDAR Data Processing 17
2.3.4 Data Analysis 26
2.4 Results 27
2.4.1 Field trees matching with LiDAR data 27
2.4.2 2D CNN classification results 30
2.4.3 Classification for the Study Area 32
2.5 Discussion 33
2.5.1 Forest Complexity 34
2.5.2 Limitations and Uncertainties 35
2.6 Conclusion 36
CHAPTER 3 - Spatial patterns of pine in mixedwood forests of the southern Appalachians 48
3.1 Abstract 48
3.2 Introduction 49
3.3 Methodology 53
3.3.1 Study Area 53
3.3.2 Data Collection 54
3.3.3 Spatial Pattern Assessment 56
3.4 Results 60
3.4.1 Quadrat test 60
3.4.2 Inhomogeneous Poisson Process with Covariates 61
3.4.3 Spatial Pattern Assessment 67
3.5 Discussion 74

3.6 Conclusion 77
References 79
CHAPTER 4 - Establishment pattern, disturbance history, and factors affecting yellow pine regeneration patterns in southern pine-oak mixedwoods 87
4.1 Abstract 87
4.2 Introduction 88
4.3 Methods 91
4.3.1 Study Site 91
4.3.2 Field Measurements 92
4.3.3 Data Analysis 93
4.4 Results 95
4.4.1 Forest Structure and Composition 95
4.4.2 Environmental Gradients and Plant Communities 100
4.4.3 Tree Establishment Pattern and Growth Release Events Over Time 105
4.4.4 Decadal Differences in Species Growth Patterns 107
4.5 Discussion 108
4.6 Conclusion 111
References 112
CHAPTER 5 - The influence of understory vegetation density on the accuracy of modeling tree structure using terrestrial and handheld mobile laser scanning 119
5.1 Abstract 119
5.2 Introduction 119
5.3 Methods 124
5.3.1 Study Site 124
5.3.2 Field Methods 125
5.3.3 Data Analysis 126
5.4 Results 133
5.5 Discussion 137
5.6 Conclusion 140
References 142
CHAPTER 6 - Synthesis 151
APPENDICES 155
S.2 CHAPTER2 – SUPPLEMENTARY MATERIAL 156
S.3 CHAPTER3 – SUPPLEMENTARY MATERIAL 159
S.4 CHAPTER 4 – SUPPLEMENTARY MATERIAL 160

S.5 CHAPTER 5 – SUPPLEMENTARY MATERIAL 163

LIST OF TABLES

Table 2.1: Matched rate for comparison of trees classified from LiDAR with field measured individuals by canopy classes. The total number of trees per class is displayed with the relative percentage of all trees shown in parentheses. 28

Table 2.2: Matched rate for pine and non-pine species class. 29

Table 2.3: Comparison of test results from the RF, SVM, and 2D CNN models.....31

Table 3.1: Parameter estimates and statistical significance of environmental and spatial covariates in the IPP Model for pine distribution. Asterisks indicate significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. 63

Table 3.2: Parameter estimates and statistical significance of environmental and spatial covariates in the IPP Model for pine distribution by height class. Asterisks indicate significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. 65

Table 4.1: Tree species level summary of forest composition in a pine-oak mixedwoods stand (n=15) in DuPont State Recreational Forest, NC. 96

Table 4.2: Sapling (stems<10 cm DBH) and seedling density in a pine-oak forest in DuPont State Recreational Forest, NC. 98

Table 4.3: Tukey HSD result for each species. In the table, FD represents the first decade, SD is the second decade, PD is the penultimate decade, and LD is the last decade. Diff is the difference in growth between two decades in mm. 102

Table 5.1: Summary of the plots classified into open, medium and dense. 133

Table 5.2: Summary statistics for the DBH estimation by MLS and TLS under different understory density. 136

LIST OF FIGURES

Figure 2.1: Study area at DuPont State Recreational Forest. The blue outlined area is the study area of 26 ha. The red circle represents locations of the fixed area plots used for ground truth data. 16

Figure 2.2: LiDAR point density of the study area. 18

Figure 2.3: Workflow of the research methodology used up to the point of classification comparisons. 18

Figure 2.4: An example of one-meter horizontal sectioning of the tree crown point cloud. 20

Figure 2.5: Layer 12 showing clustering by K-means and DBSCAN. The rectangle box shows how K-means assigned the same cluster id (represented by the convex hull) to the points which was later separated into two clusters by DBSCAN clustering. Black points represent the tree locations. 22

Figure 2.6: Segmentation resulting from three different algorithms: a) our approach vs b) dalponte vs c) watershed. 24

Figure 2.7: Comparison of a) Complete Tree Data and b) Extracted Crown Data (Top 60% of Tree Height). 25

Figure 2.8: Comparison of diameter ranges between pine and non-pine groups across crown class. 29

Figure 2.9: Confusion matrix for the 2D CNN model used to classify 38 test trees into either pine or non-pine categories. 32

Figure 2.10: Map showing individual trees classified into pine and non-pine categories. 33

Figure 3.1: Study area at Dupont State Recreational Forest, NC. Red points represent individual pine trees including a mix of pitch pine, shortleaf pine, and Eastern white pine. 54

Figure 3.2: Environmental variables including elevation (meters), slope (degrees), aspect (degrees), and topographic position index used as covariates in the study to examine the influence of topographic heterogeneity on pine distribution in the southern Appalachian region. Red points overlaid in the elevation map represent pine trees. 56

Figure 3.3: Quadrat test for spatial heterogeneity of pine tree distribution. For example, in the top left quadrat, 366 pines were observed, while the expected counts based on the area of the quadrat was 169.8 and the Pearson residuals was 15. Pearson residual closer to 0 represents that the observed value is close to the expected value while a positive residual represents more observed points than expected, and a negative residual represents less observed points than expected. 61

Figure 3.4: Pine and non-pine species distribution across TPI by height class. Plot (a)-(c) show pine distribution for small, medium, and large height classes respectively, while plot (d) shows the combined distribution of non-pine species. 66

Figure 3.5: Comparison of homogeneous and inhomogeneous spatial point pattern analyses for pine distribution. Panels (a) and (b) show the homogeneous and inhomogeneous Ripley's L-functions, respectively, while panels (c) and (d) show the homogeneous and inhomogeneous pair correlation functions. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates clustering while below the envelope indicates uniform distribution. 68

Figure 3.6: Spatial distribution of pine trees by height class: a) short, b) intermediate, and c) tall using pair correlation function. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates clustering while below the envelope indicates uniform distribution. 69

Figure 3.6: Localized clusters of pine trees identified by the spatial scan statistic. a) All pines combined, b) short pine trees c) intermediate pine trees, and d) tall pine trees. Colors represent the magnitude of test statistic with yellow indicating stronger clustering intensity. 71

Figure 3.8: Bivariate spatial associations between pine height classes based on inhomogeneous pair correlation functions. Class 1 represent short pine trees, class 2 represent intermediate pine trees, and class 3 represent tall pine trees. The top panel shows the association between short and intermediate trees, the middle panel shows short and tall trees, and the bottom panel shows intermediate and tall trees. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates spatial attraction, while below the envelope indicates repulsion across spatial scales…73

Figure 3.9: Visualization of a) pine and b) non-pine trees distribution by height class. 74

Figure 4.1: Sapling density by height class in a pine-oak mixedwoods stand in DuPont State Recreational Forest. The blue point represents the mean sapling density (n=15). 100

Figure 4.2: Non-metric multidimensional scaling ordination of overstory tree communities based on species importance values. a) Axis 1 explained 51.5% and Axis 2 explained 32.7% of the variability; while b) Axis 3 explained 10.8%. 102

Figure 4.3: Non-metric multidimensional scaling ordination of understory regeneration displaying a) Axis 1 versus Axis 2 and b) Axis 2 and 3. 104

Figure 4.4: a) Number of trees established per decade in a mixed pine-oak stand at DuPont State Recreational Forest, NC. b) Percent of trees exhibiting release events by decade categorized as major or minor releases. A disturbance was considered stand wide if more than 25% of the trees exhibit growth releases. 106

Figure 4.5: Radial Growth (mm) by decade comparison between a) pine, b) oak and maple. Blue diamond shape point represents mean growth for each individual species in each decade..... 108

Figure 5.1: Visual representation of the three strata differentiated based on understory density. The top image is plot with open understory, middle image is representing medium understory, and the bottom image is representing dense understory. The images are extracted from Faro SCENE. 127

Figure 5.2: Original HMLS data (left) vs data after filtering with verticality and curvature (right). 129

Figure 5.3: DBSCAN clustering (left) vs filtered data using the structural criteria (right). 130

Figure 5.4: Detection rate by MLS and TLS under different understory density. Blue points represent mean detection rate. 134

Figure 5.5: DBH estimation error under different understory density for MLS and TLS. The mean DBH error per understory density is shown as blue point. 135

Figure 5.6: Relationship between field measured DBH and DBH estimation error across MLS and TLS. Each point represents an individual trees. The loess smoothing curve with shaded confidence interval shows the underlying trend in DBH estimation error relative to tree size by fitting a localized regressions across the data, allowing for visualization of non-linear patterns in error as tree size increases. 137

CHAPTER 1 - Introduction: Pine-oak forests of southern Appalachian region

1.1 Background

Pine-oak forest types establish during early to mid-successional stages following severe human-made, or lightning caused wildfires in the mountainous regions of the southern United States, like the Cumberland and Appalachian Mountains. These forest types are frequently found on xeric slopes and ridges, and at low and middle elevations of the southern Appalachian Mountains which receive more sunlight exposure (Clabo & Clatterbuck, 2020). In the past, the southeastern United States consisted of forest dominated by southern pine including shortleaf pine (*Pinus echinata*), pitch pine (*P. rigida*), Table Mountain pine (*P. pungens*), and Virginia pine (*P. virginiana*), which sustained a wide variety of plants and animals indigenous to this habitat such as golden winged warbler (*Vermivora chrysoptera*), mountain golden heather (*Hudsonia montana*), and red-cockaded woodpecker (*Picoides borealis*) (Lafon et al., 2010; Nordman et al., 2021). Along with high biodiversity, they offer a number of important ecosystem services, such as watershed protection, hunting and sustainable forest products. Despite their ecological and conservation significance, these open pine ecosystems are currently threatened by successional pressures from hardwoods, as a result of land use pattern change, deforestation, fire exclusion, and southern pine beetle infestation (Elliott et al., 2012; Oswalt et al., 2016; Waldron et al., 2007). Among these drivers, fire has had profound ecological consequences. In the absence of frequent fire, which historically maintained open canopy structures, stand density increased reducing the light availability in the understory (Lafon & Kutac, 2003). This suppressed pine regeneration, and promoted the dominance of shade tolerant hardwoods in the understory. Additionally, southern pine beetle infestation caused mortality of overstory pine, resulting in a rapid shift toward being dominated by less fire tolerant species ultimately altering

the forest species composition and structure (Guldin, 2019; Jenkins et al., 2011; Waldron et al., 2007). Eastern white pine (*P. strobus*) is also native in the region and can be a dominant species in mixed hardwood forests at higher elevations, particularly on north and west-facing slopes but also occur in more exposed, subxeric sites with oaks. While yellow pines have declined with the fire suppression, the more shade-tolerant and fire-sensitive Eastern white pine has increased in dominance.

Over the past few decades, management and restoration of fire-dependent xeric pine ecosystems have become a major focus for both private and federal land managers. Prescribed burning has been progressively used for the restoration of pine in southern Appalachian region (Haines et al., 2001; Randles et al., 2002). Both prescribed burning and wildfire have shown to decrease stand density, increase understory species richness, and alter vegetation composition by reducing mesophytic species which ultimately facilitates the establishment of pine in southern Appalachian region (Arthur et al., 1998; Welch et al., 2000). Research has shown successful regeneration of pine following fire to be dependent on substantial overstory and understory reduction. Specifically, seedling establishment was limited until overstory density was reduced by over 40%, understory density by over 80%, and post-burn shrub cover remained below 10% (Jenkins et al., 2011). However, the successful implementation of restoration and management strategies requires a clear and timely understanding of the spatial heterogeneity and arrangement of pine stands and its changes over time. Spatial patterns are crucial indicators of past disturbances regimes, competitive interactions and site conditions (Boyden et al., 2005; Després et al., 2017), and can therefore inform silvicultural and fire management strategies for promoting pine regeneration in southern Appalachian Region.

Traditional forest inventory techniques while robust are labor intensive, time consuming and limited in spatial extent. These constraints are especially evident in the mountainous areas such as the southern Appalachians, where the complexity of the terrain and thick understory vegetation create significant challenges for ground based monitoring. Additionally, research that focuses on understanding the distribution and status of pine in southern Appalachian region relies on sample plot data within individual forest sites (Goode et al., 2021). While these studies provide critical insights at the stand or site level, it may often obscure fine-scale variation and introduce bias in understanding the tree distribution in a heterogeneous landscapes like that of southern Appalachia. Therefore, there is a need for scalable approaches that can assess the pine distribution at broader spatial scales to better inform restoration planning and landscape-level management.

In this scenario, advancements in remote sensing technologies particularly Light Detecting and Ranging (LiDAR), offers revolutionary opportunities. LiDAR is an active remote sensing technology that can accurately capture and assess both horizontal and vertical forest structure over large areas. In forestry, LiDAR technology has been used since the early 1900s to determine tree height, forest biomass, timber volume, and vegetation canopy structure (Harding et al., 1994; Nelson et al., 1988). For instance, aerial LiDAR offers individual tree level analysis including tree species classification, characterization of canopy height, and spatial pattern analysis (Rudjord & Trier, 2016; Wiggins et al., 2019; Zhang & Qiu, 2012). Meanwhile ground based LiDAR systems such as terrestrial and mobile LiDAR offers fine-scale insights at the plot level, providing detailed measurements of tree including diameter at breast height (DBH), stem location, leaf area index, and understory conditions that are difficult to capture using aerial data alone (Bienert et al., 2021; Hopkinson et al., 2004; Liang et al., 2014; Pueschel et al., 2013).

Therefore, this dissertation seeks to fill the gap by integrating aerial, terrestrial, and mobile LiDAR with field-based ecological analysis to better understand the spatial distribution, structure, and regeneration dynamics of pine species in the southern Appalachian region. The study was conducted in Dupont State Recreational Forest, Cedar Mountain, North Carolina covering a 26 hectare forest area. While our analysis was conducted at a relatively small scale, the approach can be applied across broader landscapes. This capability offers a pathway toward developing efficient, spatially explicit forest monitoring tools that support large-scale restoration and adaptive management of pine-oak ecosystems in the southern Appalachians region.

1.2 Dissertation Objectives

The overall objective was to develop and evaluate LiDAR-based methods for detecting pine species, quantify forest structural attributes, and explore the ecological and environmental factors that influence the spatial distribution of overstory and establishment of pine regeneration. I developed four chapters to achieve this objective:

1. *Leveraging deep learning and low density aerial LiDAR to segment individual trees and distinguish pines in a mixed pine-hardwood forest.* In this chapter, we utilized the publicly available aerial LiDAR data from 3DEP LiDAR explorer to identify the pine and non-pine species using 2D Convolutional Neural Network achieving an overall accuracy of 85%.
2. *Spatial Patterns Assessment of Pine in Mixedwood Forests of the Southern Appalachians*. In this chapter, we explored the environmental drivers responsible for the distribution of pine in southern Appalachian region. Additionally, we analyzed the spatial pattern of pine in that area. Our result showed that environmental factors significantly affected the pine

distribution. Additionally, when the effect of those environmental variables were removed, pine were still clustered at small spatial scale showing interaction between trees.

3. *Establishment pattern, disturbance history, and factors affecting yellow pine regeneration patterns in southern pine-oak mixedwoods*. In this chapter, we explored the establishment pattern of pine and the disturbance history. Additionally, we dived into how the environmental variables, and metrics such as canopy openness is affecting the overstory and understory plant composition. Our result showed that the overstory composition were mostly affected by the environmental variables such as slope and elevation while the understory vegetation were strongly influenced by overstory and understory canopy openness and soil variables.
4. *The influence of understory vegetation density on the accuracy of modeling tree structure using terrestrial and handheld mobile laser scanning*. In this chapter, we looked at how the varying understory in a pine-oak mixedwood forest would influence the stem detection and DBH estimation accuracy of the ground based laser scanner. Our results showed that understory vegetation significantly affected the stem detection rate and DBH estimation accuracy for both ground based laser scanner.

References

Arthur, M. A., Paratley, R. D., & Blankenship, B. A. (1998). Single and Repeated Fires Affect Survival and Regeneration of Woody and Herbaceous Species in an Oak-Pine Forest. *The Journal of the Torrey Botanical Society*, *125*(3), 225–236. https://doi.org/10.2307/2997220

Bienert, A., Georgi, L., Kunz, M., Von Oheimb, G., & Maas, H.-G. (2021). Automatic extraction and measurement of individual trees from mobile laser scanning point clouds of forests. *Annals of Botany*, *128*(6), 787–804. https://doi.org/10.1093/aob/mcab087

Boyden, S., Binkley, D., & Shepperd, W. (2005). Spatial and temporal patterns in structure, regeneration, and mortality of an old-growth ponderosa pine forest in the Colorado Front Range. *Forest Ecology and Management*, *219*(1), 43–55. https://doi.org/10.1016/j.foreco.2005.08.041

Clabo, D., & Clatterbuck, W. (2020). Restoration of Shortleaf Pine (Pinus echinata)-Hardwood Mixtures in Low Quality Mixed Upland Hardwood Stands Using Cluster Planting and Natural Regeneration. *Forests*, *11*(4), 457. https://doi.org/10.3390/f11040457

Després, T., Vítková, L., Bače, R., Čada, V., Janda, P., Mikoláš, M., Schurman, J. S., Trotsiuk, V., & Svoboda, M. (2017). Past disturbances and intraspecific competition as drivers of spatial pattern in primary spruce forests. *Ecosphere*, *8*(12), e02037. https://doi.org/10.1002/ecs2.2037

Elliott, K. J., Vose, J. M., Knoepp, J. D., & Clinton, B. D. (2012). Restoration of shortleaf pine (Pinus echinata)-hardwood ecosystems severely impacted by the southern pine beetle

(Dendroctonus frontalis). *Forest Ecology and Management*, *274*, 181–200. https://doi.org/10.1016/j.foreco.2012.02.034

Goode, J. D., Hart, J. L., Dey, D. C., Torreano, S. J., & Clark, S. L. (2021). Spatial Patterns of Canopy Disturbance and Shortleaf Pine in a Mixedwood Forest. *Forest Science*, *67*(4), 433–445. https://doi.org/10.1093/forsci/fxab017

Guldin, J. M. (2019). Restoration of Native Fire-Adapted Southern Pine-Dominated Forest Ecosystems: Diversifying the Tools in the Silvicultural Toolbox. *Forest Science*. https://doi.org/10.1093/forsci/fxz005

Haines, T. K., Busby, R. L., & Cleaves, D. A. (2001). Prescribed Burning in the South: Trends, Purpose, and Barriers. *Southern Journal of Applied Forestry*, *25*(4), 149–153. https://doi.org/10.1093/sjaf/25.4.149

Harding, D. J., Blair, J. B., Garvin, J. B., & Lawrence, W. T. (1994). Laser altimetry waveform measurement of vegetation canopy structure. *Proceedings of IGARSS '94 - 1994 IEEE International Geoscience and Remote Sensing Symposium*, *2*, 1251–1253. https://doi.org/10.1109/IGARSS.1994.399398

Hopkinson, C., Chasmer, L., Young-Pow, C., & Treitz, P. (2004). Assessing forest metrics with a ground-based scanning lidar. *Canadian Journal of Forest Research*, *34*(3), 573–583. https://doi.org/10.1139/x03-225

Jenkins, M. A., Klein, R. N., & McDaniel, V. L. (2011). Yellow pine regeneration as a function of fire severity and post-burn stand structure in the southern Appalachian Mountains. *Forest Ecology and Management*, *262*(4), 681–691. https://doi.org/10.1016/j.foreco.2011.05.001

Lafon, C. W., Grissino-Mayer, H. D., Horn, S. P., & Klein, R. N. (2010). *Fire Regimes of the Southern Appalachian Mountains: Temporal and Spatial Variability over Multiple Scales and Implications for Ecosystem Management*.

Lafon, C. W., & Kutac, M. J. (2003). Effects of Ice Storms, Southern Pine Beetle Infestation, and Fire on Table Mountain Pine Forests of Southwestern Virginia. *Physical Geography*, *24*(6), 502–519. https://doi.org/10.2747/0272-3646.24.6.502

Liang, X., Hyyppa, J., Kukko, A., Kaartinen, H., Jaakkola, A., & Yu, X. (2014). The Use of a Mobile Laser Scanning System for Mapping Large Forest Plots. *IEEE Geoscience and Remote Sensing Letters*, *11*(9), 1504–1508. https://doi.org/10.1109/LGRS.2013.2297418

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*(2), 247–267. https://doi.org/10.1016/0034-4257(88)90028-4

Nordman, C., Faber-Langendoen, D., & Baggs, J. (2021). Rapid Ecological Integrity Assessment Metrics to Restore Wildlife Habitat and Biodiversity for Shortleaf Pine–Oak Ecosystems. *Forests*, *12*(12), 1739. https://doi.org/10.3390/f12121739

Oswalt, C. M., Oswalt, S. N., & Meade, J. R. (2016). Species Composition and Succession in Yellow Pine Stands Following Southern Pine Beetle Outbreaks in Tennessee-Preliminary Results. *In: Proceedings of the 18th Biennial Southern Silvicultural Research Conference. e–Gen. Tech. Rep. SRS–212. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station.*, *212*.

Pueschel, P., Newnham, G., Rock, G., Udelhoven, T., Werner, W., & Hill, J. (2013). The influence of scan mode and circle fitting on tree stem detection, stem diameter and volume extraction from terrestrial laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing*, *77*, 44–56. https://doi.org/10.1016/j.isprsjprs.2012.12.001

Randles, R. B., Lear, D. H. V., Waldrop, T. A., & Simon, D. M. (2002). *PERIODIC BURNING IN TABLE MOUNTAIN-PITCH PINE STANDS*.

Rudjord, O., & Trier, O. D. (2016). Tree species classification with hyperspectral imaging and lidar. *2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1–4. https://doi.org/10.1109/WHISPERS.2016.8071665

Waldron, J. D., Lafon, C. W., Coulson, R. N., Cairns, D. M., Tchakerian, M. D., Birt, A., & Klepzig, K. D. (2007). Simulating the impacts of southern pine beetle and fire on the dynamics of xerophytic pine landscapes in the southern Appalachians. *Applied Vegetation Science*, *10*(1), 53–64. https://doi.org/10.1111/j.1654-109X.2007.tb00503.x

Welch, N. T., Waldrop, T. A., & Buckner, E. R. (2000). Response of southern Appalachian table mountain pine (*Pinus pungens*) and pitch pine (*P. rigida*) stands to prescribed burning. *Forest Ecology and Management*, *136*(1), 185–197. https://doi.org/10.1016/S0378-1127(99)00291-1

Wiggins, H. L., Nelson, C. R., Larson, A. J., & Safford, H. D. (2019). Using LiDAR to develop high-resolution reference models of forest structure and spatial pattern. *Forest Ecology and Management*, *434*, 318–330. https://doi.org/10.1016/j.foreco.2018.12.012

Zhang, C., & Qiu, F. (2012). Mapping Individual Tree Species in an Urban Forest Using Airborne Lidar Data and Hyperspectral Imagery. *Photogrammetric Engineering & Remote Sensing*, *78*(10), 1079–1087. https://doi.org/10.14358/PERS.78.10.1079

CHAPTER 2 - Leveraging deep learning and low-density aerial LiDAR to segment individual trees and distinguish pines in a mixed pine-hardwood forest

2.1 Abstract

Advances in remote sensing and machine learning offer promising tools to map and classify forest structure across large areas, although less often successful in complex mixed pine-hardwood forests. In this study, we segmented individual trees and evaluated the performance of a two-dimensional convolutional neural network (2D-CNN) in classifying trees into pine and non-pine groups using publicly available aerial LiDAR data from the USGS 3D Elevation Program. The approach was applied to a 26 hectare area of mixed pine-oak stands in the southern Appalachian Mountains. A canopy height model with a 0.4 m resolution was developed and used to identify the individual tree locations using a local maximum filtering with a variable window size. Clustering algorithms including both K-means, and density-based spatial clustering of application with noise (DBSCAN) were used to segment individual trees from the study area. The LiDAR point cloud was processed to extract individual tree crowns, and 3D crown structure was projected into a 2D raster representation for classification. To assess the effectiveness of the 2D-CNN, we compared its performance with two traditional machine learning classifiers, Random Forest (RF) and Support Vector Machine (SVM). The 2D-CNN model outperformed RF and SVM, achieving an overall accuracy of 85% in classifying both pine and non-pine species groups. In contrast, SVM achieved an accuracy of 72 and 78% for pine and non-pine species, respectively while RF achieved 74% accuracy for both groups. This study highlights the potential of leveraging deep learning and publicly available aerial LiDAR data for cost-effective forest classification, offering a scalable approach for species group identification in complex forest structures.

2.2 Introduction

Aerial LiDAR is an active remote sensing technology that is most frequently used in forest research (Donager et al., 2021) due to its capability to capture efficient and large scale forest inventory data (Wing et al., 2012). The three-dimensional (3D) structural information captured by the LiDAR can be used for tree classification due to the morphological differences in foliage arrangements and branching structures among different tree species (Shi et al., 2018). Additionally, LiDAR derived metrics such as crown shape, tree height, and crown volume have demonstrated significant potential in species differentiation due to their inter-species variability (Michałowska & Rapiński, 2021; Shi et al., 2018). For instance, coniferous species such as spruce (*Picea* spp.) generally have conical crown shapes, while deciduous trees typically exhibit more rounded or spherical forms (Ørka et al., 2009). The precise extraction of these structural features, crucial for effective species classification, is highly dependent on the accuracy of the individual tree segmentation algorithm (Ayrey et al., 2017; Li et al., 2012; Reitberger et al., 2009).

Various algorithms have been developed to segment individual trees from aerial LiDAR data. For example, individual trees were segmented directly from the LiDAR point cloud by using a top to bottom region growing method that sequentially segments trees from tallest to shortest (Li et al., 2012). In contrast, a bottom-up approach was proposed that would leverage the intensity and topological relationships of points to identify tree trunks in a leaf-off deciduous forest (Lu et al., (2014). Local maximum filter was utilized to detect and measure individual trees (Popescu & Wynne, 2004), while marker-controlled watershed segmentation has been

applied to isolate trees in savanna woodlands using treetops identified as local maxima within a canopy height model (Chen et al., 2006). Tree segmentation has also been achieved by dividing the LiDAR point cloud into 1-meter horizontal sections and performing k-means clustering in each layer using treetops as seed points (Ayrey et al., 2017). However, the accuracy of these segmentation algorithms is highly influenced by LiDAR point density, forest types, and structural complexity (Ma et al., 2022; Vauhkonen et al., 2012). Comparisons of multiple segmentation algorithms across various tree species have shown that deep learning methods, such as PointNet++ can achieve high accuracy with an F score of 0.91 (Liu et al., 2023). Furthermore, forest composition plays a crucial role in segmentation accuracy, with forests of homogeneous structure such as plantations yielding better segmentation results relative to more structurally heterogeneous forest (Ma et al., 2022).

While the importance of accurate tree segmentation in species classification is undeniable, the algorithm used for the classification also plays a critical role in determining the overall effectiveness and accuracy of the process. Techniques in traditional machine learning (ML), including logistic regression, linear discriminant analysis, support vector machine (SVM), and random forest (RF) have been widely applied for species classification tasks (Ba et al., 2020; Korpela et al., 2010; Suratno et al., 2009; Zhang & Liu, 2013). For instance, SVM demonstrated 83.5% accuracy in classifying riparian trees to genera (Ba et al., 2020), and 88% accuracy in classifying nine different species (Dalponte et al., 2019). Although traditional ML techniques provide promising results, deep learning methods, particularly convolutional neural networks (CNNs), are able to automatically extract features from raw input data eliminating the need for manual feature engineering that requires extensive domain expertise (Mäyrä et al., 2021; Sothe et al., 2020). Additionally, CNNs have been reported to provide higher accuracy in classifying tree

species than SVM and RF in some studies such as a comparison of RF and SVM with CNN, indicated the CNN led to an increase in overall accuracy of up to 30% in classifying eight dominant tree species (Hartling et al. 2019). Similarly, 2D-CNN outperformed RF and SVM with an overall accuracies of 84.4% and 74.95% in classifying sixteen tree species across two subtropical forest fragments (Sothe et al. 2020).

Given the potential advantages of CNN in classification, it is critical to assemble a method combining CNN with detection using low cost and readily available LiDAR products so that the workflow can be applied to landscapes undergoing significant ecological challenges. An example of this type of landscape includes the xeric pine (*Pinus*)-oak (*Quercus*) forest ecosystems in the southern Appalachian Mountains, USA, that are experiencing significant decline driven by factors such as southern pine beetle (*Dendroctonus frontalis)* infestation, land use pattern change, fire suppression, deforestation, and drought (Lafon et al., 2007, 2022; Lafon & Kutac, 2003; Turner et al., 2003; Waldron et al., 2007). Basic information regarding the establishment and spatial distribution pattern of the pine populations are critical for large-scale restoration efforts. While planting efforts may establish the goal of increasing species occurrence they may not successfully restore the ecosystem structure and function of these vulnerable communities if they historically maintained non-uniform distributions.

In this study, we evaluated methods for segmentation of individual trees from low density aerial LiDAR and compared the performance of 2D-CNN, RF, and SVM for distinguishing pines from hardwoods in a mixed pine-oak stand in the southern Appalachian Mountains. To mitigate the costly challenges of acquiring multiple remote sensing datasets, this research leveraged publicly available leaf-off aerial LiDAR data from USGS 3DEP LiDAR explorer for segmentation and classification. Field data collected in 2023 were used for model training and

accuracy assessments, helping to validate the classification results. To the best of our knowledge, this is the first application of 3DEP products for classification of pine and non-pine species. Specifically, the study seeks to address the following question:

1. What is the accuracy of 2D-CNN in identifying the occurrence of pine and non-pine species using the USGS aerial LiDAR data?
2. What is the relative performance of 2D-CNN for distinguishing broad species types such as pines from hardwoods compared to RF and SVM?

2.3 Methodology

2.3.1 Study Area

The DuPont State Recreational Forest encompasses more than 4000 hectares in the upland plateau of the Little River Valley within Transylvania and Henderson counties, North Carolina. Across the entire forest, elevation ranges from 701 to 1097 meters. We selected a study area 26 hectares in size (Figure 2.1) that consists of mixed oak and pine, including shortleaf pine (*Pinus echinata*), pitch pine (*Pinus rigida*), Table Mountain pine (*Pinus pungens*), Virginia pine (*Pinus virginiana*), and eastern white pine (*P. strobus*). From 1981-2022, the mean annual temperature ranged from 12.98 to 15.68 °C, mean annual precipitation from 843.75 to 1792.97 mm, and mean relative humidity ranged from 69 to 81.38% (NASA Langley Research Center, 2022). This area encompasses a diverse range of soil types which are characterized by varying slopes, drainage capacities, and compositions. Predominantly, the soils in the region are well-drained with 8 to 15 percent slopes. Soils present in both moderate and steep slopes (8 to 95 percent), are somewhat excessively drained, and feature rocky and very rocky textures.

Additionally, areas with fine sandy loam are notable for its frequent flooding at gentle slopes, 0 to 2 percent (*Web Soil Survey*, n.d.).

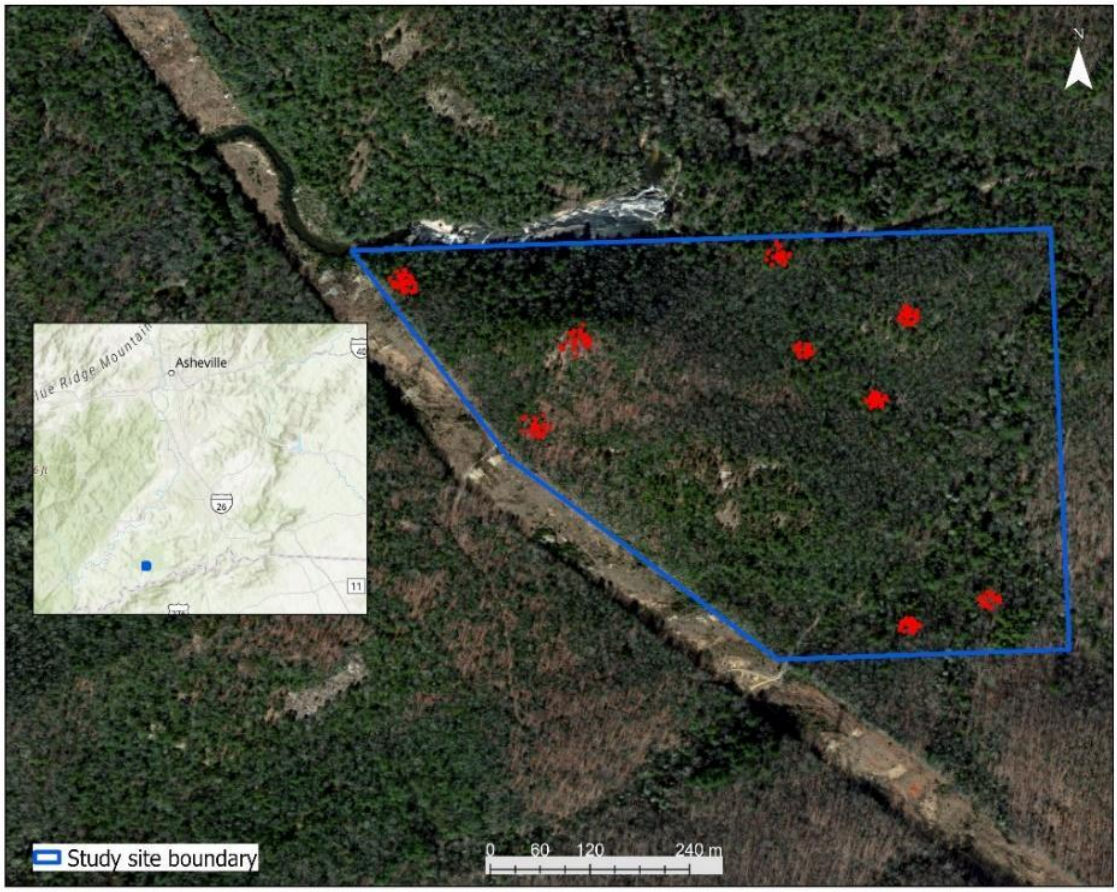

Figure 2.1: Study area at DuPont State Recreational Forest. The blue outlined area is the study area of 26 ha. The red circles represent locations of the fixed area plots used for ground truth data.

2.3.2 Field Data Collection

Nine circular fixed radius plots 0.04 hectares in size were randomly placed throughout the study area. Geographic coordinates of the plot center were recorded using a GNSS GPS unit (Bad Elf GNSS Surveyor). Within each plot, all trees with a diameter at breast height (DBH,

measured at 1.37 m from ground) greater than 10 cm were surveyed. For these trees, distance and azimuth from plot center, species, DBH to the nearest tenth cm, and crown class was recorded. The crown class was determined through visual assessment and was categorized as dominant, co-dominant, intermediate, and suppressed. In addition, height to the nearest meter for the three tallest trees was recorded.

2.3.3 LiDAR Data Processing

Aerial LiDAR data was extracted from the 3DEP LiDAR explorer (U.S. Geological Survey, 2021) for the entire study area. The point density averaged 73.85 points per square meter across the study area (Figure 2.2). The LiDAR point cloud data were pre-classified by USGS into 5 different classes: unclassified, ground, low vegetation, medium vegetation, and high vegetation and defined as:

- Unclassified: points which were processed but not assigned to a category.
- Ground: points representing the bare earth surface.
- Low Vegetation: points representing vegetation less than 1 meter in height.
- Medium Vegetation: points representing vegetation between 1 and 3 meters in height.
- High Vegetation: points representing vegetation greater than 3 meters in height.

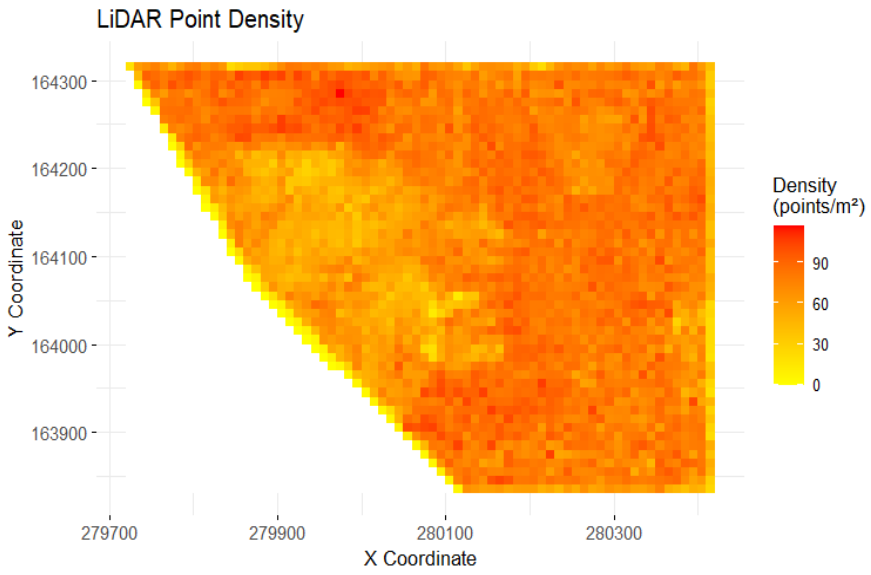

Figure 2.2: LiDAR point density of the study area.

Data was processed and analyzed using the lidR (Roussel et al., 2020), stats (R Core Team, 2024), and terra (Hijmans, 2020) packages in R v.4.4.0 (R Core Team, 2024), and the keras (Chollet & others, 2015) and scikit-learn library (Pedregosa et al., 2011) in Python (Van Rossum & Drake Jr, 1995) (Figure 2.3).

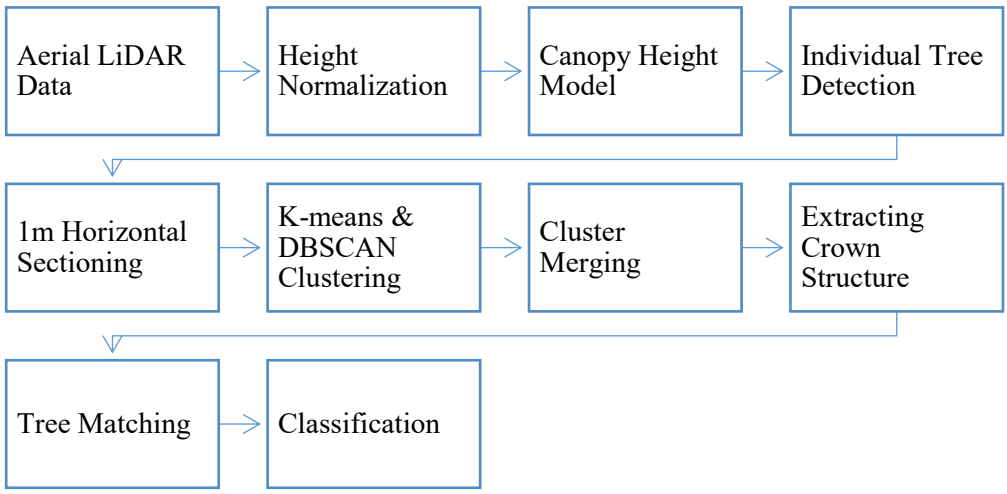

Figure 2.3: Workflow of the research methodology used up to the point of classification comparisons

2.3.3.1 Height Normalization and Canopy Height Model

To accurately analyze the vertical structure of the forest, the LiDAR point cloud was normalized by substituting the vertical position, or z coordinate, of each point with its respective distance from the ground using the normalize height function within the lidR package. Subsequently, a canopy height model (CHM) with 0.4 m resolution was generated from the normalized data using the rasterize canopy function and the pitfree algorithm (Khosravipour et al., 2014). The CHM was smoothed using a 3×3 moving window, where the mean height value was calculated for each neighborhood to reduce noise and enhance the detection of individual tree crowns.

2.3.3.2 Individual Tree Crown Detection and Delineation

Individual tree crown (ITC) detection and delineation using remotely sensed data is vital for efficient, precise, and thorough forest monitoring (Zhen et al., 2016). In this study, treetops were detected from the smoothed CHM using the locate trees function and Local Maximum Filtering (LMF) algorithm within the lidR package. The effectiveness of LMF algorithm in identifying tree location depends on selecting an appropriate window size. A fixed window size may lead to errors in complex forest structures such as hardwood or mixed pine-hardwood forest. To address this issue, a variable radius window was employed which adapts to diverse tree crown shapes and sizes (Popescu et al., 2002). The variable window size was defined by a linear function that adjusts the search radius based on the pixel height in the CHM. The window size was defined as:

$$r = 0.08h + 2$$

where, r is the window radius and h is the pixel height in CHM.

Various ITC delineation methods have been developed to identify tree crowns (Dalponte et al., 2015; Li et al., 2012), each with varying effectiveness depending on the forest characteristics being analyzed (Vauhkonen et al., 2014). We customized the approach proposed by Ayrey et al., (2017) modifying it to suit the characteristics of our data and study objectives. Firstly, LiDAR points representing unclassified, ground, low and medium vegetation were removed from the normalized point cloud, resulting in dataset containing only points above 5m in height. The remaining point cloud was divided into 1m horizontal sections (Figure 2.4), creating 36 layers in total. Within each layer, k-means clustering was performed (Hartigan & Wong, 1979) using *kmeans* function from the stats package in R, with previously detected treetops used as initial cluster centers. However, not all detected treetops were used as cluster centers across all layers. For example, if tree height was 25.4 m, the treetop was excluded as a cluster center for layers 25 and above. This modification prevents the algorithm from erroneously clustering points above the actual height of a tree.

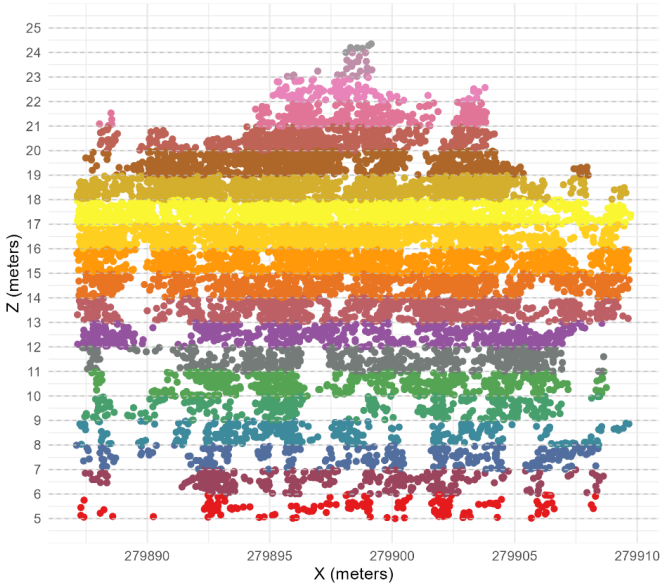

Figure 2.4: An example of one-meter horizontal sectioning of the tree crown point cloud.

K-means works by iteratively updating cluster centers to minimize within-cluster variance. However, since the algorithm relies on predefined cluster centers, it forces the data into a specified number of clusters, even when there are clear gaps, particularly in layers where data might be missing for certain trees. This limitation can lead to over-segmentation or invalid clusters in such cases. To address this issue, the Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al., 1996) was applied to refine the clusters identified by k-means in each layer (Figure 2.5) using the dbscan package in R (Hahsler & Piekenbrock, 2015). DBSCAN clusters points based on two key parameters: epsilon (eps), which defines the maximum distance between individual points in order for them to be considered part of the same cluster, and MinPts, the minimum number of points required to form a cluster. For this study, the eps and MinPts values were set to 0.4 and 5 respectively. For the clustering process, only the X and Y coordinates of each point were used within each horizontal layer as the layers were so finely defined. After the clustering, all points identified as noise by the DBSCAN algorithm were removed.

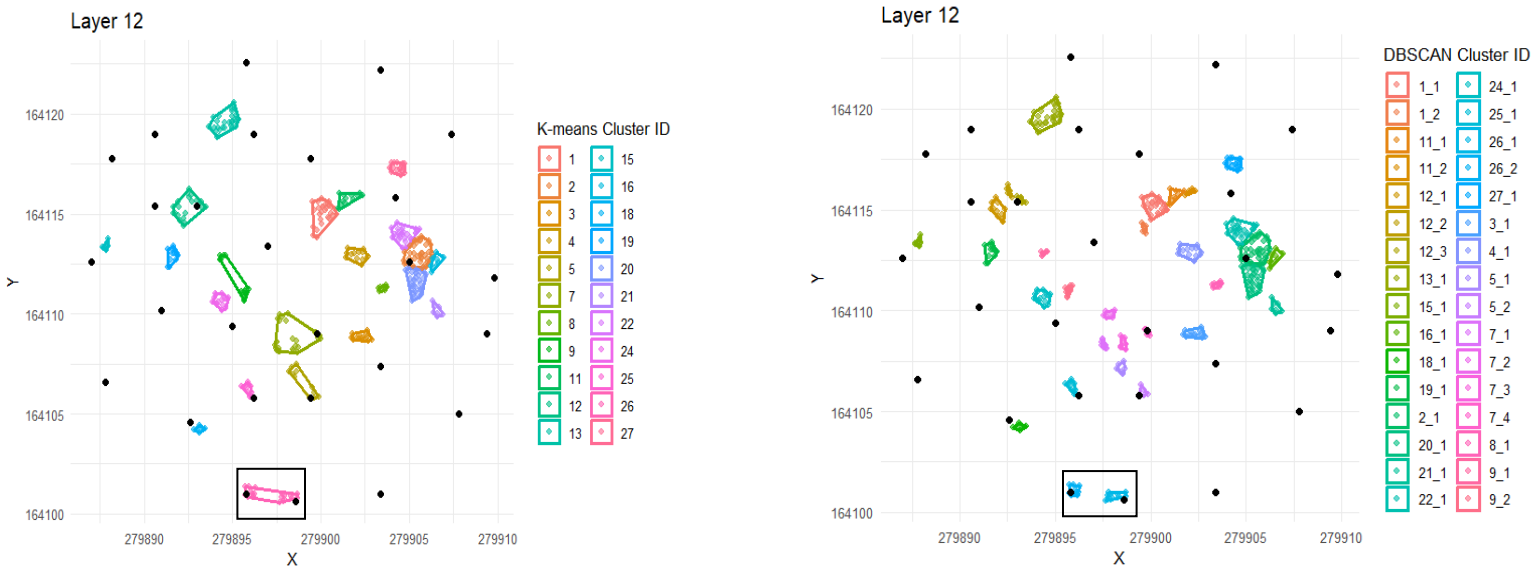

Figure 2.5: Layer 12 showing clustering by K-means and DBSCAN. The rectangle box shows how K-means assigned the same Cluster ID (represented by the convex hull) to the points which was later separated into two clusters by DBSCAN clustering. Black points represent the tree locations.

A unique identifier, referred to as a Cluster ID, was assigned to each cluster within each independent layer. However, because this process was layer specific, the Cluster IDs varied between layers, even for clusters belonging to the same tree. To address this issue and ensure continuity across layers, clusters were merged based on their spatial relationships. A convex hull was created for each cluster and where a treetop intersected the hull that cluster was assigned the treetop Tree ID. For the remaining clusters that did not intersect with treetops, a centroid was calculated for each layer. The distances between the cluster centroids and the treetop coordinates derived from the CHM were measured, and the cluster centroid closest to a treetop was grouped and assigned the treetop Tree ID. However, clusters that were further than 2 m from the treetop were removed as it often tended to have crown segments from neighboring tree. This process ensured that clusters spanning multiple layers were accurately grouped as a single tree, thereby enhancing the reliability of the delineation results (Figure 2.6).

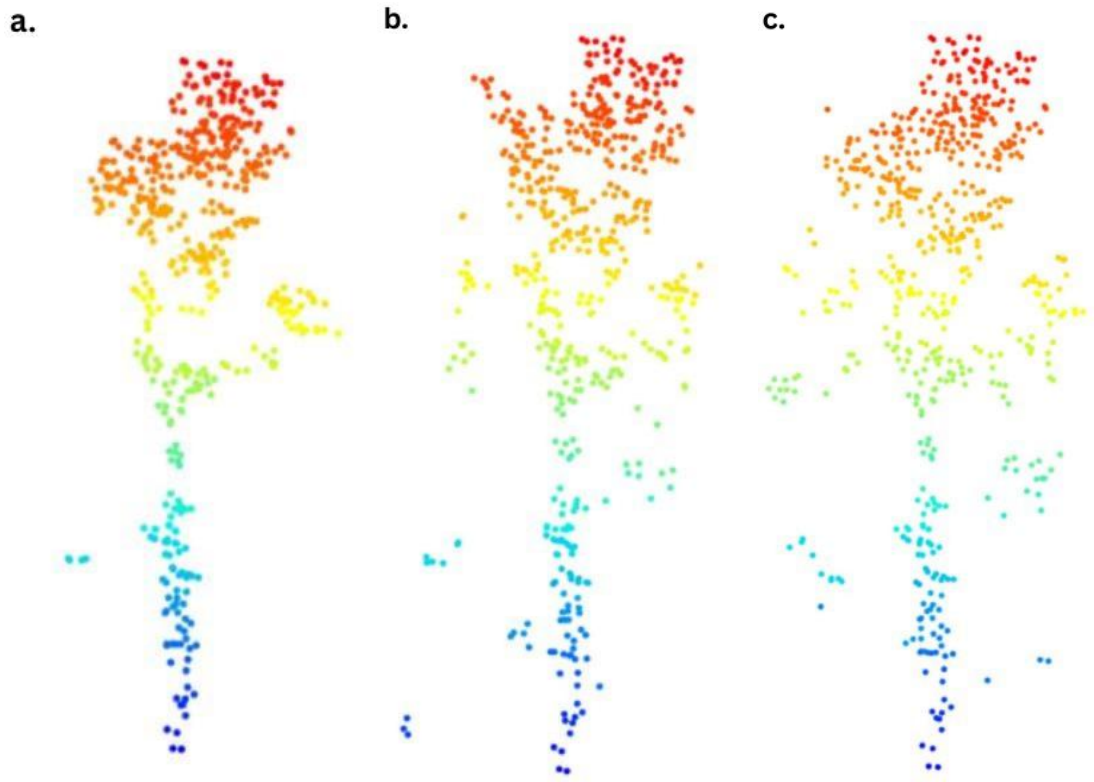

Figure 2.6: Segmentation resulting from three different algorithms: a) our approach vs b) dalponte vs c) watershed.

2.3.3.3 Classification

Following the segmentation procedures, we developed methods to classify individuals into pine and non-pine species categories. The distinctive conical shape of pine crowns help to differentiate them from broadleaved species. To focus on this structural feature, crown data was extracted by filtering all the points belonging to the top 60% of tree height for each tree. This threshold was selected because it effectively captured the crown shape of individual trees minimizing the inclusion of lower stem data where tree crown shape becomes irregular, making it less valuable for the species classification (Figure 2.7).

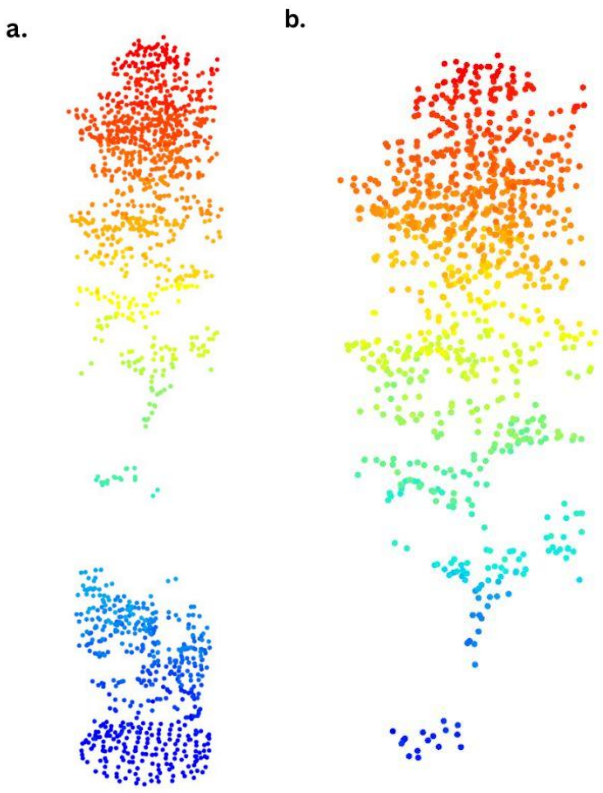

Figure 2.7: Comparison of a) Complete Tree Data and b) Extracted Crown Data (Top 60% of Tree Height).

We imported the crown segments, LiDAR derived tree locations, and field tree locations into ArcGIS Pro to visually pair each segmented tree crown, with a corresponding field measured tree. If a single field tree point overlapped a crown segment, we labelled that tree crown with that matched field point. In cases where a crown segment contained more than one field-measured tree location, we calculated the distance between the LiDAR derived tree locations and field tree locations and assigned the crown segment to the field tree nearest the centroid. This approach ensured that each tree crown segment was matched to the most appropriate field-measured tree, even in areas where tree locations were densely packed.

The extracted 3D crown data was converted into multi-band raster, with pixel values representing mean intensity, point density, and Z values. These rasters served as input for a 2D CNN to perform categorical classification of individual trees into pine and non-pine species. The raster data was split into training and testing sets, comprising 70% and 30% of the data, respectively. During training, 20% of the training data was further set aside as a validation dataset to monitor model performance. The CNN architecture, built using Keras Sequential API, consisted of three convolutional layers with kernel size of 3x3 with ReLU activation (Nair & Hinton, 2010). Each convolutional layer was followed by max pooling the layers to reduce the spatial dimensions of the feature maps and limit overfitting. The final feature maps were flattened into a 1D vector and passed through a dense layer with 64 neurons. A softmax activation function was used in the output layer to perform binary classification. The model was compiled using the Adam optimizer (Kingma & Ba, 2017), and the binary cross-entropy loss function, appropriate for binary classification tasks. The model was trained for 50 epochs with a batch size of 32, using both training and validation datasets. Early stopping was used to prevent overfitting (Prechelt, 2012).

The RF (Breiman, 2001) and SVM (Awad & Khanna, 2015) classifications were implemented using the sckit-learn package in python. The 2D raster was flattened into a 1D vector to use as an input and hyperparameters were optimized using randomized search with 5-fold cross validation (Bergstra & Bengio, 2012).

2.3.4 Data Analysis

The performance of the algorithms were assessed by comparing the classification output from RF, SVM and 2D-CNN model to the test dataset. A confusion matrix was constructed to

identify true positive (TP), false positive (FP), true negative (TN), and false negative (FN) classifications for the test set. To assess the classification process, accuracy, precision, sensitivity, and F-score performance metrics were calculated. Accuracy was defined as the proportion of observations accurately predicted (De Diego et al., 2022):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \#(1)$$

Precision was defined as the ratio of correctly identified positive instances to the total number of instances that the model classified as positive:

$$Precision = \frac{TP}{TP + FP} \quad \#(2)$$

Sensitivity/Recall was defined as the ratio of correctly identified positive instances to the total number of actual positive instances in the dataset (Sokolova & Lapalme, 2009):

$$\frac{Sensitivity}{Recall} = \frac{TP}{TP + FN} \quad \#(3)$$

The F-score is a metric that combines both precision and sensitivity as follows:

$$F - score = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity} \quad \#(4)$$

2.4 Results

2.4.1 Field trees matching with LiDAR data

A total of 306 trees (including pine and hardwood individuals) were measured in the inventory plots, out of which 123 trees were matched with the LiDAR derived segments (Table

2.1). Trees in dominant crown positions had the highest matching rate of 94.3% followed by co-dominant of 47.0%. Mid-canopy trees matched less frequently and no suppressed or overtopped individuals matched to the LiDAR segments.

Table 2.1: Matched rate for comparison of trees classified from LiDAR with field measured individuals by canopy classes. The total number of trees per class is displayed with the relative percentage of all trees shown in parentheses.

Canopy Class	**Measured Tree Count**	**Matched Tree Count**	**Unmatched Tree Count**
Dominant (D)	70 (22.88%)	66 (94.29%)	4 (5.71%)
Co-dominant (CD)	83 (27.12%)	39 (46.99%)	44 (53.01%)
Intermediate (I)	81 (26.47%)	18 (22.2%)	63 (77.7%)
Suppressed (S)	72 (23.53%)	0 (0%)	72 (100%)
Total	306 (100%)	123 (40.20%)	183 (59.80%)

Although pines were less frequent in the dataset (40% versus 60% hardwoods), they had a higher matching rate of 55.7% compared to the non-pine group with 30.1% (Table 2.2). Additionally, 72% of the matched pine trees were of dominant crown class with 23.53% and 4.41% being of co-dominant and intermediate crown class respectively. For the remaining 54 pines that we were not able to match, 7.4% were dominant, 42.6% were co-dominant, 14.8% were intermediate, and 35.2% were suppressed. Matching was again higher for non-pine trees in dominant (n=17), co-dominant (n=23), and intermediate (n=15) crown positions.

Table 2.2: Matched rate for pine and non-pine species class

Species Class	Matched	Unmatched	Total
Pine	68 (55.74%)	54 (44.26%)	122 (39.86 %)
Non-Pine	55 (30.05%)	129 (69.94%)	184 (60.13%)

The mean diameter of matched trees exceeded that of unmatched trees across all canopy classes and species groups, with the sole exception of pines in the intermediate crown class (Figure 2.8). While the pine group exhibited larger mean diameters in the dominant crown class, this pattern reversed in codominant, intermediate, and suppressed classes, where non-pine species had greater mean diameters.

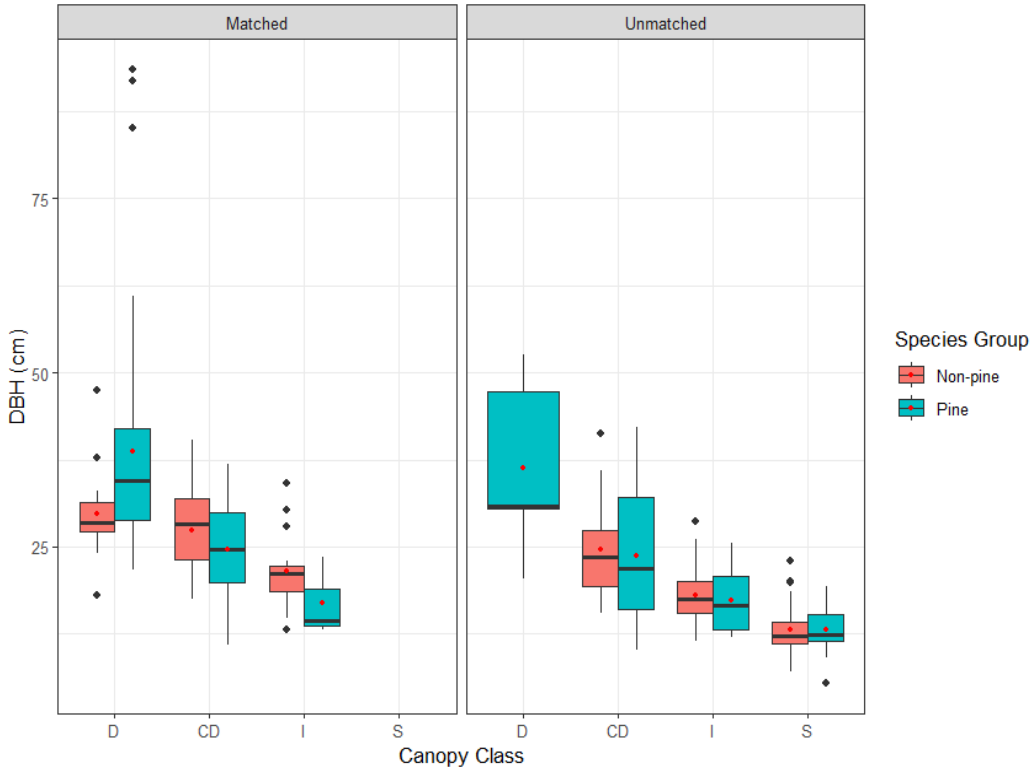

Figure 2.8: Comparison of diameter ranges between pine and non-pine groups across crown class.

2.4.2 2D CNN classification results

2D CNN outperformed both RF and SVM in classifying both pine and non-pine trees (Table 2.3). Out of 38 test samples, the 2D CNN model correctly classified 33, achieving an overall accuracy of 85%. The confusion matrix shows robust predictive capabilities of CNN, with 18 pine individuals correctly identified and only one misclassified as non-pine trees (Figure 2.9). For non-pines, 15 were correctly classified while 4 were misidentified as pine. For pine classification, the 2D CNN model achieved a precision of 84%, recall of 97%, and F1-score of 90%. The precision, recall, and F1-score for the non-pine classification were 97%, 81%, and 88% respectively.

In contrast, SVM had the lowest overall accuracy of 75%. It performed better in classifying the non-pine trees with an F1-score of 0.78 compared to pines which had an F1-score of 0.72. The RF model's performance was between that of the 2D CNN and SVM with an overall accuracy of 77%. It showed balanced performance across both classes, with identical F1-scores of 0.74 for both pine and non-pine classification.

Table 2.3: Comparison of test results from the RF, SVM, and 2D CNN models.

Species	**RF**			**SVM**			**2D CNN**		
Group	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Pine	0.72	0.89	0.74	0.70	0.63	0.72	0.81	0.91	0.86
Non-Pine	0.86	0.65	0.74	0.84	0.63	0.78	0.90	0.79	0.84
Overall Accuracy			0.77			0.75			0.85
Macro Average	0.79	0.77	0.77	0.77	0.75	0.75	0.91	0.89	0.89
Weighted Average	0.79	0.77	0.77	0.77	0.75	0.75	0.90	0.89	0.89

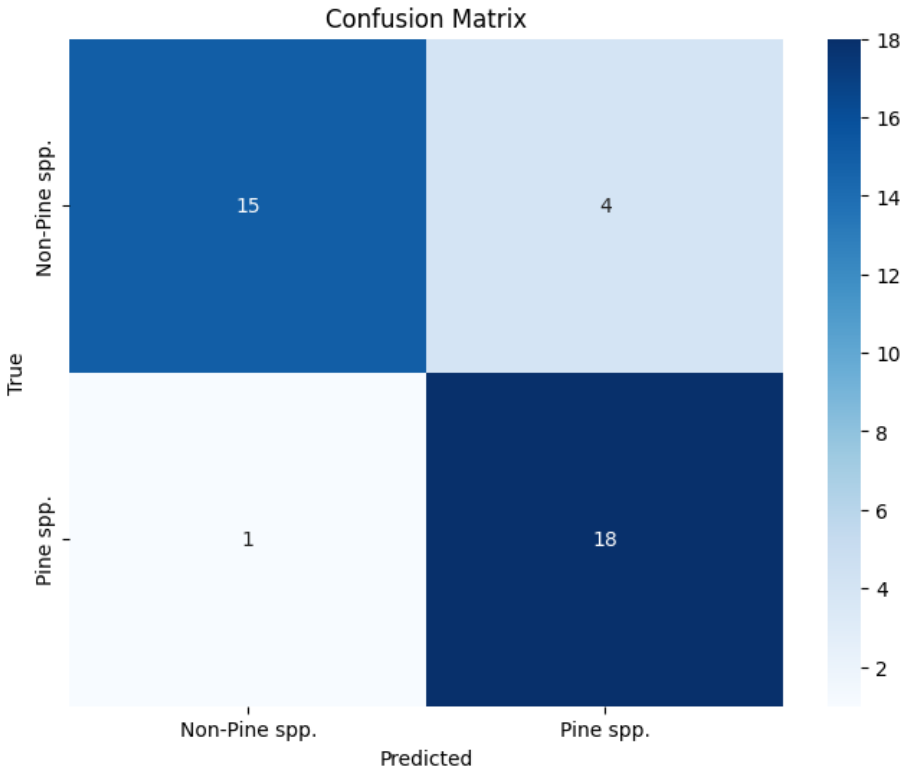

Figure 2.9: Confusion matrix for the 2D CNN model used to classify 38 test trees into either pine or non-pine categories.

2.4.3 Classification for the Study Area

We selected the 2D CNN model to classify tree types for the whole study area. Figure 2.10 shows the model prediction of pine and non-pine types. Out of the 7,142 trees, model predicted 25.29% (1,806) to be pines and the remaining 74.71% (5,336) to be non-pines.

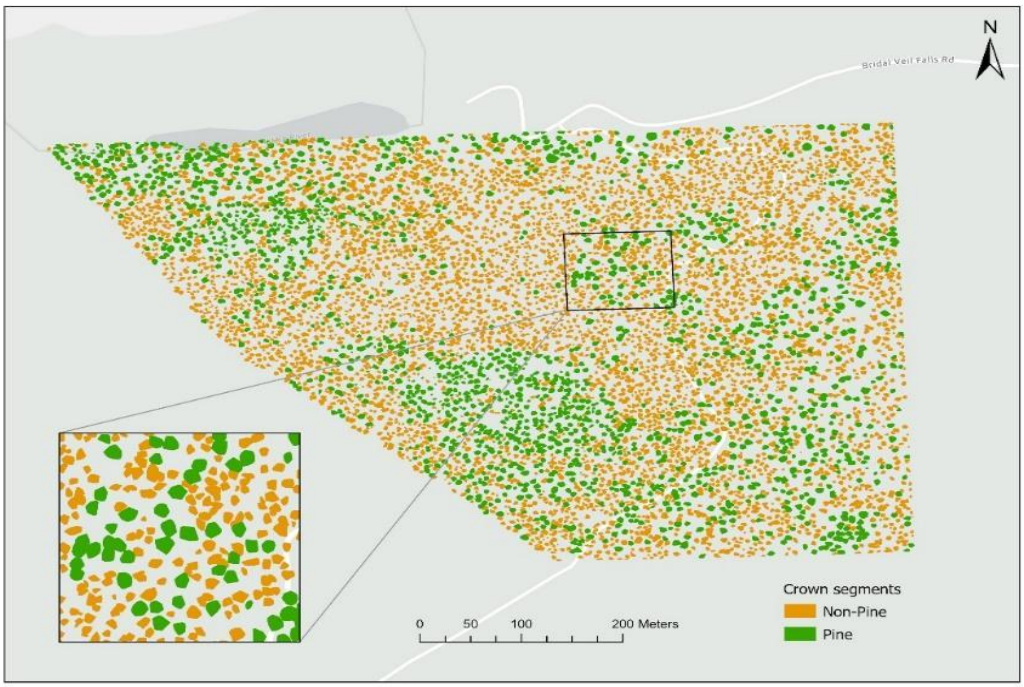

Figure 2.10: Map showing individual trees classified into pine and non-pine categories.

2.5 Discussion

This study highlights the potential of leveraging 2D-CNN and publicly available aerial LiDAR data for efficient individual tree segmentation and classification in forest ecosystems, achieving relative success in a mixed pine-hardwood forest in the southern Appalachian Mountains. Our results demonstrated the capacity of 2D-CNN in identifying pine with high accuracy, outperforming traditional machine learning algorithms such as RF and SVM. The 2D-CNN model performed better on differentiating pine from hardwood trees in our study area, similar to success reported in previous studies (Hartling et al., 2019; Mäyrä et al., 2021; Sothe et al., 2020). SVM had the lowest performance out of the three models and classified non-pine trees better than pine. In the case of RF, it showed a balanced performance across both classes, with identical F1-scores for both pine and non-pine classification.

The matching rate of trees in this study was 40.2% which is comparable though on the low side of previous studies that used aerial LiDAR data to match LiDAR derived tree locations to field tree locations. For example, prior research has reported matching rates ranging from 40% to 69% across various forest conditions and a high proportion when considering crown positions (60-70% of dominant trees matched) (Eysn et al., 2015; Pang et al., 2021; Pitkänen et al., 2004; Yu et al., 2011). Notably, our study achieved even higher matching rate for dominant trees at 94.3%, indicating that the discrepancy in overall matching rates was largely driven by inherent limitation of CHM in detecting suppressed trees (Pitkänen et al., 2004). Prior comparisons of tree detection algorithms indicated that the method using local maxima with a 3×3 fixed radius window resulted in the highest matching rate of 54% but also the highest commission rate of 113%, indicating a tendency to overestimate the numbers of trees (Eysn et al., 2015). Conversely, the method that used local maxima with filtering had the matching rate of 45% with the lowest commission error of 9%. Additionally, the method that used local maxima with 7×7 fixed radius window had the highest omission rate of 63%. These findings underscore the inherent trade-offs in tree detection methods, where higher matching rates often come at the cost of increased commission errors, and vice versa.

2.5.1 Forest Complexity

The complexity of mixed-species forests poses further challenges for accurate tree detection. Forests with high species diversity, significant variations in tree ages, and heterogeneous growth conditions tend to exhibit higher commission and omission errors in individual tree detection (Sun et al., 2022). Additionally, forest type and structure have been found to have significant impacts on the detection rates, with omission errors ranging from 25% in a single layered mixed forests to as high as 75% in a multi layered mixed forests (Torresan et al., 2020). For instance, in

mixed forests, commission error mostly occurred for broadleaf trees which have multiple peaks leading to multiple local maxima detection (Qin et al., 2022). In contrast, coniferous trees typically exhibit a well-defined crown shape, reducing the commission error (Eysn et al., 2015).

To mitigate these errors, hybrid methods combining trunk detection and crown segmentation have been proposed (Deng et al., 2024). These methods first identify potential trunk position from the LiDAR data which may include false trunks. Detected trunks are validated using the CHM and crown segments, ensuring only trunks with corresponding crowns are retained. This reduces the risks of over segmentation in a mixed-natural stands where trees with varying crown shape and sizes exists. However, such approaches have primarily been applied to highly dense point clouds (892 to 1090 points/m^2), which are not applicable for a USGS aerial LiDAR data due to its lower point density.

2.5.2 Limitations and Uncertainties

A crucial factor to consider in our study is the temporal discrepancy between the aerial LiDAR data acquisition and the field data collection. The LiDAR data obtained from USGS was collected in 2017, while the field data was gathered in 2023. This time discrepancy could have resulted in the lower matching rate, particularly for the sub-canopy individuals, raising the question of whether this level of accuracy is sufficient for certain applications. While the matching rate may not be ideal, it is important to note that even with low-density LiDAR, incorporating additional ground measurements could enhance model performance and improve matching success. To avoid using the biased tree heights, we were limited to using categorical canopy positions or DBH distributions of both matched and unmatched pine and non-pine species.

The limited number of field observations combined with the available LiDAR data reduced our ability for species-level classification. High density LiDAR is often required for species level classification. For instance, an overall classification accuracy of 90.8% for pine species, spruce species and deciduous trees was achieved using LiDAR data with over 500 points/m^2 (Hakula et al., 2023). In contrast, the USGS LiDAR dataset used in this study is a single-return LiDAR, meaning each emitted pulse produced only a single return point. This limitation significantly constrained the ability to capture detailed vertical canopy structure, particularly in dense areas where multiple vegetation layers exist.

Despite these challenges, our study demonstrates that distinguishing pine and non-pine species is feasible using single-return LiDAR. Future research could improve classification performance by incorporating multi-return, higher-density LiDAR and expanding field observations to enhance model training and validation.

2.6 Conclusion

One of the most significant contributions of this research is the demonstration of how publicly available LiDAR data can be effectively utilized for segmentation and classification in species rich forests occurring on complex terrain. By relying on a single, widely accessible dataset, this study addresses the challenges associated with integrating multiple remote sensing datasets, such as high costs, computational complexities, and inconsistencies in spatial and temporal resolutions (Balestra et al., 2024). We propose a workflow that provides a cost effective and scalable solution for forest inventory and monitoring, particularly for regions with limited access to expensive hyperspectral or multispectral imagery.

References

Awad, M., & Khanna, R. (2015). Support Vector Machines for Classification. In M. Awad & R. Khanna, *Efficient Learning Machines* (pp. 39–66). Apress. https://doi.org/10.1007/978-1-4302-5990-9_3

Axelsson, A., Lindberg, E., & Olsson, H. (2018). Exploring Multispectral ALS Data for Tree Species Classification. *Remote Sensing*, *10*(2), 183. https://doi.org/10.3390/rs10020183

Ayrey, E., Fraver, S., Kershaw, J. A., Kenefic, L. S., Hayes, D., Weiskittel, A. R., & Roth, B. E. (2017). Layer Stacking: A Novel Algorithm for Individual Forest Tree Segmentation from LiDAR Point Clouds. *Canadian Journal of Remote Sensing*, *43*(1), 16–27. https://doi.org/10.1080/07038992.2017.1252907

Ba, A., Laslier, M., Dufour, S., & Hubert-Moy, L. (2020). Riparian trees genera identification based on leaf-on/leaf-off airborne laser scanner data and machine learning classifiers in northern France. *International Journal of Remote Sensing*, *41*(5), 1645–1667. https://doi.org/10.1080/01431161.2019.1674457

Balestra, M., Marselis, S., Sankey, T. T., Cabo, C., Liang, X., Mokroš, M., Peng, X., Singh, A., Stereńczak, K., Vega, C., Vincent, G., & Hollaus, M. (2024). LiDAR Data Fusion to Improve Forest Attribute Estimates: A Review. *Current Forestry Reports*, *10*(4), 281–297. https://doi.org/10.1007/s40725-024-00223-7

Bergstra, J., & Bengio, Y. (2012). *Random Search for Hyper-Parameter Optimization*.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data. *Photogrammetric Engineering & Remote Sensing*, *72*(8), 923–932. https://doi.org/10.14358/PERS.72.8.923

Dalponte, M., Frizzera, L., & Gianelle, D. (2019). Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. *PeerJ*, *6*, e6227. https://doi.org/10.7717/peerj.6227

Dalponte, M., Orka, H. O., Gobakken, T., Gianelle, D., & Naesset, E. (2013). Tree Species Classification in Boreal Forests With Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(5), 2632–2645. https://doi.org/10.1109/TGRS.2012.2216272

Dalponte, M., Reyes, F., Kandare, K., & Gianelle, D. (2015). Delineation of Individual Tree Crowns from ALS and Hyperspectral data: A comparison among four methods. *European Journal of Remote Sensing*, *48*(1), 365–382. https://doi.org/10.5721/EuJRS20154821

De Diego, I. M., Redondo, A. R., Fernández, R. R., Navarro, J., & Moguerza, J. M. (2022). General Performance Score for classification problems. *Applied Intelligence*, *52*(10), 12049–12063. https://doi.org/10.1007/s10489-021-03041-7

Deng, S., Jing, S., & Zhao, H. (2024). A Hybrid Method for Individual Tree Detection in Broadleaf Forests Based on UAV-LiDAR Data and Multistage 3D Structure Analysis. *Forests*, *15*(6), 1043. https://doi.org/10.3390/f15061043

Donager, J. J., Sánchez Meador, A. J., & Blackburn, R. C. (2021). Adjudicating Perspectives on Forest Structure: How Do Airborne, Terrestrial, and Mobile Lidar-Derived Estimates Compare? *Remote Sensing*, *13*(12), 2297. https://doi.org/10.3390/rs13122297

Ester, M., Kriegel, H.-P., & Xu, X. (1996). *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*.

Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D., & Pfeifer, N. (2015). *A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space*.

Hahsler, M., & Piekenbrock, M. (2015). *dbscan: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Related Algorithms* (p. 1.2.1) [Dataset]. https://doi.org/10.32614/CRAN.package.dbscan

Hakula, A., Ruoppa, L., Lehtomäki, M., Yu, X., Kukko, A., Kaartinen, H., Taher, J., Matikainen, L., Hyyppä, E., Luoma, V., Holopainen, M., Kankare, V., & Hyyppä, J. (2023). Individual tree segmentation and species classification using high-density close-range multispectral laser scanning data. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, *9*, 100039. https://doi.org/10.1016/j.ophoto.2023.100039

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, *28*(1), 100. https://doi.org/10.2307/2346830

Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., & Carron, J. (2019). Urban Tree Species Classification Using a WorldView-2/3 and LiDAR Data Fusion Approach and Deep Learning. *Sensors*, *19*(6), 1284. https://doi.org/10.3390/s19061284

Hijmans, R. J. (2020). *terra: Spatial Data Analysis* (p. 1.8-15) [Dataset]. https://doi.org/10.32614/CRAN.package.terra

Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. (2014). Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing*, *80*(9), 863–872. https://doi.org/10.14358/PERS.80.9.863

Kingma, D. P., & Ba, J. L. (2017). *Adam: A Method for Stochastic Optimization* (arXiv:1412.6980). arXiv. http://arxiv.org/abs/1412.6980

Korpela, I., Ørka, H., Maltamo, M., Tokola, T., & Hyyppä, J. (2010). Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, *44*(2). https://doi.org/10.14214/sf.156

Lafon, C. W., DeWeese, G. G., Flatley, W. T., Aldrich, S. R., & Naito, A. T. (2022). Historical Fire Regimes and Stand Dynamics of Xerophytic Pine–Oak Stands in the Southern Appalachian Mountains, Virginia, USA. *Annals of the American Association of Geographers*, *112*(2), 387–409. https://doi.org/10.1080/24694452.2021.1935206

Lafon, C. W., & Kutac, M. J. (2003). Effects of Ice Storms, Southern Pine Beetle Infestation, and Fire on Table Mountain Pine Forests of Southwestern Virginia. *Physical Geography*, *24*(6), 502–519. https://doi.org/10.2747/0272-3646.24.6.502

Lafon, C. W., Waldron, J. D., Cairns, D. M., Tchakerian, M. D., Coulson, R. N., & Klepzig, K. D. (2007). Modeling the Effects of Fire on the Long-Term Dynamics and Restoration of Yellow Pine and Oak Forests in the Southern Appalachian Mountains. *Restoration Ecology*, *15*(3), 400–411. https://doi.org/10.1111/j.1526-100X.2007.00236.x

Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M. (2012). A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogrammetric Engineering & Remote Sensing*, *78*(1), 75–84. https://doi.org/10.14358/PERS.78.1.75

Liu, Y., You, H., Tang, X., You, Q., Huang, Y., & Chen, J. (2023). Study on Individual Tree Segmentation of Different Tree Species Using Different Segmentation Algorithms Based on 3D UAV Data. *Forests*, *14*(7), 1327. https://doi.org/10.3390/f14071327

Lu, X., Guo, Q., Li, W., & Flanagan, J. (2014). A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *94*, 1–12. https://doi.org/10.1016/j.isprsjprs.2014.03.014

Ma, K., Chen, Z., Fu, L., Tian, W., Jiang, F., Yi, J., Du, Z., & Sun, H. (2022). Performance and Sensitivity of Individual Tree Segmentation Methods for UAV-LiDAR in Multiple Forest Types. *Remote Sensing*, *14*(2), 298. https://doi.org/10.3390/rs14020298

Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T., & Vihervaara, P. (2021). Tree

species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sensing of Environment*, *256*, 112322. https://doi.org/10.1016/j.rse.2021.112322

Michałowska, M., & Rapiński, J. (2021). A Review of Tree Species Classification Based on Airborne LiDAR Data and Applied Classifiers. *Remote Sensing*, *13*(3), 353. https://doi.org/10.3390/rs13030353

Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27 Th International Conference on Machine Learning*.

Ørka, H. O., Næsset, E., & Bollandsås, O. M. (2009). Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment*, *113*(6), 1163–1174. https://doi.org/10.1016/j.rse.2009.02.002

Pang, Y., Wang, W., Du, L., Zhang, Z., Liang, X., Li, Y., & Wang, Z. (2021). Nyström-based spectral clustering using airborne LiDAR point cloud data for individual tree segmentation. *International Journal of Digital Earth*, *14*(10), 1452–1476. https://doi.org/10.1080/17538947.2021.1943018

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.

Persson, M., Lindberg, E., & Reese, H. (2018). Tree Species Classification with Multi-Temporal Sentinel-2 Data. *Remote Sensing*, *10*(11), 1794. https://doi.org/10.3390/rs10111794

Pitkänen, J., Maltamo, M., Hyyppä, J., & Yu, X. (2004). *ADAPTIVE METHODS FOR INDIVIDUAL TREE DETECTION ON AIRBORNE LASER BASED CANOPY HEIGHT MODEL*.

Popescu, S. C., & Wynne, R. H. (2004). Seeing the Trees in the Forest: Using Lidar and Multispectral Data Fusion with Local Filtering and Variable Window Size for Estimating Tree Height. *Photogrammetric Engineering & Remote Sensing*, *70*(5), 589–604. https://doi.org/10.14358/PERS.70.5.589

Popescu, S. C., Wynne, R. H., & Nelson, R. F. (2002). Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture*, *37*(1–3), 71–95. https://doi.org/10.1016/S0168-1699(02)00121-7

Prechelt, L. (2012). Early Stopping—But When? In G. Montavon, G. B. Orr, & K.-R. Müller (Eds.), *Neural Networks: Tricks of the Trade* (Vol. 7700, pp. 53–67). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35289-8_5

Pu, R. (2021). Mapping Tree Species Using Advanced Remote Sensing Technologies: A State-of-the-Art Review and Perspective. *Journal of Remote Sensing*, *2021*, 2021/9812624. https://doi.org/10.34133/2021/9812624

Qin, H., Zhou, W., Yao, Y., & Wang, W. (2022). Individual tree segmentation and tree species classification in subtropical broadleaf forests using UAV-based LiDAR, hyperspectral, and ultrahigh-resolution RGB data. *Remote Sensing of Environment*, *280*, 113143. https://doi.org/10.1016/j.rse.2022.113143

R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Reitberger, J., Schnörr, Cl., Krzystek, P., & Stilla, U. (2009). 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(6), 561–574. https://doi.org/10.1016/j.isprsjprs.2009.04.002

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., Boissieu, F. de, & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, *251*. https://doi.org/10.1016/j.rse.2020.112061

Shi, Y., Wang, T., Skidmore, A. K., & Heurich, M. (2018). Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS Journal of Photogrammetry and Remote Sensing*, *137*, 163–174. https://doi.org/10.1016/j.isprsjprs.2018.02.002

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Sothe, C., De Almeida, C. M., Schimalski, M. B., La Rosa, L. E. C., Castro, J. D. B., Feitosa, R. Q., Dalponte, M., Lima, C. L., Liesenberg, V., Miyoshi, G. T., & Tommaselli, A. M. G. (2020). Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. *GIScience & Remote Sensing*, *57*(3), 369–394. https://doi.org/10.1080/15481603.2020.1712102

Sun, C., Huang, C., Zhang, H., Chen, B., An, F., Wang, L., & Yun, T. (2022). Individual Tree Crown Segmentation and Crown Width Extraction From a Heightmap Derived From Aerial Laser Scanning Data Using a Deep Learning Framework. *Frontiers in Plant Science*, *13*, 914974. https://doi.org/10.3389/fpls.2022.914974

Suratno, A., Seielstad, C., & Queen, L. (2009). Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(6), 683–693. https://doi.org/10.1016/j.isprsjprs.2009.07.001

Terryn, L., Calders, K., Disney, M., Origo, N., Malhi, Y., Newnham, G., Raumonen, P., Å Kerblom, M., & Verbeeck, H. (2020). Tree species classification using structural features derived from terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, 170–181. https://doi.org/10.1016/j.isprsjprs.2020.08.009

Torresan, C., Carotenuto, F., Chiavetta, U., Miglietta, F., Zaldei, A., & Gioli, B. (2020). Individual Tree Crown Segmentation in Two-Layered Dense Mixed Forests from UAV LiDAR Data. *Drones*, *4*(2), 10. https://doi.org/10.3390/drones4020010

Turner, M. G., Pearson, S. M., Bolstad, P., & Wear, D. N. (2003). Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). *Landscape Ecology*, *18*(5), 449–464. https://doi.org/10.1023/A:1026033116193

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam. http://www.python.org

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M., Lien, V., Packalen, P., Gobakken, T., Koch, B., Naesset,

E., Tokola, T., & Maltamo, M. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, *85*(1), 27–40. https://doi.org/10.1093/forestry/cpr051

Vauhkonen, J., Maltamo, M., McRoberts, R. E., & Næsset, E. (2014). Introduction to Forestry Applications of Airborne Laser Scanning. In M. Maltamo, E. Næsset, & J. Vauhkonen (Eds.), *Forestry Applications of Airborne Laser Scanning* (Vol. 27, pp. 1–16). Springer Netherlands. https://doi.org/10.1007/978-94-017-8663-8_1

Waldron, J. D., Lafon, C. W., Coulson, R. N., Cairns, D. M., Tchakerian, M. D., Birt, A., & Klepzig, K. D. (2007). Simulating the impacts of southern pine beetle and fire on the dynamics of xerophytic pine landscapes in the southern Appalachians. *Applied Vegetation Science*, *10*(1), 53–64. https://doi.org/10.1111/j.1654-109X.2007.tb00503.x

Web Soil Survey. (n.d.). Retrieved January 24, 2025, from https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx

Wing, B. M., Ritchie, M. W., Boston, K., Cohen, W. B., Gitelman, A., & Olsen, M. J. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment*, *124*, 730–741. https://doi.org/10.1016/j.rse.2012.06.024

Yao, W., Krzystek, P., & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, *123*, 368–380. https://doi.org/10.1016/j.rse.2012.03.027

Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., & Viitala, R. (2011). Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 28–37. https://doi.org/10.1016/j.isprsjprs.2010.08.003

Zhang, Z., & Liu, X. (2013). Support vector machines for tree species identification using LiDAR-derived structure and intensity variables. *Geocarto International*, *28*(4), 364–378. https://doi.org/10.1080/10106049.2012.710653

Zhen, Z., Quackenbush, L., & Zhang, L. (2016). Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. *Remote Sensing*, *8*(4), 333. https://doi.org/10.3390/rs8040333

CHAPTER 3 - Spatial patterns of pine in mixedwood forests of the southern Appalachians

3.1 Abstract

The spatial distribution of trees within a forest is determined by complex interactions between species and their environment. Pines occurring on ridges and peaks across a wide elevational gradient are an important component of the southern Appalachian forests both ecologically and economically. This study investigated the spatial patterns and environmental drivers of pine (*Pinus* spp.) distribution in the mixed oak-pine forests of Southern Appalachian Mountain. Individual pine locations were mapped using a combination of field surveys and aerial LiDAR data and classified into three height classes. Environmental variables including elevation, slope, aspect and topographic position index (TPI) were used to model the influence of environment in the distribution of pine using inhomogeneous Poisson process model. Inhomogeneous Ripley's L function and inhomogeneous pair correlation function were used to determine the spatial pattern of pines in the study area. We found environmental variables strongly influenced the heterogeneous distribution of pine (p-value < 0.05). However, the significance of the environmental variables differed by height classes. The density of pines in dominant and co-dominant canopy positions decreased with elevation, while the density of intermediate positioned pines increased with elevation. Intermediate pines were concentrated on xeric ridges, codominant pines favored both ridges and upper slope while dominant pines were more evenly distributed across topographic positions. Second order analysis indicated clustering at small scales, <10 m for intermediate pines, and < 25 m for both co-dominant and dominant pines. Bivariate pair correlation function showed positive association between intermediate and co-dominant pines at scale > 15 and < 20 m. However, both intermediate and dominant, and co-dominant and dominant pine had negative association. These findings highlight the critical role

of topographic heterogeneity and spatial processes in structuring the pine populations in the southern Appalachians.

3.2 Introduction

The spatial pattern, including the distribution and arrangement of trees within a forest, results from long-term interactions between the trees and the environment and direct tree-tree interactions. Spatial patterns play a critical role in shaping ecosystem functions and ecological processes (Larson & Churchill, 2012; Maestre et al., 2012) such as water retention capacity (Guo et al., 2003), net primary production and water fluxes (Simioni et al., 2003), and understory light availability which influences growth and competition in the regeneration layer leading to variation of the understory plant composition (Battaglia et al., 2002; Turner & Franz, 1986). While plantations are designed to have evenly spaced regular distributions, naturally regenerated or unmanaged stands are characterized by irregular and often complex aggregated or clustered patterns often driven by factors such as limited seed dispersal mechanisms and habitat heterogeneity (Liu et al., 2021). These spatial patterns reflect life history strategies that are intricately connected to resource availability. Separating the effects of large-scale variation in habitat quality (referred to as first-order effects) from direct tree-tree interaction (referred to as second-order effects) on the resulting spatial patterns remains a major challenge (Getzin et al. 2008).

Tree distributions are scale dependent with ecological processes and patterns manifesting at different spatial scales (Ostapowicz et al., 2008). At a landscape scale, factors such as topography, climate and disturbance shape the overall forest structure and distribution, whereas at a stand or fine scale, factors like microclimate, mortality and competition between species and

individuals become more prominent (Addington et al., 2018; Shive et al., 2018). This sensitivity complicates the selection of an optimal scale recommended for assessing the condition of vegetation (Lindenmayer et al., 2008). For example, tree distributions in a naturally regenerated forest changed from a random distribution within a smaller observation window (80m by 80m) to a clustered distribution using an expanded observation window of 500m by 500m (Wang et al., 2020).

The spatial distribution of trees in a forest can shift over time from clumped to random to regular patterns due to mortality caused by competition for resources among neighboring trees and localized disturbances (Peet & Christensen, 1987). For instance, in a naturally regenerated pure jack pine stand that originated after a wildfire, live jack pine showed a regular distribution while dead ones were clumped (Kenkel, 1988). This pattern was attributed to intraspecific competition occurring in two phases: initially a two-sided competition for nutrients and water leading to the mortality of the clumped, suppressed young trees followed by a one-sided competition for light. The shifts are often in response to a combination of natural and anthropogenic disturbances caused by climate change, land use change, mortality, fire, timber harvest, past management, silvicultural treatments, microsite variability, and resource availability (Cannon et al., 2022; Franklin et al., 2002; Hessburg et al., 1999; Rodman et al., 2017; Wolf, 2005). For instance, in a semi-natural, mixed deciduous forest in Denmark, distribution patterns shifted from regular to random occurrences over a period of 50 years due to gap recruitment and management history (Wolf, 2005).

The regional distribution of pine species in southern Appalachians has been described in well-known detail by Whittaker (1956). Contemporary pine ecosystems historically dominated by *Pinus* subg. *Pinus* or the yellow pines, including shortleaf (*P. echinata*), pitch (*P. rigida*),

Virginia (*P. virginiana*), and Table Mountain pine (*P. pungens*) are experiencing significant decline due to factors such as fire suppression, land use change, deforestation, drought, and southern pine beetle (*Dendroctonus frontalis)* infestation (Lafon et al., 2007, 2022; Lafon & Kutac, 2003; Turner et al., 2003; Waldron et al., 2007). Fire suppression has led to increased density in the midstory and overstory, fuel accumulation, failed regeneration of oak and pine species, and encroachment by the mesic hardwood species (Elliott et al., 1999). These mixedwoods ecosystems require intricate management strategies to restore and maintain a balance of species mixtures (Goode et al., 2021). The pine component in mixed forests, contribute to increased vertical structural complexity and ecosystem functions as they have high live crown ratios, maintain foliage throughout the year and can extend 10-15 m above the hardwood canopy creating distinct habitats and adding substantial deadwood inputs (Fahey & Lorimer, 2013; Goode et al., 2021).

Given this importance, restoration efforts are being intensified through joint public-private sector initiatives centered on reforestation, prescribed burning, and sustainable forest management practices to conserve and restore these critical more open pine habitats (Matthews et al., 2020). Restoration efforts often focus on managing the spatial distribution of trees, by species, size, or age. In a fire frequent forest, restoring complex mosaics of individual trees, tree clumps, and openings can enhance forest resilience and maintain ecosystem functionality (Churchill et al., 2013). In the eastern United States, where fire suppression has long been the norm, prescribed burning is being reintroduced in pine and mixed-conifer forest to reduce the hazardous fuel buildup, lower high severity fire risks, and promote pine-oak regeneration (Kobziar et al., 2015; Schwartz et al., 2016). However, implementing effective management and restoration practices requires understanding the spatial and temporal variation within a forested

ecosystem over large areas (White & Walker, 1997; Lydersen et al., 2013) which is often unquantified and therefore spatially informed management strategies are rarely applied (Churchill et al., 2013).

Advancements in remote sensing technology such as aerial Light Detection and Ranging (LiDAR) have made it possible to map the locations of individual trees across vast forested areas. LiDAR offers detailed three-dimensional data on canopy structure, allowing accurate identification of individual tree crowns and their spatial locations. Utilizing a complete census of individual tree locations obtained from LiDAR can improves the detection of patterns by size class, leaf habit, and environmental gradients providing new insights into forest structure and dynamics at ecologically meaningful scales. Since pine distributions in the southern Appalachian Mountains are known to be influenced by both first-order environmental conditions and second-order tree-tree interactions, we sought to use a representative and previously mapped area to develop a methodology to describe the distribution and configuration of pine. Therefore, the objectives of this study were to: 1) evaluate the influence of environmental factors on the overall spatial distribution of pine and assess whether these influences differ by height class and 2) analyze the overall spatial arrangement of pine and how these arrangements vary with height class. Based on ecological history, we hypothesized that the spatial pattern of pine would exhibit a clumpy distribution due to factors such as limited seed dispersal, habitat heterogeneity, and historical disturbances. Additionally, we hypothesized that environmental factors such as slope, aspect, elevation, and topographic position index (TPI) would influence the aggregation with certain factors promoting clustering in specific topographic settings. For example, we expected pine to be more aggregated on xeric, south facing slopes and ridge positions. This study aims to provide insights into the current spatial pattern and identify processes driving these patterns,

contributing to more effective and targeted forest management strategies for pine ecosystems in the Southern Appalachian region.

3.3 Methodology

3.3.1 Study Area

The DuPont State Recreational Forest is located in Cedar Mountain, North Carolina and encompasses more than 4000 hectares. The site lies in the upland plateau of the Little River Valley within Transylvania and Henderson counties, NC. Elevation ranges from 701 to 1097 meters. The study area is 26 hectares in size (Figure 3.1) and consists of mixed oak (*Quercus* spp.) and pine (*Pinus* spp.), including shortleaf pine (*Pinus echinata*), pitch pine (*Pinus rigida*), and eastern white pine (*P. strobus*). From 1981-2022, the mean annual temperature ranged from 12.98 to 15.68 °C, mean annual precipitation from 843.75 to 1792.97 mm, and mean relative humidity ranged from 69 to 81.38% (NASA Langley Research Center, 2022). This area encompasses a diverse range of soil types which are characterized by varying slopes, drainage capacities, and compositions. Predominantly, the soils in the region are well-drained with 8 to 15 percent slopes. Soils present in both moderate and steep slopes (8 to 95 percent), are somewhat excessively drained, and feature rocky and very rocky textures. Additionally, areas with fine sandy loam are notable for its frequent flooding at gentle slopes, 0 to 2 percent (*Web Soil Survey*, n.d.).

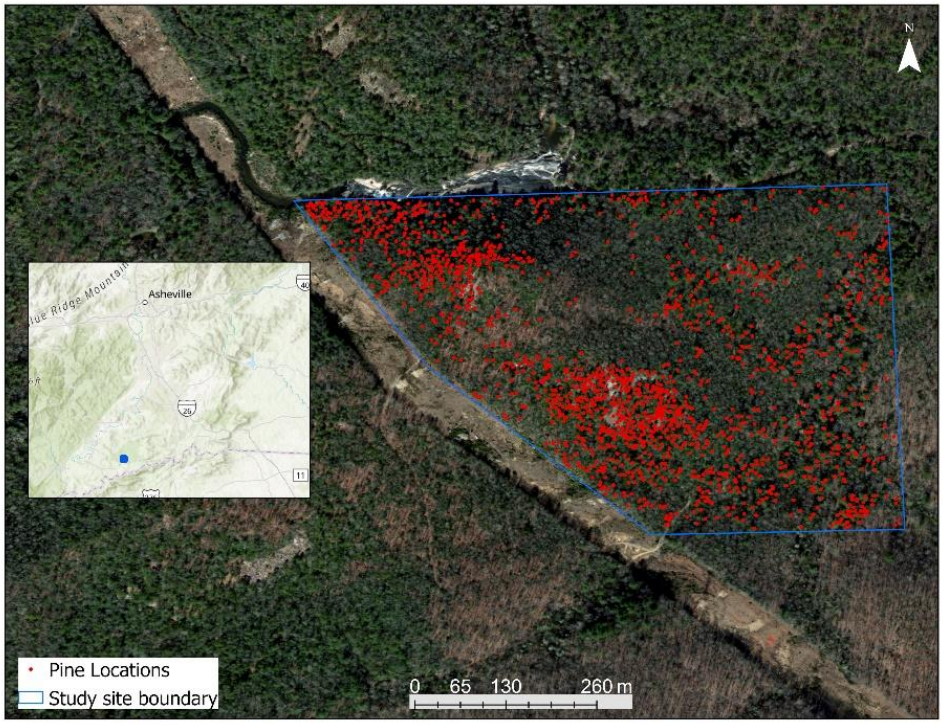

Figure 3.1: Study area at Dupont State Recreational Forest, NC. Red points represent individual pine trees including a mix of pitch pine, shortleaf pine, and Eastern white pine.

3.3.2 Data Collection

Individual pine tree locations were determined through a combination of field surveys and aerial LiDAR data, as described in Chapter 2. Each pine tree was classified into three height classes based on LiDAR-derived height metrics: small ($5.74 \leq$ height ≤ 15 m), medium (15 m $<$ height ≤ 20 m), and large (height > 20 m). These height classes were derived by comparing LiDAR tree heights to the field-recorded canopy position categories including intermediate, co-dominant and dominant. Field-measured locations were matched to LiDAR-detected tree locations manually, ensuring accurate height attribution for each tree.

Environmental variables, including elevation (meters), slope (degrees), aspect (degrees), and topographic position index (TPI) (Figure 3.2) were derived from high-resolution 3DEP

LiDAR data (*U.S. Geological Survey*, 2024). A digital terrain model (DTM) with 0.5 m resolution was generated using the grid_terrain function and the kriging algorithm within the LidR package (Roussel et al., 2020) in R (R Core Team, 2024). Slope and aspect for the entire study area was generated using the DTM and terrain() function from the terra package (Hijmans, 2020). Aspect was first converted to radians, then converted into sine and cosine components to capture the directional variation in terrain orientation. Additionally, TPI was calculated using the TPI() function within the spatialECO package (Evans et al., 2023). A focal window size of 200 was used and TPI was classified into 6 classes including ridge, upper slope, middle slope, flat area, lower slope, and valley and was based on Weiss (2001). Soil variables such as pH, organic matter, and texture were obtained from *Web Soil Survey*, (n.d.) and evaluated but found to be coarse and highly correlated with elevation so were not considered further.

Across the 26-ha study area, elevation ranged from 760 m to 882 m, and slope ranged from 0º to 86.5º reflecting the heterogeneous nature of the terrain. These topographic variables were used to assess environmental influences on pine spatial patterns and to model intensity surfaces for subsequent point pattern analysis.

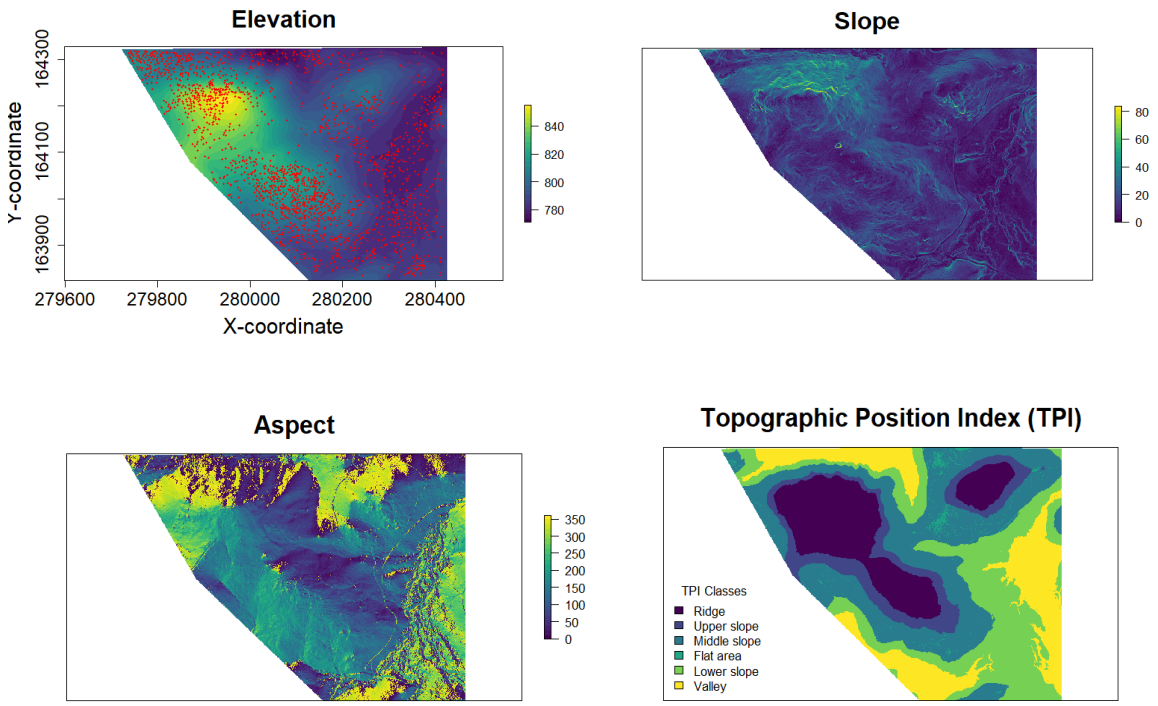

Figure 3.2: Environmental variables including elevation (meters), slope (degrees), aspect (degrees), and topographic position index used as covariates in the study to examine the influence of topographic heterogeneity on pine distribution in the southern Appalachian region. Red points overlaid in the elevation map represent pine trees.

3.3.3 Spatial Pattern Assessment

3.3.3.1 Quadrat Test for Spatial Heterogeneity

We calculated a quadrat based chi-square test of complete spatial randomness (CSR) using quadrat.test() from the spatstat package (Baddeley et al., 2025) to determine whether the distribution of pine is random, uniform or exhibits spatial heterogeneity. The study area was divided into a 3×4 grid, and the observed pine density in each quadrat was compared to expected counts under the assumption of CSR. The test statistic was calculated as $X^2 = \sum_j \frac{(n_j - e_j)^2}{e_j}$, where

n_j represents the observed count in quadrat j and e_j is the expected count under CSR. A significant result ($p < 0.05$) would indicate a spatially heterogeneous pine distribution, rejecting the null hypothesis of CSR. If CSR was accepted ($p > 0.05$), we would conclude pine was randomly distributed and no further spatial analysis would be needed.

3.3.3.2 Inhomogeneous Poisson Process with Covariates

To evaluate whether the heterogeneous pine distribution was related to environmental variables, we calculated inhomogeneous Poisson process with covariates using the ppm() function in the spatstat package in R. This approach models the intensity of pine occurrences as a function of environmental covariates as well as spatial coordinates, to account for both environmental drivers and broad spatial trends. The analysis was guided by the hypothesis that environmental variables and spatial location do not significantly influence pine distribution (H_o: $\beta_i = 0$) versus the alternative hypothesis that at least one environmental variable or spatial location significantly influenced the distribution pattern (H_a: At least one $\beta_i \neq 0$), where, β_i is the coefficients of the environmental covariates in the regression model.

The regression model for the analysis is given by:

$$\log\{\lambda(s)\} = \beta_0 + \sum_{i=1}^{p} X_i(s)\beta_i$$

where, β_0, β_1, ... , β_p are the coefficients to be estimated, $X_i(s)$ are the environmental covariates at location s, and $\lambda(s)$ denotes the spatially varying intensity of pine trees.

To evaluate the overall significance of environmental factors and spatial trends, we compared the inhomogeneous model to a homogeneous Poisson process using likelihood ratio tests via anova.ppm(). This comparison tested whether the inclusion of spatial and environmental

variables significantly improved the model fit compared to a model assuming complete spatial randomness.

3.3.3.3 Point Pattern Analysis

We employed a comprehensive analytical framework to quantify the spatial distribution patterns of pine trees using second order statistics that account for environmental heterogeneity. Our hierarchical approach examined: overall pine distribution, distribution within each height class, and spatial associations between different height classes.

To describe the overall pine distribution, we used both inhomogeneous pair correlation function g(r) and a transformation of inhomogeneous Ripley's K-function known as the Ripley's L-function and compared this with a homogeneous model to quantify the influence of environmental factors in spatial distribution of pine. The inhomogeneous pair correlation function was computed by pcfinhom() function and inhomogeneous Ripley's L-function was computed by Linhom() function from the spatstat package. Both functions are widely used spatial statistical tools to objectively assess point patterns with respect to the spatial distribution of geographic units (Wang et al., 2020).

Inhomogeneous Ripley's L function is given as:

$$L_{inhom}(r) = \sqrt{\frac{K_{inhom}(r)}{\pi}}$$

where, $K_{inhom}(r)$ is the inhomogeneous Ripley's K-function. For easier interpretation, $L_{inhom}(r)$ was transformed to $L_{inhom}(r) - r$ which centers the function around zero. $L_{inhom}(r) - r > 0$ indicates

clustering of pine at distance r, $L_{inhom}(r) - r = 0$ indicates random distribution, and $L_{inhom}(r) - r < 0$ indicates regularity or uniform distribution.

Similarly, inhomogeneous pair correlation function $g_{inhom}(r)$ is derived from the derivative of the inhomogeneous K-function and provides a more localized measure of spatial dependence

$$g_{inhom}(r) = \frac{K'_{inhom}(r)}{2\pi r}$$

where, $g_{inhom}(r)$ represents the expected density of points at a given distance r. If $g_{inhom}(r) = 1$, the species exhibit CSR at distance r, if < 1 the species are uniformly distributed and if > 1 the species are clustered at distance r.

To assess the pine distribution within each height class, we fitted three class specific inhomogeneous Poisson process models and computed the spatial arrangement using the pcfinhom() function. To evaluate the spatial association between height classes, we applied bivariate pair correlation function ($g_{ij}(r)$) using the pcfcross.inhom() function within the spatstat.explore package (Baddeley et al., 2022). Additionally, all analysis incorporated Monte Carlo significance testing (n = 199 simulations) with edge correction.

To support our visual observation of clusters of pine across the study area, we applied spatial scan statistic using scan.test function in the spatstat package. This method moves a circular scanning window of radius r across the study area and compares the observed number of pines to the number expected under the null hypothesis. The null hypothesis was the IPP model with the predicted intensity surface from the fitted model used as the baseline to account for the spatial variation in the environmental conditions. The radius for the circle was based on the pair

correlation function, where it showed significant clustering between specific distance ranges. Monte Carlo simulations (n = 199) were used to assess the statistically significant clusters.

3.4 Results

3.4.1 Quadrat test

The quadrat test indicated a heterogeneous distribution of pine trees (p-value < 0.001) across the study area (Figure 3.3). In over 70% of the sections, the observed number of trees was lower than the expected, leaving 25% of the sections to have an aggregation of occurrences higher than what would be expected. Two of the quadrats had Pearson residuals of 15 and 17 respectively, indicating strong clustering of pine trees in these areas likely reflecting underlying environmental gradients or microsite suitability. These results provide initial evidence of the heterogeneity in pine distribution justifying the use of further spatial analysis to identify the drivers of this heterogeneity.

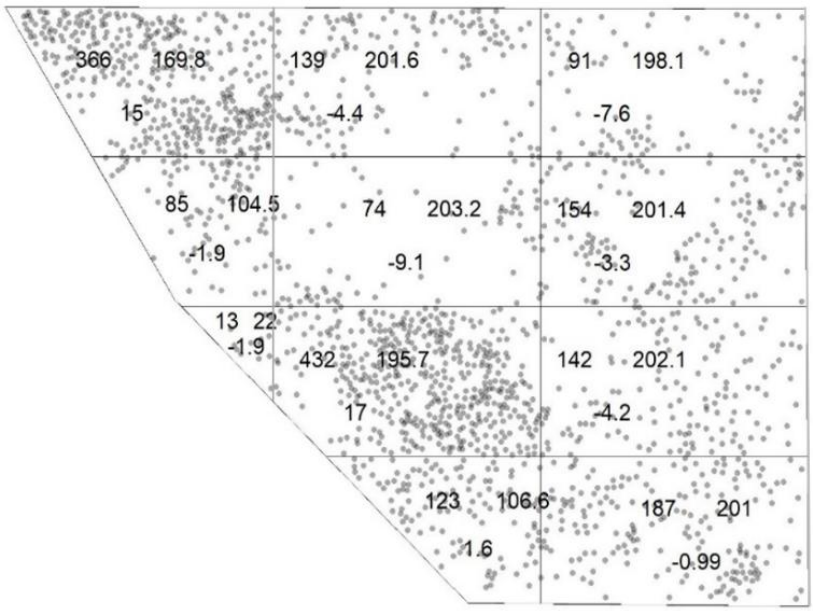

Figure 3.3: Quadrat test for spatial heterogeneity of pine tree distribution. For example, in the top left quadrat, 366 pines were observed, while the expected counts based on the area of the quadrat was 169.8 and the Pearson residuals was 15. Pearson residual closer to 0 represents that the observed value is close to the expected value while a positive residual represents more observed points than expected, and a negative residual represents less observed points than expected.

3.4.2 Inhomogeneous Poisson Process with Covariates

The IPP model, incorporating environmental covariates and spatial location, showed a highly significant fit over the homogeneous Poisson model ($p < 0.001$). Several of the environmental conditions were shown to significantly influence the occurrence of pine across the site (Table 3.1). The spatial location also significantly influenced the distribution, with both the interaction term longitude * latitude (x*y) and the quadratic term for the latitude (y^2) relating to pine intensity. However, the quadratic term for the longitude (x^2) was non-significant, suggesting

that the relationship between pine distribution and spatial variables were linear in the longitude, but exhibited curvature in the latitude.

The occurrence of pine was negatively related with elevation and aspect, but positively related to slope. The landform categories all showed significant negative coefficients compared to the reference category ridges. This is interpreted that pines were least common in valleys, lower slopes, and middle slopes relative to ridge positions (Table 3.1).

Table 3.1: Parameter estimates and statistical significance of environmental and spatial covariates in the IPP Model for pine distribution. Asterisks indicate significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Covariates	**Estimate**	**S.E.**	**Wald test**	**Z value**
x (Longitude)	-0.0029	0.00033	***	-8.68
y (Latitude)	-0.0021	0.00029	***	-7.14
$I(x^2)$	0.0000001	0.000001		-0.16
I(x*y)	-0.00001	0.000002	***	-5.64
$I(y^2)$	-0.000004	0.000002	*	-2.04
X1 (Elevation)	-0.0237	0.0041	***	-5.85
X2 (Slope)	0.0157	0.00244	***	6.44
X3 (Aspect_cos)	-0.2957	0.0093	***	-7.38
X4 (Aspect_sin)	-0.1878	0.0385	***	-4.88
X52 (Upper Slope)	-0.5719	0.1003	***	-5.70
X53 (Middle Slope)	-0.9556	0.1090	***	-8.77
X54 (Flat Area)	-0.4709	0.2431	*	-2.10
X55 (Lower Slope)	-1.0059	0.1401	***	-7.18
X56 (Valley)	-1.1026	0.1528	***	-7.22

The influence of environmental conditions varied among height classes (Table 3.2). Elevation significantly influenced the distribution of all three height classes but with contrasting

relationships. The occurrence of shorter trees increased with elevation, whereas the distribution of the intermediate and tallest trees decreased. Slope negatively influenced the shorter pine while no significant effect was observed for the taller classes. The shortest and tallest pine occurred on west facing slopes, while the intermediate height pine were found on both south and west facing slopes.

Reinforcing the elevation pattern, the occurrence of the shortest pine trees was greatest on the ridges and significantly declined in all other positions. Intermediately tall pines occurred less on upper slopes relative to ridges, though the difference in abundance was not significant. However, ridges had significantly greater pine density compared to middle slope, lower slope, flat area, and valley. In contrast, for large pine, none of the topographic positions were significantly different than ridges although the pine density increased for middle slope, lower slope, and flat area.

Table 3.2: Parameter estimates and statistical significance of environmental and spatial covariates in the IPP Model for pine distribution by height class. Asterisks indicate significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Covariates	**Height classes**		
	Short (<15 m)	Intermediate (15 < 20 m)	Tall (20+ m)
x	-0.0101***	-0.0075***	-0.0026***
y	0.00761***	-0.0028***	-0.0015***
$I(x^2)$	-0.00003***	0.0000	-0.0000004
I(x*y)	-0.00007***	-0.000001*	-0.000005*
$I(y^2)$	-0.00002*	0.00001**	0.000002
X1 (Elevation)	0.0258***	-0.0595***	-0.0398***
X2 (Slope)	-0.0116*	0.0092	0.0018
X3 (Aspect_cos)	-0.1105	-0.6125***	-0.0506
X4 (Aspect_sin)	-0.2408**	-0.2336**	-0.3118***
X52 (Upper Slope)	-1.1317***	-0.1823	-0.1645
X53 (Middle Slope)	-2.5183***	-1.0930***	0.0699
X54 (Flat Area)	-2.5396***	-1.2516*	0.1893
X55 (Lower Slope)	-3.2513***	-1.7010***	0.1722
X56 (Valley)	-3.1043***	-1.4789***	-0.1416

These modeled patterns are visually supported by the spatial overlay of pine on TPI by height classes (Figure 3.4). Shorter stature pines were predominantly located on ridges (Figure 3.4a), intermediate height pine trees were more dispersed but still frequent on ridges and upper slopes (Figure 3.4b), while the tallest pines were less related to particular topographic positions (Figure 3.4c).

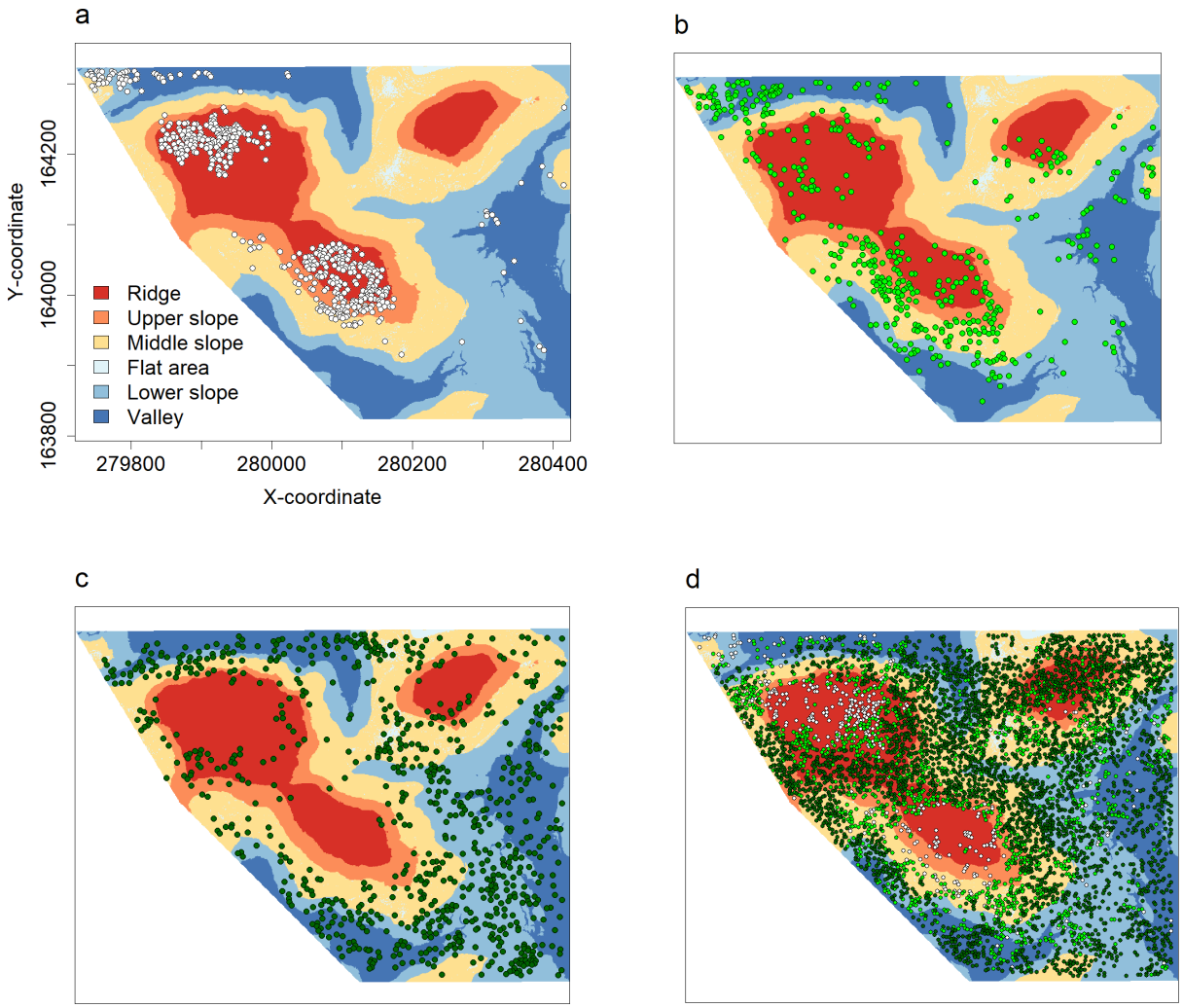

Figure 3.4: Pine and non-pine species distribution across TPI by height class. Plot (a)-(c) show pine distribution for small, medium, and large height classes respectively, while plot (d) shows the combined distribution of non-pine species.

3.4.3 Spatial Pattern Assessment

3.4.3.1 Overall Pine Distribution

At the smallest scale (0-5 m), pine had a uniform distribution (Figure 3.5), though at intermediate distances (10-70 m), pine was more aggregated. In contrast, inhomogeneous pair correlation function shows clustering of pine species at the smallest scale ($r < 20$) beyond what is explained by the environmental covariates. For intermediate to larger distances ($r > 20$), the observed g_{inhom} closely follows the expected g_{inhom} and remains within the simulation enveloped suggesting a random distribution. Additionally, no evidence of significant dispersion is found at any distance since the observed g_{inhom} does not fall significantly below the expected g_{inhom}, at any distance.

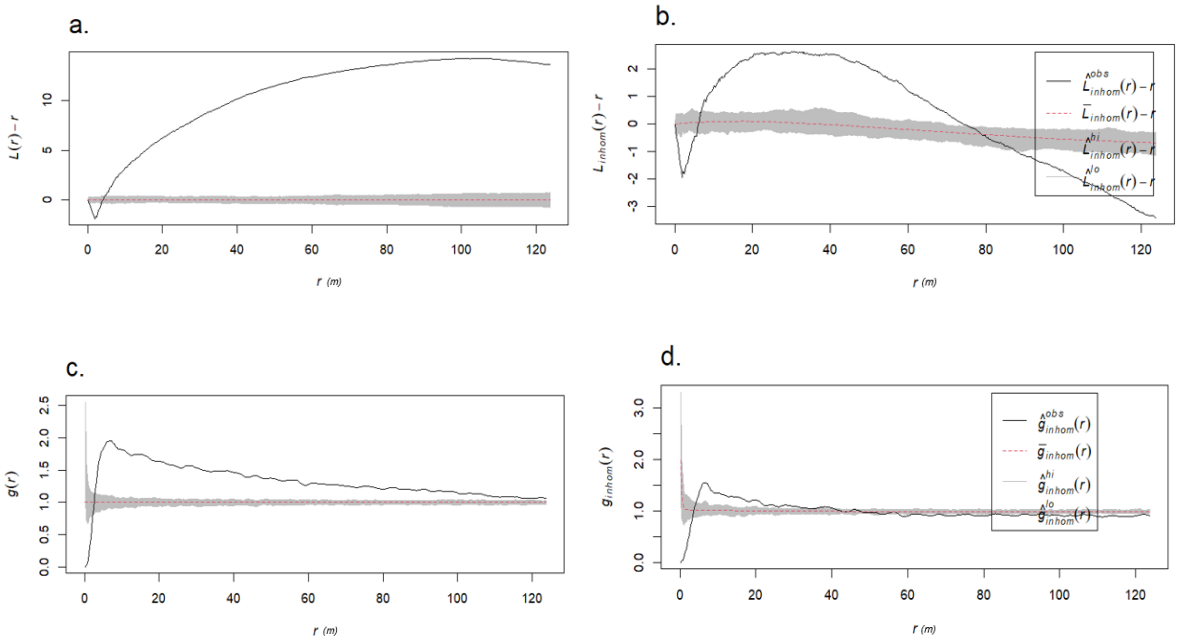

Figure 3.5: Comparison of homogeneous and inhomogeneous spatial point pattern analyses for pine distribution. Panels (a) and (b) show the homogeneous and inhomogeneous Ripley's L-functions, respectively, while panels (c) and (d) show the homogeneous and inhomogeneous pair correlation functions. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates clustering while below the envelope indicates uniform distribution.

3.4.3.2 Pine Distribution by Height Class

The inhomogeneous pair correlation function was calculated for each height class, revealing clustering at shorter distance across all height class (Figure 3.6). Shorter pine exhibited clustering at very short distances (r < 10) (Figure 3.6a). Both the intermediate and tallest trees were clustered at distances ranging from 5 to 25 m. These patterns suggest varying spatial distribution among the height classes, with the shortest trees showing a stronger tendency for proximity compared to intermediate and taller trees though no height classes exhibit uniformity.

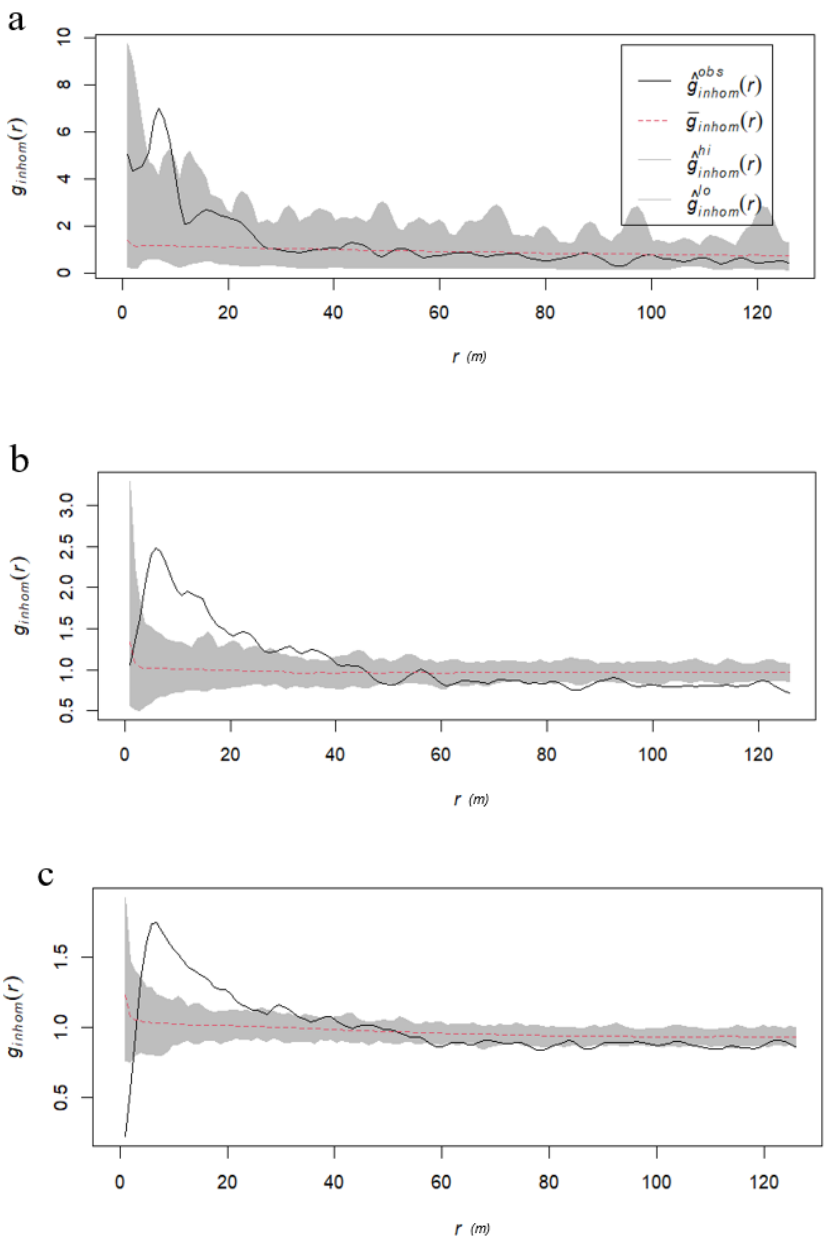

Figure 3.6: Spatial distribution of pine trees by height class: a) short, b) intermediate, and c) tall using pair correlation function. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates clustering while below the envelope indicates uniform distribution.

3.4.3.3 Localized Clustering of Pines

The spatial scan statistic identified significant clusters of pine trees across the study area (Figure 3.7). For the overall pine distribution (Figure 3.7a), two primary hotspots of clustering were identified. These clusters represent the clusters with a significantly higher number of pines than expected under the fitted IPP model. When analyzed by height class, the spatial clustering varied. Shorter pines showed strongly defined clusters particularly concentrated on the ridges (Figure 3.7 b). Intermediate pines were clustered closer to the shorter pine and more concentrated in two areas indicated by the bright yellow color (Figure 3.7 c). The tall pines however, were widely distributed with numerous smaller clusters distributed across the site with one area of higher concentration suggesting more frequent but less dense aggregations compared to the shorter and intermediate height classes (Figure 3.7 d).

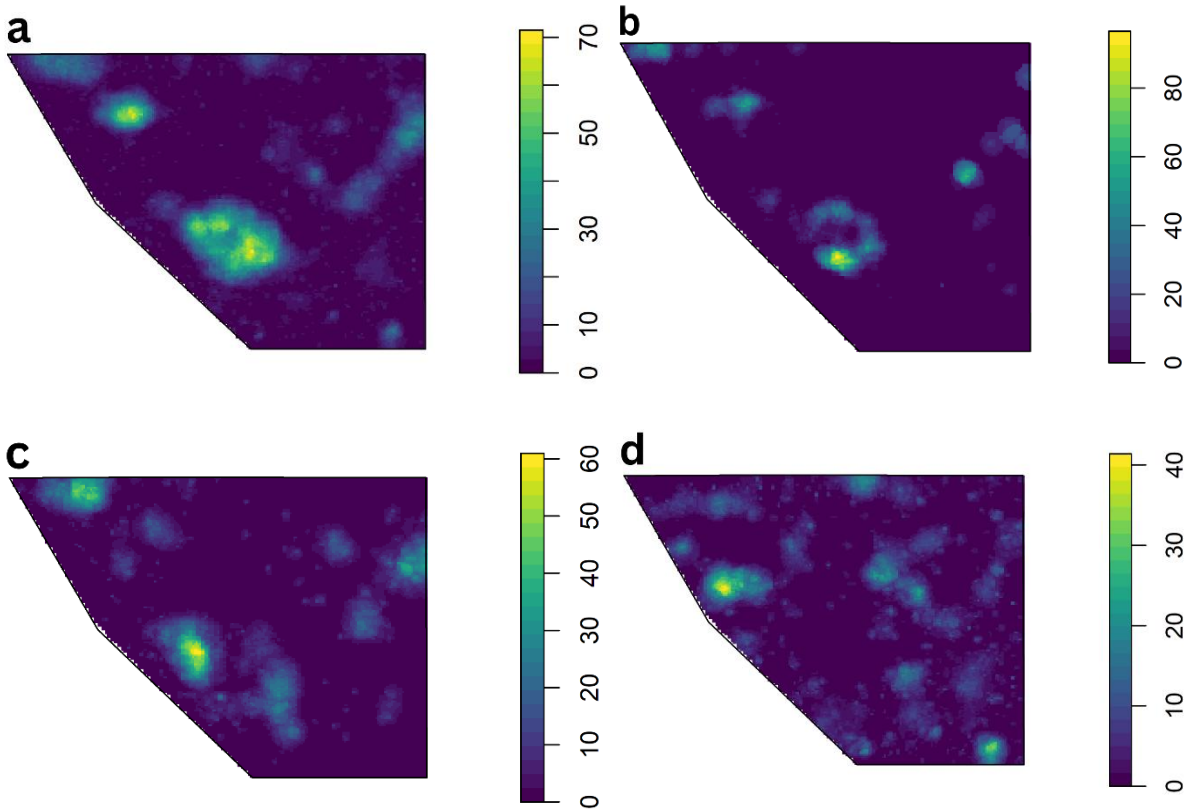

Figure 3.7: Localized clusters of pine trees identified by the spatial scan statistic. a) All pines combined, b) short pine trees c) intermediate pine trees, and d) tall pine trees. Colors represent the magnitude of test statistic with yellow indicating stronger clustering intensity.

3.4.3.4 Spatial Associations Between Height Classes

Distinct spatial relationships among different tree height classes were evident across multiple spatial scales indicated (Figure 3.8). The short and intermediate pine demonstrated a statistically significant positive spatial association (g_inhom > 1) at scales of 15-20 m, indicating aggregation at this specific distance range. This relationship shifted markedly at larger scales (> 40 m), where these two height classes exhibited significant repulsion. In contrast, small and large trees displayed consistent negative spatial association (g_inhom < 1) across all examined spatial scales. For medium and large trees, the analysis revealed a random spatial relationship (g_inhom ≈ 1, within Monte Carlo simulation envelopes) at smaller scales (< 20 m), which transitioned to

significant negative association at scales exceeding 20 m. Visually these patterns are apparent in the mapped distribution, where short and intermediate height trees appear near one another, while the tallest trees are more dispersed throughout the area (Figure 3.9).

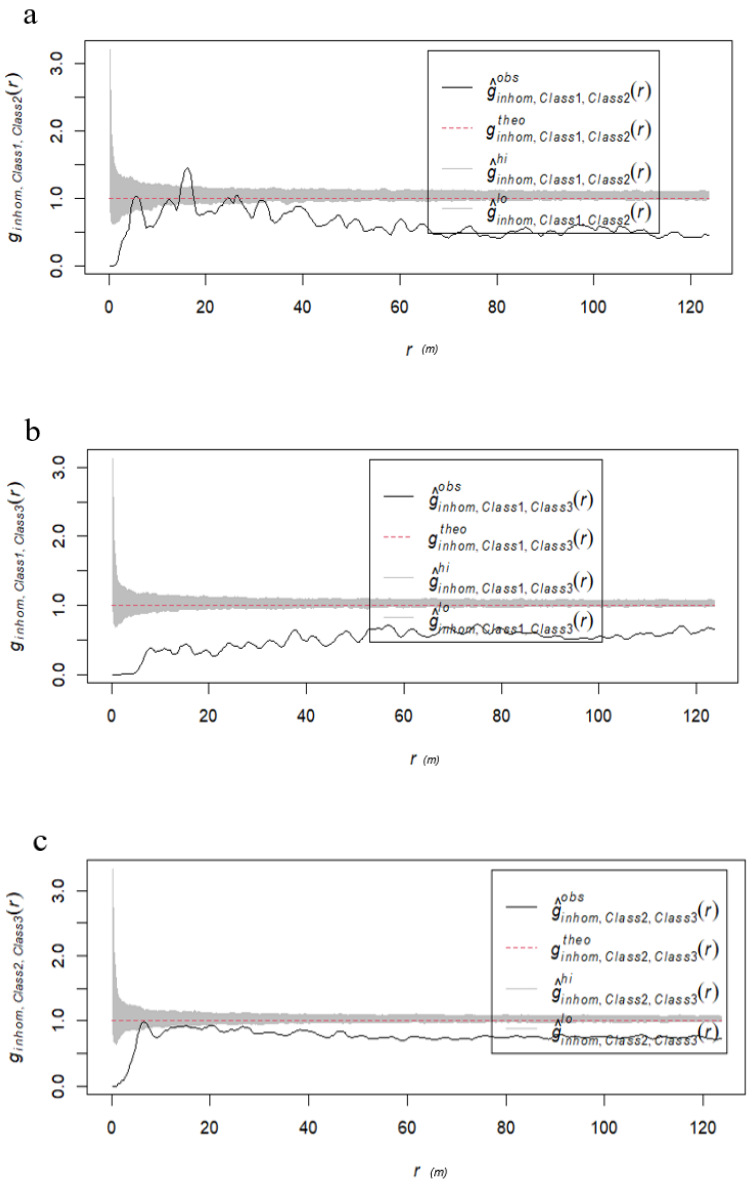

Figure 3.8: Bivariate spatial associations between pine height classes based on inhomogeneous pair correlation functions. Class 1 represent short pine trees, class 2 represent intermediate pine trees, and class 3 represent tall pine trees. The top panel shows the association between short and intermediate trees, the middle panel shows short and tall trees, and the bottom panel shows intermediate and tall trees. The gray shaded area represents confidence envelopes generated from 199 Monte Carlo simulations of the null model. If the black solid line is above the envelope, it indicates spatial attraction, while below the envelope indicates repulsion across spatial scales.

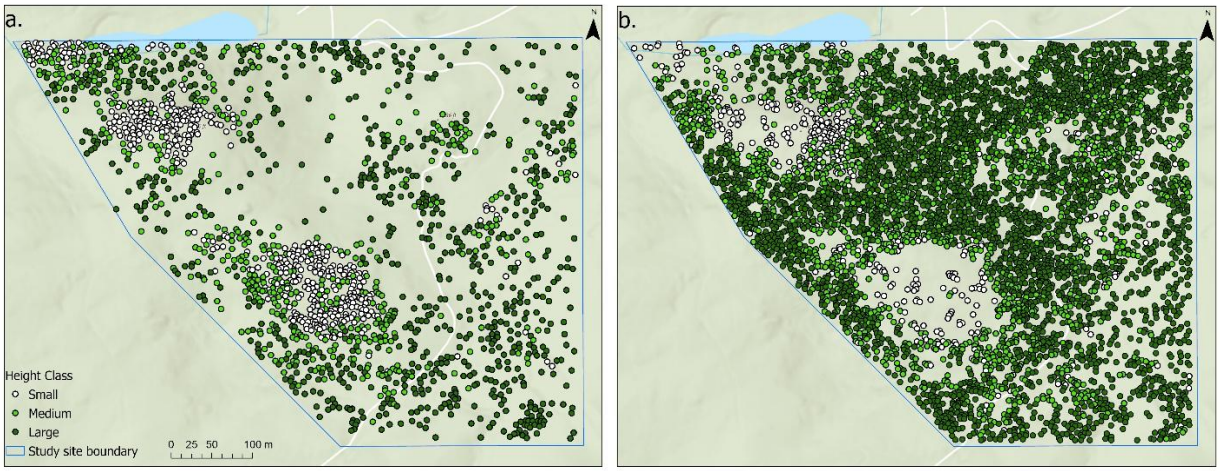

Figure 3.9: Visualization of a) pine and b) non-pine trees distribution by height class.

3.5 Discussion

In the southern Appalachian region, pine species including Eastern white pine, pitch pine, Table Mountain pine, Virginia pine, and shortleaf pine co-occur regionally but are typically associated with distinct geographic and environmental settings (Williams, 1998). Eastern white pine is common on moist slopes, coves and valleys at lower elevation within the study area in DuPont State Forest. Pitch pine occurs on mid-elevation slopes and ridges while shortleaf pine are primarily found on low elevation mountain slopes, ridges, and plateaus (Whittaker, 1956). Consistent with these distributional descriptions, our results supported that pine distributions are influenced by a complex interplay of environmental factors including elevation, slope, aspect, and TPI. A negative association between pine and elevation suggests preference for lower elevation ridges with well-drained soils, consistent with previous research in the region (Harrod et al., 1998). Increasing pine occurrences with steeper slopes likely reflects both reduced competition from mesophytic hardwoods and historical fire patterns that favored pine establishment in dry upland environments (Lafon et al., 2007). Additionally, the negative

estimates of the sine and cosine aspects confirm that pine occur more frequently on south and west facing slopes (Lafon et al., 2019). Furthermore, the lower pine occurrence in all landform categories compared to ridges supports the idea that xeric site conditions are critical for pine survival in this landscape.

Distributional patterns varied by height class. Short statured pine trees were significantly concentrated on the ridges (Figure 3.4a), characterized by very gravelly sandy loam, steep slopes (30 to 95%), and excessively drained soil (*Web Soil Survey*, n.d.). These conditions align with xeric environments, where pines tend to invest more in belowground biomass to support establishment and survival at the expense of aboveground growth (Comeau & Kimmins, 1989). Compared to mesic sites, pine on xeric sites had significantly lower aboveground biomass (116.5 vs 313.1 Mg/ha) and total new production (7.9 vs 11.9 Mg/ha/yr), but allocated a higher proportion of production to belowground biomass (55-66% vs 38-46%), with fine and small roots making up 4% of total biomass on xeric sites compared to only 1.5% on mesic sites (Comeau & Kimmins, 1989). Thus, the shorter height of these pines may reflect site-imposed growth limitations and competition for those limited resources. While intermediate pine trees also occurred on ridges, their distribution was not significantly different on upper slope (Figure 3.4b). Large pine trees were more evenly distributed across all topographic positions, reflecting their ability to grow on less xeric microsite conditions despite facing increased competition (Figure 3.4c). This pattern may also reflect the influence of historical fire activity that extended beyond drought-prone ridgetops into areas dominated by mesic hardwoods, creating canopy openings and favorable regeneration condition for pine (Fesenmyer & Christensen, 2010).

Understanding how plants are distributed across different environments is crucial as it provides valuable insights into stand history, ecological dynamic and interaction between plants.

In our study, the shift from large-scale to small-scale clustering in pine distributions after accounting for environmental heterogeneity suggests that observed spatial patterns are driven by multiple processes operating at different spatial scales. Initially, under CSR, large-scale clustering reflected the influence of environmental gradients, which aggregated pine trees in favorable sites (Figure 3.5). However, once these abiotic factors were removed, the residual spatial pattern revealed small-scale clustering. Such clustering has also been observed in old-growth ponderosa pine forest where upper canopy trees were clustered at scale >2 m (Youngblood et al., 2004). In a mixed oak-shortleaf pine forest, the distribution of shortleaf pine was significantly clustered at stand scale but random at neighborhood scale (Goode et al., 2021). The stand scale clustering, however, was the result of stand-wide disturbance and fire helping in shortleaf pine establishment and recruitment. Additionally, localized biological processes, such as limited seed dispersal or facilitative interactions among neighboring trees may also result in small scale clustering. For instance, although pine species like pitch pine possess winged seeds adapted for wind dispersal, Fowells (1965) reported their spread to be limited with natural regeneration occurring within just 300 feet of parent trees. Additionally, species relying on wind for seed dispersal usually show more compact spatial patterns than animal-dispersed species (Seidler & Plotkin, 2006).

Bivariate analysis revealed strong negative correlations between small and large pine trees across all scales, suggesting competition for resources among mature individuals. In contrast, the positive correlation between small and medium pine trees at small scales indicates potential facilitative interactions or shared microhabitat preferences. This scale-dependent pattern suggests that competitive interactions between these larger tree classes intensify with increasing distance, potentially reflecting crown competition dynamics or resource partitioning

strategies. These findings support that environmental heterogeneity not only shapes overall spatial patterns but also mediates spatial correlations between different canopy strata (Liu et al., 2021). As individual trees cluster in favorable habitats, environmental filtering likely drives aggregated distribution patterns (Getzin et al., 2008), while tree-tree interactions further refine local spatial structures.

3.6 Conclusion

Field-based and labor intensive sampling methods restricted the datasets available to test spatial patterns at multiple scales and orders. The advancement of remote sensing and statistical tools offers new opportunities to confirm and refine patterns. Our study mapped individual pine locations, classified them by height class, and analyzed their spatial pattern, quantifying the topographical factors shaping pine distribution in the Southern Appalachian region. Our findings indicate that environmental variables significantly influence pine distribution to various degrees and the effects change with height class. Understanding the spatial dynamics of pine populations and the relative influence of environmental conditions aids in predicting forest responses to environmental changes and informs strategies to maintain and restore declining pine ecosystem. While this study focused on defining the spatial arrangement of pine and the relationship of topographical variables in the distribution of pine species, future studies incorporating additional covariates such as land-use changes, and past disturbance history could provide a more comprehensive understanding of pine distribution dynamics. Give that land-use alterations have significantly impacted the oak-pine communities in southern Appalachian region including these

variables may enhance predictions of pine occurrence and provide deeper ecological insights into the processes influencing pine distribution, recruitment and landscape level dynamics.

References

Addington, R. N., Aplet, G. H., Battaglia, M. A., Briggs, J. S., Brown, P. M., Cheng, A. S., Dickinson, Y., Feinstein, J. A., Pelz, K. A., Regan, C. M., Thinnes, J., Truex, R., Fornwalt, P. J., Gannon, B., Julian, C. W., Underhill, J. L., & Wolk, B. (2018). *Principles and practices for the restoration of ponderosa pine and dry mixed-conifer forests of the Colorado Front Range* (RMRS-GTR-373; p. RMRS-GTR-373). U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/RMRS-GTR-373

Baddeley, A., Turner, R., & Rubak, E. (2022). *spatstat.explore: Exploratory Data Analysis for the "spatstat" Family* (p. 3.4-2) [Dataset]. https://doi.org/10.32614/CRAN.package.spatstat.explore

Baddeley, A., Turner, R., & Rubak, E. (2025). *spatstat: Spatial Point Pattern Analysis, Model-Fitting, Simulation, Tests* (Version 3.3-2) [Computer software]. https://cran.r-project.org/web/packages/spatstat/index.html

Battaglia, M. A., Mou, P., Palik, B., & Mitchell, R. J. (2002). The effect of spatially variable overstory on the understory light environment of an open-canopied longleaf pine forest. *Canadian Journal of Forest Research*, *32*(11), 1984–1991. https://doi.org/10.1139/x02-087

Cannon, J. B., Bigelow, S. W., Hiers, J. K., & Jack, S. B. (2022). Effects of silvicultural selection treatments on spatial pattern and dynamics in a Pinus palustris Mill. Woodland. *Forest Ecology and Management*, *505*, 119888. https://doi.org/10.1016/j.foreco.2021.119888

Churchill, D. J., Larson, A. J., Dahlgreen, M. C., Franklin, J. F., Hessburg, P. F., & Lutz, J. A. (2013). Restoring forest resilience: From reference spatial patterns to silvicultural prescriptions and monitoring. *Forest Ecology and Management*, *291*, 442–457. https://doi.org/10.1016/j.foreco.2012.11.007

Comeau, P. G., & Kimmins, J. P. (1989). Above- and below-ground biomass and production of lodgepole pine on sites with differing soil moisture regimes. *Canadian Journal of Forest Research*, *19*(4), 447–454. https://doi.org/10.1139/x89-070

Elliott, K. J., L. Hendrick, R., Major, A. E., Vose, J. M., & Swank, W. T. (1999). Vegetation dynamics after a prescribed fire in the southern Appalachians. *Forest Ecology and Management*, *114*(2–3), 199–213. https://doi.org/10.1016/S0378-1127(98)00351-X

Evans, J. S., Murphy, M. A., & Ram, K. (2023). *spatialEco: Spatial Analysis and Modelling Utilities* (Version 2.0-2) [Computer software]. https://cran.r-project.org/web/packages/spatialEco/index.html

Fahey, R. T., & Lorimer, C. G. (2013). Restoring a midtolerant pine species as a component of late-successional forests: Results of gap-based planting trials. *Forest Ecology and Management*, *292*, 139–149. https://doi.org/10.1016/j.foreco.2012.12.026

Fesenmyer, K. A., & Christensen, N. L. (2010). Reconstructing Holocene fire history in a southern Appalachian forest using soil charcoal. *Ecology*, *91*(3), 662–670. https://doi.org/10.1890/09-0230.1

Fowells, H. A. (1965). *Silvics of forest trees of the United States | Fire Research and Management Exchange System*.

Franklin, J. F., Spies, T. A., Pelt, R. V., Carey, A. B., Thornburgh, D. A., Berg, D. R., Lindenmayer, D. B., Harmon, M. E., Keeton, W. S., Shaw, D. C., Bible, K., & Chen, J. (2002). Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *Forest Ecology and Management*, *155*(1–3), 399–423. https://doi.org/10.1016/S0378-1127(01)00575-8

Getzin, S., Wiegand, T., Wiegand, K., & He, F. (2008). Heterogeneity influences spatial patterns and demographics in forest stands. *Journal of Ecology*, *96*(4), 807–820. https://doi.org/10.1111/j.1365-2745.2008.01377.x

Goode, J. D., Hart, J. L., Dey, D. C., Torreano, S. J., & Clark, S. L. (2021). Spatial Patterns of Canopy Disturbance and Shortleaf Pine in a Mixedwood Forest. *Forest Science*, *67*(4), 433–445. https://doi.org/10.1093/forsci/fxab017

Guo, Z., Li, Y., & Gan, Y. (2003). Spatial Pattern of Ecosystem Function and Ecosystem Conservation. *Environmental Management*, *32*(6), 682–692. https://doi.org/10.1007/s00267-003-2911-z

Harrod, J., White, P. S., & work(s):, M. E. H. R. (1998). Changes in Xeric Forests in Western Great Smoky Mountains National Park, 1936-1995. *Castanea*, *63*(3), 346–360.

Hessburg, P. F., Smith, B. G., & Salter, R. B. (1999). DETECTING CHANGE IN FOREST SPATIAL PATTERNS FROM REFERENCE CONDITIONS. *Ecological Applications*, *9*(4), 1232–1252. https://doi.org/10.1890/1051-0761(1999)009[1232:DCIFSP]2.0.CO;2

Hijmans, R. J. (2020). *terra: Spatial Data Analysis* (p. 1.8-15) [Dataset]. https://doi.org/10.32614/CRAN.package.terra

Kenkel, N. C. (1988). Pattern of Self-Thinning in Jack Pine: Testing the Random Mortality Hypothesis. *Ecology*, *69*(4), 1017–1024. https://doi.org/10.2307/1941257

Kobziar, L., Godwin, D., Taylor, L., & Watts, A. (2015). Perspectives on Trends, Effectiveness, and Impediments to Prescribed Burning in the Southern U.S. *Forests*, *6*(3), 561–580. https://doi.org/10.3390/f6030561

Lafon, C. W., DeWeese, G. G., Flatley, W. T., Aldrich, S. R., & Naito, A. T. (2022). Historical Fire Regimes and Stand Dynamics of Xerophytic Pine–Oak Stands in the Southern

Appalachian Mountains, Virginia, USA. *Annals of the American Association of Geographers*, *112*(2), 387–409. https://doi.org/10.1080/24694452.2021.1935206

Lafon, C. W., Hanson, A. A., & Dwight, R. A. (2019). Geographic variations in fine-scale vegetation patterns: Aspect preferences of montane pine stands over Southern Appalachian landscapes. *Physical Geography*, *40*(5), 433–460. https://doi.org/10.1080/02723646.2019.1576013

Lafon, C. W., & Kutac, M. J. (2003). Effects of Ice Storms, Southern Pine Beetle Infestation, and Fire on Table Mountain Pine Forests of Southwestern Virginia. *Physical Geography*, *24*(6), 502–519. https://doi.org/10.2747/0272-3646.24.6.502

Lafon, C. W., Waldron, J. D., Cairns, D. M., Tchakerian, M. D., Coulson, R. N., & Klepzig, K. D. (2007). Modeling the Effects of Fire on the Long-Term Dynamics and Restoration of Yellow Pine and Oak Forests in the Southern Appalachian Mountains. *Restoration Ecology*, *15*(3), 400–411. https://doi.org/10.1111/j.1526-100X.2007.00236.x

Larson, A. J., & Churchill, D. (2012). Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. *Forest Ecology and Management*, *267*, 74–92. https://doi.org/10.1016/j.foreco.2011.11.038

Lindenmayer, D., Hobbs, R. J., Montague-Drake, R., Alexandra, J., Bennett, A., Burgman, M., Cale, P., Calhoun, A., Cramer, V., Cullen, P., Driscoll, D., Fahrig, L., Fischer, J., Franklin, J., Haila, Y., Hunter, M., Gibbons, P., Lake, S., Luck, G., … Zavaleta, E. (2008). A checklist for ecological management of landscapes for conservation. *Ecology Letters*, *11*(1), 78–91. https://doi.org/10.1111/j.1461-0248.2007.01114.x

Liu, J., Bai, X., Yin, Y., Wang, W., Li, Z., & Ma, P. (2021). Spatial patterns and associations of tree species at different developmental stages in a montane secondary temperate forest of northeastern China. *PeerJ*, *9*, e11517. https://doi.org/10.7717/peerj.11517

Lydersen, J. M., North, M. P., Knapp, E. E., & Collins, B. M. (2013). Quantifying spatial patterns of tree groups and gaps in mixed-conifer forests: Reference conditions and long-term changes following fire suppression and logging. *Forest Ecology and Management*, *304*, 370–382. https://doi.org/10.1016/j.foreco.2013.05.023

Maestre, F. T., Castillo-Monroy, A. P., Bowker, M. A., & Ochoa-Hueso, R. (2012). Species richness effects on ecosystem multifunctionality depend on evenness, composition and spatial pattern. *Journal of Ecology*, *100*(2), 317–330. https://doi.org/10.1111/j.1365-2745.2011.01918.x

Matthews, J. M., Hinchee, J., & Guldin, J. M. (2020). Restoration of longleaf pine in the southern region of the U.S. Forest Service: An overview of the million-acre challenge. *In Proceedings of the 20th Biennial Southern Silvicultural Research Conference*.

Morrow, R. J. (1985). *Age structure and spatial pattern of old-growth ponderosa pine in Pringle Falls Experimental Forest, central Oregon*. Oregon State University.

Ostapowicz, K., Vogt, P., Riitters, K. H., Kozak, J., & Estreguil, C. (2008). Impact of scale on morphological spatial pattern of forest. *Landscape Ecology*, *23*(9), 1107–1117. https://doi.org/10.1007/s10980-008-9271-2

Peet, R. K., & Christensen, N. L. (1987). Competition and Tree Death. *BioScience*, *37*(8), 586–595. https://doi.org/10.2307/1310669

R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Rodman, K. C., Sánchez Meador, A. J., Moore, M. M., & Huffman, D. W. (2017). Reference conditions are influenced by the physical template and vary by forest type: A synthesis of Pinus ponderosa-dominated sites in the southwestern United States. *Forest Ecology and Management*, *404*, 316–329. https://doi.org/10.1016/j.foreco.2017.09.012

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., Boissieu, F. de, & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, *251*. https://doi.org/10.1016/j.rse.2020.112061

Schwartz, N. B., Urban, D. L., White, P. S., Moody, A., & Klein, R. N. (2016). Vegetation dynamics vary across topographic and fire severity gradients following prescribed burning in Great Smoky Mountains National Park. *Forest Ecology and Management*, *365*, 1–11. https://doi.org/10.1016/j.foreco.2016.01.027

Seidler, T. G., & Plotkin, J. B. (2006). Seed Dispersal and Spatial Pattern in Tropical Trees. *PLoS Biology*, *4*(11), e344. https://doi.org/10.1371/journal.pbio.0040344

Shive, K. L., Preisler, H. K., Welch, K. R., Safford, H. D., Butz, R. J., O'Hara, K. L., & Stephens, S. L. (2018). From the stand scale to the landscape scale: Predicting the spatial patterns of forest regeneration after disturbance. *Ecological Applications*, *28*(6), 1626–1639. https://doi.org/10.1002/eap.1756

Simioni, G., Gignoux, J., & Le Roux, X. (2003). TREE LAYER SPATIAL STRUCTURE CAN AFFECT SAVANNA PRODUCTION AND WATER BUDGET: RESULTS OF A 3-D MODEL. *Ecology*, *84*(7), 1879–1894. https://doi.org/10.1890/0012-9658(2003)084[1879:TLSSCA]2.0.CO;2

Szwagrzyk, J., & Czerwczak, M. (1993). Spatial patterns of trees in natural forests of East-Central Europe. *Journal of Vegetation Science*, *4*(4), 469–476. https://doi.org/10.2307/3236074

Turner, D. P., & Franz, E. H. (1986). The Influence of Canopy Dominants on Understory Vegetation Patterns in an Old-Growth Cedar-Hemlock Forest. *American Midland Naturalist*, *116*(2), 387. https://doi.org/10.2307/2425747

Turner, M. G., Pearson, S. M., Bolstad, P., & Wear, D. N. (2003). Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). *Landscape Ecology*, *18*(5), 449–464. https://doi.org/10.1023/A:1026033116193

U.S. Geological Survey. (2024, August 28). https://www.usgs.gov/tools/lidarexplorer

Vose, J. M., CIinton, B. D., & Swank, W. T. (1993). *FIRE, DROUGHT, AIND FOREST hlANAGEME,NT IINFLUENCES ON PINE/HARDWOOD ECOSYSTEMS IN THE SOUTHERN APPALA'CHIANSL*.

Waldron, J. D., Lafon, C. W., Coulson, R. N., Cairns, D. M., Tchakerian, M. D., Birt, A., & Klepzig, K. D. (2007). Simulating the impacts of southern pine beetle and fire on the dynamics of xerophytic pine landscapes in the southern Appalachians. *Applied Vegetation Science*, *10*(1), 53–64. https://doi.org/10.1111/j.1654-109X.2007.tb00503.x

Wang, X., Zheng, G., Yun, Z., & Moskal, L. M. (2020). Characterizing Tree Spatial Distribution Patterns Using Discrete Aerial Lidar Data. *Remote Sensing*, *12*(4), 712. https://doi.org/10.3390/rs12040712

Web Soil Survey. (n.d.). Retrieved January 24, 2025, from https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx

Weiss, A. D. (2001). *Topographic position and landforms analysis*. https://env761.github.io/assets/files/tpi-poster-tnc_18x22.pdf

White, P. S., & Walker, J. L. (1997). Approximating Nature's Variation: Selecting and Using Reference Information in Restoration Ecology. *Restoration Ecology*, *5*(4), 338–349. https://doi.org/10.1046/j.1526-100X.1997.00547.x

Whittaker, R. H. (1956). Vegetation of the Great Smoky Mountains. *Ecological Monographs*, *26*(1), 1–80. https://doi.org/10.2307/1943577

Williams, C. E. (1998). History and Status of Table Mountain Pine–Pitch Pine Forests of the Southern Appalachian Mountains (USA). *Natural Areas Association*, *18*(1), 81–90.

Wolf, A. (2005). Fifty year record of change in tree spatial patterns within a mixed deciduous forest. *Forest Ecology and Management*, *215*(1–3), 212–223. https://doi.org/10.1016/j.foreco.2005.05.021

Yarnell, S. L. (1998). *The Southern Appalachians: A History of the Landscape*.

Youngblood, A., Max, T., & Coe, K. (2004). Stand structure in eastside old-growth ponderosa pine forests of Oregon and northern California. *Forest Ecology and Management*, *199*(2–3), 191–217. https://doi.org/10.1016/j.foreco.2004.05.056

CHAPTER 4 - Establishment pattern, disturbance history, and factors affecting yellow pine regeneration patterns in southern pine-oak mixedwoods

4.1 Abstract

Understory tree regeneration is crucial for the sustainability of southern pine-oak mixedwood forest, yet recent decades have experienced a significant decline in pine recruitment due to altered disturbance regimes and increased competition from shade-tolerant species. This study investigated the factors influencing pine regeneration in a 26 ha pine-oak mixedwood stand by analyzing the overstory and understory composition, soil and topographic variables, and disturbance history. Our results revealed that while shortleaf, Eastern white, and pitch pines dominated the overstory, the understory was composed primarily of mountain laurel and other hardwoods with minimal pine regeneration. Non-metric multidimensional scaling ordination indicated that overstory composition was primarily structured by abiotic factors such as slope and aspect, whereas understory regeneration patterns were strongly linked to canopy openness and shrub cover. Temporal analysis of tree establishment and growth release events highlighted a peak in pine recruitment during the 1950s, coinciding with disturbance events, followed by a shift toward shade-tolerant species as disturbance frequency declined. These findings highlight the importance of disturbance, particularly fire and canopy gap creation, in maintaining pine populations and inform management strategies aimed at restoring pine regeneration and sustaining mixedwood forest ecosystems in the southern Appalachians.

4.2 Introduction

Disturbances whether they be windstorms, fire, climatic events, species loss or human land-use, have a legacy effect on the composition and function of the contemporary forests. Studying past disturbances provides a long-term context for present-day conditions and helps determine whether current changes fall within natural ecosystem dynamics or represent something novel (Turner et al., 2003). Disturbances are key drivers of forest dynamics integral to succession as they create space and alter resource availability to allow for the regeneration and establishment of new species or individuals (Seidl et al., 2014). Stand-replacing disturbance events, significantly impact the forest age structure and creates spatial landscape heterogeneity of different habitats and successional stages (Pickett & White, 1985). Knowledge of historical disturbance regimes is important to developing forest management and restoration strategies. Management activities are often designed with the historical natural disturbance regime to maintain ecosystem vigor and diversity.

In the southern United States, pine (*Pinus* spp.) - oak (*Quercus* spp.) mixedwoods (where pine comprise 25-50%) consists of 9% of the total forest area (Oswalt et al., 2019). Predominantly located on xeric slopes and ridges at mid-elevations in the southern Appalachian Mountains (Lafon et al., 2007), these mixedwoods thrive in dry, nutrient-poor conditions (Waldron et al., 2007). Xeric sites characterized by lower site moisture, shallower soil depth, and lower nutrient levels lead to the selection of xeric pines over their mesic counterparts. However, pine forests in these mixedwoods have experienced significant decline due to the combined effects of fire suppression and extensive overstory mortality caused by periodic southern pine beetle (*Dendroctonus frontalis*) outbreaks (Jenkins et al., 2011). Additionally, the southern Appalachian Mountains have a long history of disturbance including insect infestation (Khodaee

et al., 2020), mountain top mining (Bernhardt & Palmer, 2011; Wickham et al., 2007), wildfires (Caldwell et al., 2020; Hagan et al., 2015; Reilly et al., 2006), logging (Benfield et al., 1991), and livestock grazing (Pyle 1988). These disturbances have significantly influenced the structure, function, composition, and diversity of the forest ecosystems in these areas (Khodaee et al., 2020; Waldron et al., 2007).

Historically, fire played a pivotal role in maintaining the abundance of pine species by creating favorable conditions for pine regeneration (Gilliam & Platt, 1999). However, years of fire suppression coupled with landscape fragmentation and insect infestation have led to an increase of overstory density, proliferation of less fire-tolerant hardwood species in the understory, and accumulation of heavy leaf litter, all of which have obstructed the regeneration of pine species (Harrod et al., 1998; Williams, 1998) placing them at risk of being encroached by shade tolerant species (Dey et al., 2019). For example, long-term absence of fire has led to the dominance of mountain laurel (Kalmia latifolia) and Rhododendron spp. in the understories of pine-oak forests and abundance of these species creates low light conditions in the understory thereby inhibiting the regeneration of shade intolerant species (Brose & Waldrop, 2006, 2010; Phillips & Murdy, 1985).

Restoring the declining pine species in the southern pine-oak mixedwoods is crucial both ecologically and economically, as these forests offer a range of benefits such as vital food sources, diverse wildlife habitat, timber production, fuelwood, resins, fodder, and biodiversity conservation (Brose & Waldrop, 2010; Kenefic et al., 2021; Martin et al., 2021). Therefore, effective restoration strategies aimed at conserving species diversity, promoting sustainable forest management, and optimizing the use of plant resources requires understanding the factors that affect growth and establishment of regenerating species that ultimately shapes the

regeneration patterns in the understory (Chai & Wang, 2016). Among the various factors influencing patterns, disturbance emerges as the primary driver affecting the spatial pattern of canopy trees in mixedwoods forest (Goode et al., 2021; Schwarz et al., 2003) which in turn plays a crucial role in shaping the understory regeneration pattern. Therefore, quantifying the extent of disturbance is vital from both a management perspective and for forest carbon cycle modeling (Frolking et al., 2009). Post-disturbance, the rate of forest recovery depends on the extent and severity of the disturbance with large disturbances (disturbance that generate gaps larger than 0.1 ha) requiring a longer recovery period (Frolking et al., 2009). Large scale forest restoration relies on natural regeneration as these methods are significantly more cost-effective than artificial regeneration (Chazdon & Guariguata, 2016). However, key ecological processes such as soil carbon and nitrogen cycling that enable natural regeneration are impacted by forest disturbances (Mlambo et al., 2023). Restoration activities are best determined with a clear understanding of the disturbance and establishment of the natural system. In pine-oak mixedwoods this involves reconstructing the historical disturbance regime such as canopy gaps and fires that determined the patterns of stand development and succession (Goode et al., 2021).

The objective of this study was to determine the factors affecting the pine regeneration in a pine-oak mixedwoods stand in the southern Appalachian Mountains. Specifically, the research aims to answer the following questions:

1. Do the establishment patterns of pine relate to contemporary overstory pine distributions?
2. Which environmental factors most strongly influence understory pine regeneration?
3. What proportion of the disturbances are widespread across all stands versus localized to specific areas?

The environmental factors considered in this study include a range of topographic variables (elevation, aspect and slope), soil properties (pH, organic matter, nitrogen, phosphorous and potassium content) (Table S.4.1), forest stand characteristics (shrub and herb cover, overstory and understory canopy cover, hardwood density, and stand density), and past disturbance history. We hypothesized that the variation in stand structure, particularly canopy cover, will influence the pine regeneration more relative to soil and topographic factors. We expect to see higher abundance of understory pine in areas with canopy gaps as these gaps promote light availability, thereby facilitating the growth or survival of shade intolerant species such as pine. We also hypothesized that the variation in topographic variables will impact the understory pine abundance. The results may be used to inform forest management practices aimed at enhancing pine regeneration in the mixedwoods stands. By identifying the key factors that influence pine regeneration, this study can provide guidance to forest managers on where to focus restoration efforts, thereby promoting a more uniform recovery across the landscape.

4.3 Methods

4.3.1 Study Site

Forest vegetation and soil samples were collected within a 26-hectare study area in DuPont State Recreational Forest in Cedar Mountain, North Carolina (Fig. 1). The site lies in the upland plateau of the Little River Valley with elevation ranging from 701 to 1097 meters. From 1981-2022, the mean annual temperature ranged from 12.98 to 15.68°C, mean annual precipitation from 843.75 to 1792.97 mm, and mean relative humidity ranged from 69 to 81.38% (NASA Langley Research Center, 2022). The study area encompasses a diverse range of soil

types characterized by varying slopes, drainage capacities, and compositions. Predominantly, the soils in the study area are well-drained with 8 to 15 percent slopes. Soils present in both moderate and steep slopes (8 to 95 percent), are somewhat excessively drained, and feature rocky and very rocky textures. Additionally, areas with fine sandy loam are notable for its frequent flooding at gentle slopes, 0 to 2 percent (Web Soil Survey, n.d.).

4.3.2 Field Measurements

Fifteen fixed area plots (0.04 ha, 11.35 m radius) were established to quantify species composition and structure, with varied sizes per lifeform. In each plot, the species, Diameter at Breast Height (DBH) and height of all stems ≥10 cm DBH were recorded. All trees ≥ 20 cm and a subset of stems between 5 to 20 cm were cored (two per plot) at 1.37 m height, for a total of 263 cores collected. Cores were stored in paper straws until they could be glued to tree mounts in the lab.

At each plot location, a 0.004 ha circular subplot (3.6 m radius) located 7.6 m north of the plot center was established to measure saplings (woody stems ≥ 1.4 m height and < 10 cm DBH. Species, DBH and height class (defined as 1.4 – 3m, 3-5m, 5-10m and >10 m) was recorded for each individual sapling. Additionally, 4 m from the plot center regeneration subplots (2 m radius) were established in the cardinal directions. In each subplot, we recorded the count of woody stems by species for individuals < 1.4 m tall and the percent cover of shrub species was estimated visually to the nearest 10%. A soil sample was extracted using a soil probe (2.5 cm diameter, 15 cm depth) from each of the regeneration subplots, composited for a plot sample, and submitted to Brookside Laboratories (New Bremen, OH) for standard soil nutrient analysis (pH, soil organic matter, cation exchange capacity, micro and macro-nutrients).

A Terrestrial Laser Scanner (TLS) was used to take a 360 scan at plot center (Faro Focus 3D X330). We used the Faro Scene software to convert the .fls files to .las and join scans of a plot where more than one scan was taken. Most of the time, automatic processing adequately joined multiple scans when there was sufficient overlap between the scans. When the automatic registration failed, we used target-based registration using the targets set on the plots during scan collection. Faro provides a month of free subscription if signed up using a student id and we utilized that to process our scans. After the scans were combined and converted into las format, they were further processed using the lidR package (Roussel et al., 2020) in R programming (as explained in Chapter 3 methods) (R Core Team, 2024). To determine the overstory and understory canopy cover, we used the area-based approach by Vatandaslar et al., (2024). In this approach, the normalized lidar data for each plot was used to create a canopy height model of 0.1 m resolution using the rasterize_canopy function from the lidR package. All pixels with value ≤ 2 are assigned a zero value to represent the ground and to remove the understory. The number of pixels with values equal to 0 are summed and the total number of pixels with value > 0 are summed. Pixels with NA values were not counted. To determine the overstory canopy cover, we divided the number of pixels with 0 value by the number of pixels with value > 0. The process was repeated to estimate the understory canopy cover after filtering the lidar data to contain points with $Z \leq 2$ m only.

4.3.3 Data Analysis

Vegetation measurements were summarized by plot for all lifeforms. The Importance Values for tree species was calculated as an average of the relative density, relative basal area (BA) and frequency. Sapling density and BA were summarized by height classes. The percent

cover of shrubs was averaged across subplots, and counts of woody stems from regeneration subplots were combined and scaled to seedlings per hectare for comparison.

Tree cores were glued to wooden mounts with cells aligned vertically (Stokes & Smiley, 1996). The samples were sanded using progressively finer sandpaper to expose the cellular structure and scanned to create a digital file for each core. Growth increments were measured using CooRecorder software (Cybis Elektronik & Data AB; Maxwell and Larsson, 2021). Records were combined in CDendro and establishment years were calculated based on age estimates calculated for each core. Abrupt and sustained increases in radial growth were identified using the Tree Ring Analysis of Disturbance Events (TRADER) package in R (Altman et al., 2014). The package provided results from several proposed methodologies for defining disturbance events. We used the Nowacki and Abrams method and identified release events as periods in which the raw ring width of a year was >25% (moderate) or >50% (major) of the mean ring width of the preceding and superseding ten years, sustained for at least five years with a 10-year buffer (Nowacki & Abrams, 1997). Stand wide disturbances were identified as release events for a given decade detected in at least 25% of the individuals or a simultaneous release detected in at least 25% of the plots.

Growth increments (mm per year) were averaged by decade to compare whether there were any significant differences in growth across time periods. A one-way ANOVA was used to determine whether the mean growth differed significantly across decades. Where significant differences were found, Tukey's Honest Significant Difference post hoc test was applied to identify which specific decades showed significant pairwise differences in growth.

We explored the overstory and understory species composition across the plots using nonmetric multidimensional scaling (NMS) ordination in PC-ORD (Version 7.07, McCune & Mefford, 1999). NMS ordination was used to visualize the differences in the species composition based on different variables such as soil pH, organic matter content, slope, aspect, topographic position index (TPI), overstory and understory openness, shrub cover, sapling and regeneration density.

4.4 Results

4.4.1 Forest Structure and Composition

Pine accounted for approximately 38% of all trees recorded across the 15 plots, indicating a mixedwood forest composition dominated by hardwood species. Among the pine species, shortleaf pine had the highest basal area (BA) of 11.9 m^2/ha followed by white pine and pitch pine (Table 4.1). Although pitch pine had higher basal area than red maple, red maple had a higher IV indicating that it was more abundant and more frequently distributed across the sampled plots despite having smaller individual tree sizes.

Table 4.1: Tree species level summary of forest composition in a pine-oak mixedwoods stand (n=15) in DuPont State Recreational Forest, NC

Tree species	**BA (m2/ha)**	**Density (stems/ha)**	**Relative density (%)**	**Relative Frequency (%)**	**Relative Dominance (%)**	**IV (%)**
Shortleaf pine	11.9	121.7	16.2	9.4	28.6	18.0
E. white pine	8.4	86.7	11.5	8.5	20.1	13.4
Red maple	3.4	93.3	12.4	11.1	8.1	10.5
Pitch pine	5.0	78.3	10.4	6.8	12.0	9.7
Sourwood	1.8	88.3	11.7	11.1	4.3	9.0
Northern red oak	3.6	71.7	9.5	7.7	8.7	8.6
Scarlet oak	1.9	31.7	4.2	6.8	4.6	5.2
White oak	0.9	35.0	4.6	6.8	2.2	4.6
Chestnut oak	1.2	40.0	5.3	5.1	2.9	4.4
Yellow poplar	0.9	23.3	3.1	5.1	2.3	3.5
Eastern hemlock	0.9	18.3	2.4	3.4	2.1	2.7
Fraser magnolia	0.7	13.3	1.8	4.3	1.6	2.5
Hickory	0.4	13.3	1.8	2.6	1.0	1.8
Black gum	0.3	11.7	1.5	2.6	0.7	1.6
Yellow birch	0.2	10.0	1.3	1.7	0.5	1.2
Serviceberry	0.04	5.0	0.7	2.6	0.1	1.1
Sassafrass	0.04	3.3	0.4	1.7	0.1	0.8

Table 4.1 (continued):

Mountain laurel	0.03	3.3	0.4	1.7	0.1	0.7
Black oak	0.1	5.0	0.7	0.9	0.2	0.6

Shortleaf pine saplings were not present, despite having the highest basal area and stem density among the overstory trees (Table 2). In contrast, mountain laurel had the lowest basal area (0.03 m^2/ha) among trees but was the most dominant species in the understory. It was present in 14 of the 15 sample plots and frequently formed dense thickets, occupying much of the sapling layer. Saplings in the shortest (1.4 - 3 m) height class were the most abundant, with a mean density of 230 stems/ha (Fig. 1). As the height class increased, the sapling density declined with the tallest class (>10 m) having the fewest stems (mean density of 42.6 stems/ha).

Similar to the sapling layer, mountain laurel also had the highest seedling density across all sampled plots (Table 2). Among the pine species, white pine had the highest seedling density of 1100.8 stems/ha. While we didn't find any saplings of shortleaf pine, seedling density was 172.4 stems/ha. Pitch pine regeneration was minimal with only one seedling recorded across all 15 plots.

Table 4.2: Sapling (stems<10 cm DBH) and seedling density in a pine-oak forest in DuPont State Recreational Forest, NC

Species	**Mean Sapling DBH (cm)**	**Sapling Density (stems/ha)**	**Seedling Density (stems/ha)**
Mountain laurel	2.15	4150	16751.06
Sparkleberry	1.03	250	-
Rhododendron	1.87	233.33	344.84
Red maple	4.28	200	7520.07
Sweet leaf	3.23	133.33	848.83
Sourwood	6.13	66.67	79.57
Black gum	5.13	50	66.32
Chestnut oak	3.47	50	106.10
American holly	5.15	33.33	13.26
Black haw	6.85	33.33	-
Fraser magnolia	2.70	33.33	358.10
Serviceberry	4.45	33.33	397.89
E. white pine	8.85	33.33	1100.82
Witch hazel	1.4	33.33	119.37
Carolina hemlock	0.2	16.67	-
Northern red oak	0.4	16.67	1618.08
Sassafras	0.7	16.67	252
White oak	5	16.67	225.47

Table 4.2 (continued):

Yellow birch	7.3	16.67	2559.74
Blueberry	-	-	10716.43
Yellow poplar	-	-	198.94
Shortleaf	-	-	172.42
Hickory	-	-	132.629
Black oak	-	-	92.84
Black cherry	-	-	79.58
Mountain sweet pepper bush	-	-	79.58
Carolina buckthorn	-	-	39.79
Eastern hemlock	-	-	39.79
American chestnut	-	-	26.53
Scarlet oak	-	-	26.53
Dogwood	-	-	13.26
Balsam fir	-	-	13.26
Green ash	-	-	13.26
Pitch pine	-	-	13.26
Southern red oak	-	-	13.26

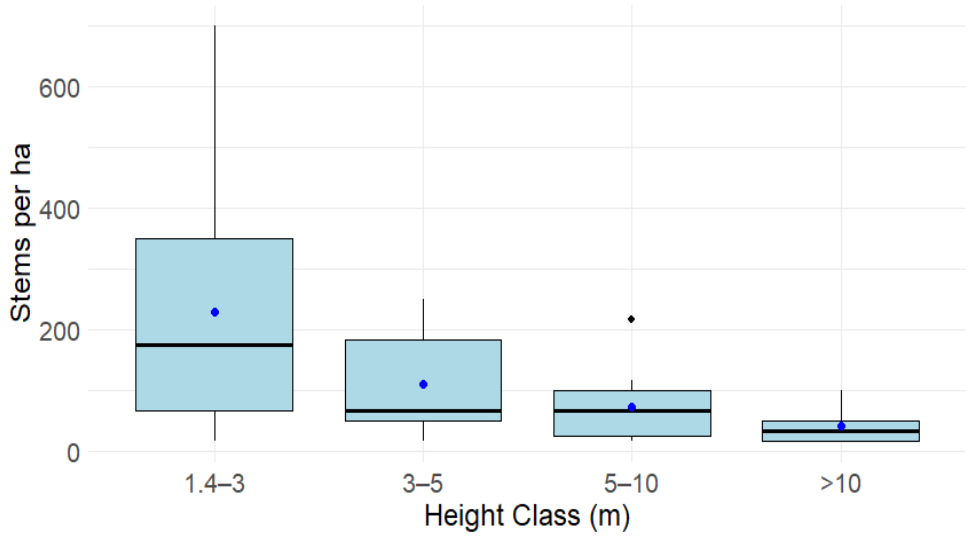

Figure 4.1: Sapling density by height class in a pine-oak mixedwoods stand in DuPont State Recreational Forest. The blue point represents the mean sapling density (n=15).

4.4.2 Environmental Gradients and Plant Communities

Environmental conditions across the 15 samples plots reflected the heterogeneous topography typical of the southern Appalachian landscape. Elevation ranged from 783.72 to 866.25 m, with slope between 4.57 to 17.35 degree. The mean overstory canopy openness percent was 11.87 ± 1.36 while the understory canopy openness percent was 33.53 ± 3.55 (Table S.4.1). Soil pH across plots ranged from 3.9 to 5.0, indicating acidic conditions which are typical to the pine-oak forests in southern Appalachians.

Three overstory plant communities were suggested based on the NMS solution based on the tree IV. The final solution had a three-axis stress of 5.502 and explained 95% of the variability in the overstory data with Axis 1-3 explaining 51.5%, 32.7%, and 10.8%, respectively. Shortleaf pine was in highest abundance with species such as northern red oak,

hickory, sassafrass, and blackgum (Fig. 2) and was negatively correlated with Axis 2 (r^2=0.79). Eastern white pine was most abundant with tulip-poplar, fraser magnolia, yellow birch and Eastern hemlock and positively correlated with Axis 2 (r^2=0.5). Pitch pine and chestnut oak were negatively correlated with Axis 1 (r^2=0.86 and 0.688, respectively).

The overstory composition was strongly influenced by abiotic variables including aspect (Axis 2 r^2=0.383), elevation (Axis 2 r^2=0.313) and several soil nutrients (r^2 >0.2 for nutrients in Fig. 2). The Eastern white pine community was related to lower slope and northern aspect and at relatively lower elevations. The pitch pine community occurred on ridges, while shortleaf pine community was on middle slope, with both at relatively higher elevations.

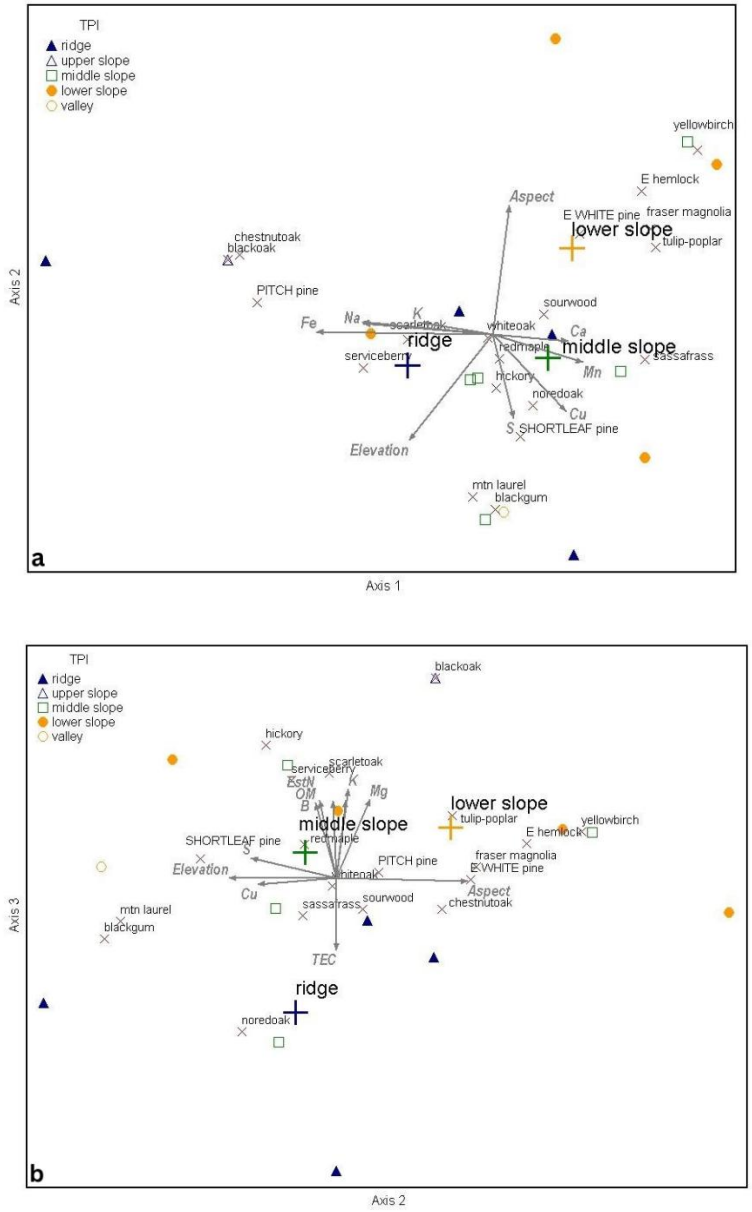

Figure 4.2: Non-metric multidimensional scaling ordination of overstory tree communities based on species importance values. a) Axis 1 explained 51.5% and Axis 2 explained 32.7% of the variability; while b) Axis 3 explained 10.8%.

The NMS ordination for understory regeneration showed strong ecological gradients structuring the species composition explaining 97.6% of the total variance in understory species composition with a 3-axis solution (final stress=1.069). Axis 1 explained 60.5% of the total

variation and was strongly correlated with overstory openness (r2 = 0.302) indicating that gradients in canopy cover and light availability were key drivers of regeneration community composition (Fig. 3). Additionally, the axis was correlated with soil magnesium content (r2 = 0.320), and understory openness (r2=0.264) while being negatively associated with shrub cover, sapling and regeneration density (r2 = -0.483). Axis 2 explained 28.7% of the variation and was correlated with several soil nutrients including calcium (r2= 0.47), and potassium (r2= 0.53). Axis 3 explained 8.4% of variation and was associated with variables such as soil zinc and potassium but contributed less to the overall community pattern.

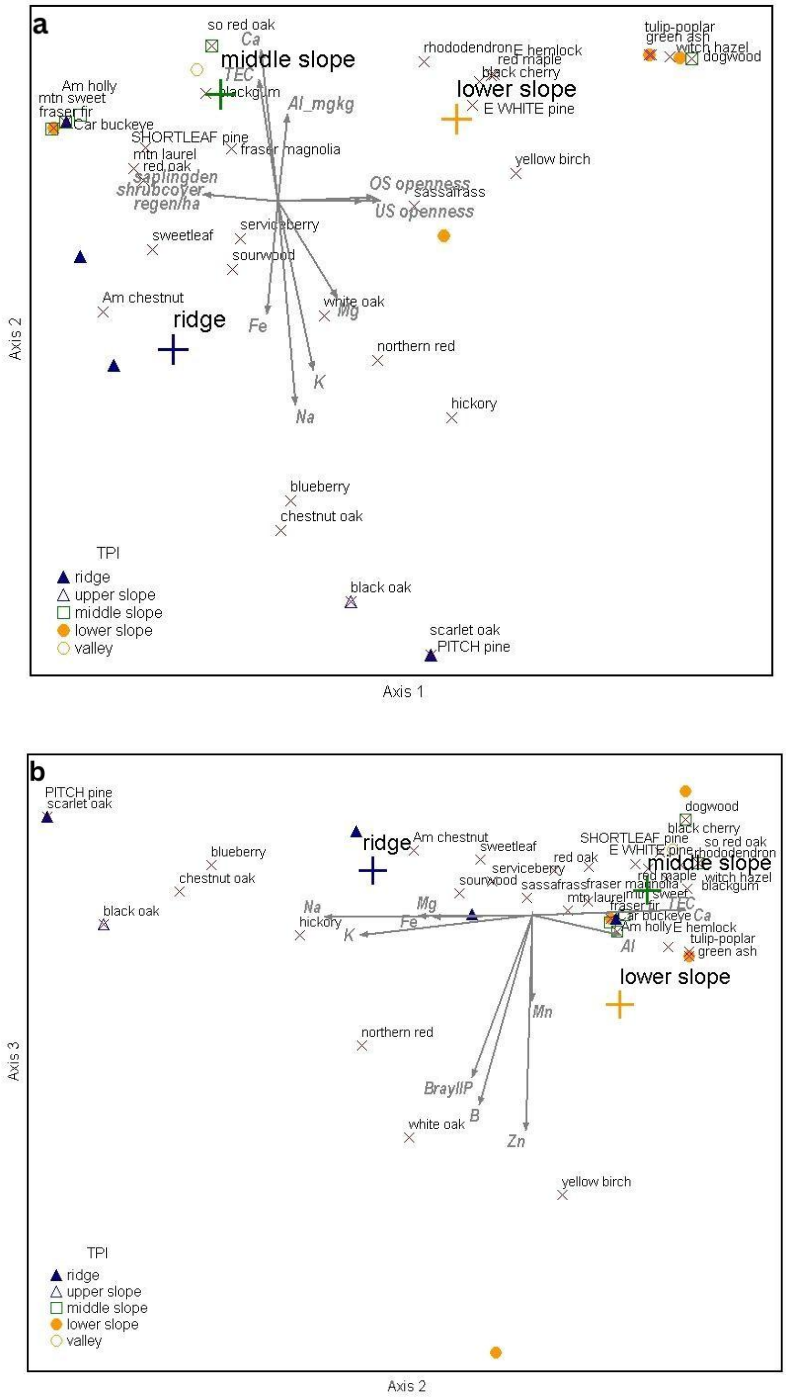

Figure 4.3: Non-metric multidimensional scaling ordination of understory regeneration displaying a) Axis 1 versus Axis 2 and b) Axis 2 and 3.

4.4.3 Tree Establishment Pattern and Growth Release Events Over Time

Tree recruitment peaked in the 1950s, dominated by shortleaf pine, pitch pine, and oak (Figure 4.4). From the early 1900s through the 1950s, pine species, particularly shortleaf pine, were most frequently establishing. Red maple began establishing in the 1940s and Eastern white pine in the 1960s and as these species increased, the establishment of oaks, pitch and shortleaf pine ceased to occur.

Release frequency increased from 1930 onwards, with the highest proportion of releases occurring in the 1940s (Figure 4.4). During this decade, over 20% of trees exhibited some form of major and moderate release. While major releases were more common overall, moderate releases made up a substantial portion in several decades, particularly in the 1950s, 1970s, and 1990s.

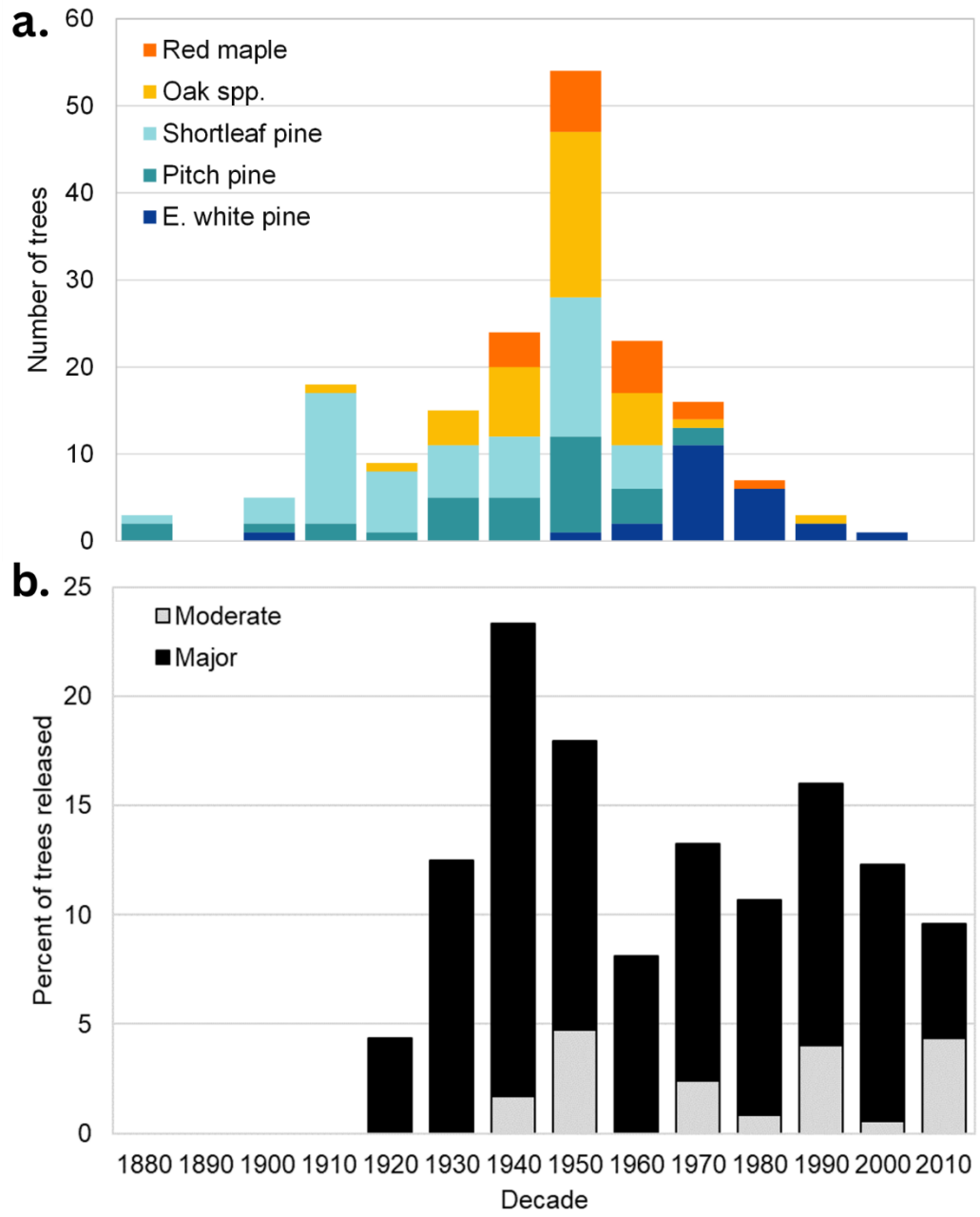

Figure 4.4: a) Number of trees established per decade in a mixed pine-oak stand at Dupont State Recreation Forest, NC. b) Percent of trees exhibiting release events by decade categorized as major or minor releases. A disturbance was considered stand wide if more than 25% of the trees exhibit growth.

4.4.4 Decadal Differences in Species Growth Patterns

ANOVA test showed significant differences among decadal growth for all the species. However, Tukey HSD test indicated that the significance of each decade varied between species (Table S.4.2). For both pitch and shortleaf pine, the growth differed significantly between most decades except between the last and penultimate decades. For white pine however, the difference in growth between second and first decade, penultimate and second decade, and last and penultimate decade were also not significant. For oaks, only the growth difference between the second and first decade was not significant. For red maple, the only significant difference occurred between the penultimate and last decades. For all the species, the mean growth reduced for each following decade; however, red maple had higher growth in the second decade than the first decade (Figure 4.5).

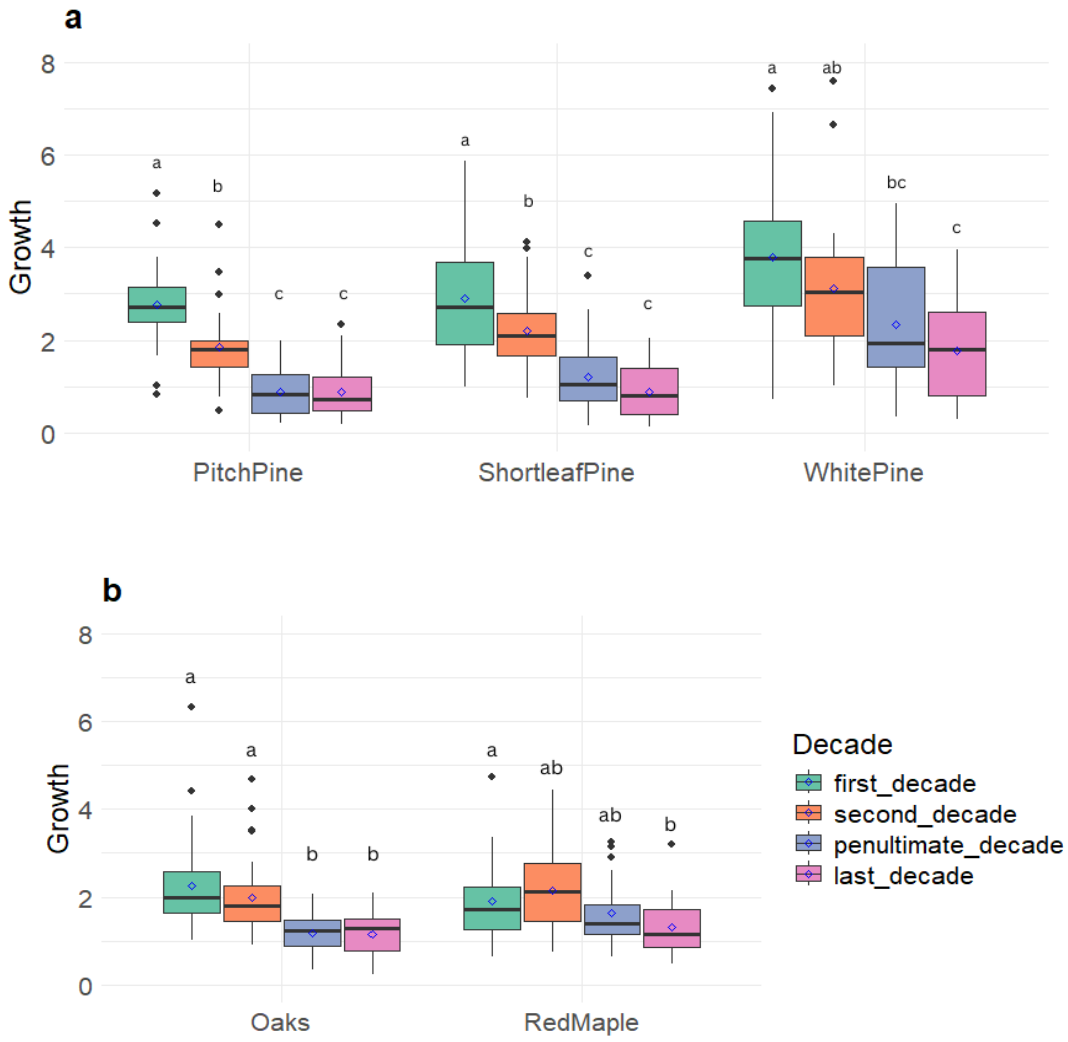

Figure 4.5: Radial Growth (mm) by decade comparison between a) pine, b) oak and maple. Blue diamond shape point represents mean growth for each individual species in each decade.

4.5 Discussion

The age and size structure of the site in DuPont State Forest reflect patterns exhibited in other xeric pine-hardwood forests in the southern Appalachians. The disturbance regime in the

19th and early 20th century appeared to favor pine and oak species, with cohorts of these species establishing up until the 1960s. During this period, there were several lumber mills active in the area that were regularly harvesting the timber and several wildfires were known to have occurred. The open conditions following the logging and fires created favorable conditions for these shade intolerant species to establish. In the 1950s, DuPont Chemical Corporation acquired 4,000 hectares for recreational purposes for their employees, effectively conserving the property. This timing corresponds to a peak in recruitment during the 1950s, primarily involving shortleaf pine, pitch pine, and oaks. DuPont established a forest management program in the late 1950s permitting harvest for one mill. However, after the 1960s, the canopy composition began shifting toward fire-sensitive species of red maple and eastern white pine suggesting a change in the disturbance history. These species can establish under denser canopies when fires are not present to kill off seedlings.

In the contemporary forest pine species including shortleaf, Eastern white, and pitch pine had the highest basal area in the overstory. However, the understory was dominated by mountain laurel creating shade in the understory thus restricting the regeneration of shade intolerant species (Dobbs & Parker, 2004). Only two white pine saplings were measured in the inventory, and pitch and shortleaf pine saplings were absent. This discrepancy between the overstory composition and understory regeneration further supports a lack of pine recruitment as a result of increased stand density, and extensive understory hardwood and shrub cover that prevent light from reaching the forest floor which is an important factor for pine regeneration (Jenkins et al., 2011). Although a few plots in our study area had an open understory, the overstory canopy openness in those plots were still low creating conditions that inhibit pine from regenerating. Studies have shown an increase of overstory pine mortality due to the outbreak of southern pine

beetle (SPB), and reduction of pine regeneration as a result of the fire suppression (Elliott et al., 2008).

Multivariate analyses demonstrated that both overstory and understory species compositions were structured along strong environmental gradients. While overstory composition was more influenced by abiotic variables such as slope, TPI, and aspect, the understory composition was more related to the understory and overstory canopy openness. In contrast, Goode et al., (2021) found that the environmental variables such as elevation and aspect were not correlated to the overstory composition in a more mature shortleaf pine stand in Tennessee. The ordination analysis in this study reinforces the patterns of pine distribution identified in Chapter 2. Eastern white pine occurred on lower slope, while pitch and shortleaf pine occurred on ridges and middle slopes respectively. These findings are consistent with the spatial pattern of pine occurrences mapped in Chapter 2, which documented pine in the same landscape positions. Previous studies on xeric pine-oak ecosystems in southern Appalachians have long recognized the influence of topography and disturbance history in shaping the pine distributions, particularly the dominance of xeric pines in ridgetops and south to southwest facing slopes at mid elevations (Williams, 1998; Whittaker, 1956). Building upon these established understanding, our results provide quantitative evidence linking the contemporary pine establishment patterns and overstory distributions to key environmental gradients and disturbance history demonstrating that both localized and stand-level processes continue to shape the structure and dynamics of xeric pine-oak ecosystems in the southern Appalachians.

Pine-oak forests depend on disturbance (both natural and anthropogenic) for development and persistence. As a result, fires are being reintroduced as a restoration tool to reduce stand density, open up the canopy, and create light-rich conditions necessary for pine seedlings to

establish and grow. Stand replacement fire has been found to be beneficial for pine because it created a mosaic of fire intensities that promotes seedbed conditions suitable for pine regeneration and reduction of mountain laurel from the understory (Vose et al., 1999). This approach closely mimics the natural wildfires conditions needed for pine establishment, while causing only limited nutrient losses and minimal negative impacts on soil and stream chemistry (Vose et al., 1999). Additionally, research have found that maintaining a pine ecosystem in the southern Appalachians requires both fire and SPB disturbances because fire alone supports pine regeneration and establishment while SPB alone may lead to decline of pine forests, but together they can sustain open, pine dominated communities over the long term (Waldron et al., 2007).

4.6 Conclusion

This study highlights the critical challenges pine communities are facing in the southern Appalachian region. Despite pine domination in the overstory, it is ceasing to exist in the understory due to the encroachment by hardwood species such as mountain laurel. Our result further shows that while the overstory species composition is primarily shaped by abiotic factors such as slope, and aspect, the understory composition is strongly influenced by overstory and understory canopy openness, reflecting to the importance of light availability for pine seedling survival and establishment. In conclusion, sustaining pine-oak ecosystems in the southern Appalachian will require active management including reintroduction of fire and creation of canopy gaps to promote pine regeneration and maintain ecosystem resilience.

References

Altman, J., Fibich, P., Dolezal, J., & Aakala, T. (2014). TRADER: A package for Tree Ring Analysis of Disturbance Events in R. Dendrochronologia, 32(2), 107–112. https://doi.org/10.1016/j.dendro.2014.01.004

Benfield, E. F., Webster, J. R., Golladay, S. W., Peters, G. T., & Stout, B. M. (1991). Effects of forest disturbance on leaf breakdown in southern Appalachian streams. SIL Proceedings, 1922-2010, 24(3), 1687–1690. https://doi.org/10.1080/03680770.1989.11899049

Bernhardt, E. S., & Palmer, M. A. (2011). The environmental costs of mountaintop mining valley fill operations for aquatic ecosystems of the Central Appalachians. Annals of the New York Academy of Sciences, 1223(1), 39–57. https://doi.org/10.1111/j.1749-6632.2011.05986.x

Brose, P. H., & Waldrop, T. A. (2006). Fire and the origin of Table Mountain pine pitch pine communities in the southern Appalachian Mountains, USA. Canadian Journal of Forest Research, 36(3), 710–718. https://doi.org/10.1139/x05-281

Brose, P. H., & Waldrop, T. A. (2010a). A dendrochronological analysis of a disturbance–succession model for oak–pine forests of the Appalachian Mountains, USA. Canadian Journal of Forest Research, 40(7), 1373–1385. https://doi.org/10.1139/X10-077

Brose, P. H., & Waldrop, T. A. (2010b). A dendrochronological analysis of a disturbance–succession model for oak–pine forests of the Appalachian Mountains, USA. Canadian Journal of Forest Research, 40(7), 1373–1385. https://doi.org/10.1139/X10-077

Caldwell, P. V., Elliott, K. J., Liu, N., Vose, J. M., Zietlow, D. R., & Knoepp, J. D. (2020). Watershed-scale vegetation, water quantity, and water quality responses to wildfire in the

southern Appalachian mountain region, United States. Hydrological Processes, 34(26), 5188–5209. https://doi.org/10.1002/hyp.13922

Chai, Z., & Wang, D. (2016). Environmental influences on the successful regeneration of pine-oak mixed forests in the Qinling Mountains, China. Scandinavian Journal of Forest Research, 31(4), 368–381. https://doi.org/10.1080/02827581.2015.1062912

Chazdon, R. L., & Guariguata, M. R. (2016). Natural regeneration as a tool for large-scale forest restoration in the tropics: Prospects and challenges. Biotropica, 48(6), 716–730. https://doi.org/10.1111/btp.12381

Dey, D. C., Knapp, B. O., Battaglia, M. A., Deal, R. L., Hart, J. L., O'Hara, K. L., Schweitzer, C. J., & Schuler, T. M. (2019). Barriers to natural regeneration in temperate forests across the USA. New Forests, 50(1), 11–40. https://doi.org/10.1007/s11056-018-09694-6

Dobbs, M. M., & Parker, A. J. (2004). Evergreen Understory Dynamics in Coweeta Forest, North Carolina. Physical Geography, 25(6), 481–498. https://doi.org/10.2747/0272-3646.25.6.481

Elliott, K. J., Vose, J. M., & White, A. S. (2008). Pine regeneration following wildland fire. In: Olberding, Susan D., and Moore, Margaret M., Tech. Coords. Fort Valley Experimental Forest-A Century of Research 1908-2008. Conference Proceedings; August 7-9, 2008; Flagstaff, AZ. Proc. RMRS-P-55. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 162-167, 055. https://research.fs.usda.gov/treesearch/32509

Frolking, S., Palace, M. W., Clark, D. B., Chambers, J. Q., Shugart, H. H., & Hurtt, G. C. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. Journal of

Geophysical Research: Biogeosciences, 114(G2), 2008JG000911. https://doi.org/10.1029/2008JG000911

Gilliam, F. S., & Platt, W. J. (1999). [Effects of long-term fire exclusion on tree species composition and stand structure in an old-growth Pinus palustris (Longleaf pine) forest. Plant Ecology, 140(1), 15–26. https://doi.org/10.1023/A:1009776020438

Goode, J. D., Hart, J. L., Dey, D. C., Torreano, S. J., & Clark, S. L. (2021). Spatial Patterns of Canopy Disturbance and Shortleaf Pine in a Mixedwood Forest. Forest Science, 67(4), 433–445. https://doi.org/10.1093/forsci/fxab017

Hagan, D. L., Waldrop, T. A., Reilly, M., & Shearman, T. M. (2015). Impacts of repeated wildfire on long-unburned plant communities of the southern Appalachian Mountains. International Journal of Wildland Fire, 24(7), 911. https://doi.org/10.1071/WF14143

Harrod, J., White, P. S., & work(s):, M. E. H. R. (1998). Changes in Xeric Forests in Western Great Smoky Mountains National Park, 1936-1995. Castanea, 63(3), 346–360.

Jenkins, M. A., Klein, R. N., & McDaniel, V. L. (2011). Yellow pine regeneration as a function of fire severity and post-burn stand structure in the southern Appalachian Mountains. Forest Ecology and Management, 262(4), 681–691. https://doi.org/10.1016/j.foreco.2011.05.001

Kenefic, L. S., Kabrick, J. M., Knapp, B. O., Raymond, P., Clark, K. L., D'Amato, A. W., Kern, C. C., Vickers, L. A., Dey, D. C., & Rogers, N. S. (2021). Mixedwood silviculture in North America: The science and art of managing for complex, multi-species temperate forests. Canadian Journal of Forest Research, 51(7), 921–934. https://doi.org/10.1139/cjfr-2020-0410

Khodaee, M., Hwang, T., Kim, J., Norman, S. P., Robeson, S. M., & Song, C. (2020). Monitoring Forest Infestation and Fire Disturbance in the Southern Appalachian Using a Time Series Analysis of Landsat Imagery. Remote Sensing, 12(15), 2412. https://doi.org/10.3390/rs12152412

Lafon, C. W., Waldron, J. D., Cairns, D. M., Tchakerian, M. D., Coulson, R. N., & Klepzig, K. D. (2007). Modeling the Effects of Fire on the Long-Term Dynamics and Restoration of Yellow Pine and Oak Forests in the Southern Appalachian Mountains. Restoration Ecology, 15(3), 400–411. https://doi.org/10.1111/j.1526-100X.2007.00236.x

Martin, M. P., Peters, C. M., Asbjornsen, H., & Ashton, M. S. (2021). Diversity and niche differentiation of a mixed pine–oak forest in the Sierra Norte, Oaxaca, Mexico. Ecosphere, 12(4), e03475. https://doi.org/10.1002/ecs2.3475

Maxwell, R. S., & Larsson, L. A. (2021). Measuring tree-ring widths using the CooRecorder software application. *Dendrochronologia*, *67*, 125841.

McCune, B., & Mefford, M. J. (1999). PC-ORD: Multivariate Analysis of Ecological Data; Version 4 for Windows ; [user' S Guide]. MjM Software Design.

Mlambo, D., Kumar, M., Sahoo, U. K., & Syampungani, S. (2023). Editorial: Effect of forest disturbances on natural forest regeneration in a changing tropical environment. Frontiers in Forests and Global Change, 6, 1289315. https://doi.org/10.3389/ffgc.2023.1289315

Nowacki, G. J., & Abrams, M. D. (1997). Radial-Growth Averaging Criteria for Reconstruction Disturbance Histories from Presettlement-Origin Oaks. Ecological Monographs, 67(2), 225. https://doi.org/10.2307/2963514

Oswalt, S. N., Smith, W. B., Miles, P. D., & Pugh, S. A. (2019). Forest Resources of the United States, 2017: A technical document supporting the Forest Service 2020 RPA Assessment

(No. WO-GTR-97; p. WO-GTR-97). U.S. Department of Agriculture, Forest Service. https://doi.org/10.2737/WO-GTR-97

Phillips, D. L., & Murdy, W. H. (1985). Notes: Effects of Rhododendron (Rhododendron maximum L.) on Regeneration of Southern Appalachian Hardwoods. Forest Science, 31(1), 226–233.

Pickett, S. T. A., & White, P. S. (1985). The ecology of natural disturbance and patch dynamics. Academic press.

R Core Team. (2024). R: A language and environment for statistical computing [Computer software]. https://www.R-project.org/

Reilly, M. J., Wimberly, M. C., & Newell, C. L. (2006). Wildfire effects on plant species richness at multiple spatial scales in forest communities of the southern Appalachians. Journal of Ecology, 94(1), 118–130. https://doi.org/10.1111/j.1365-2745.2005.01055.x

Schwarz, P. A., Fahey, T. J., & McCulloch, C. E. (2003). FACTORS CONTROLLING SPATIAL VARIATION OF TREE SPECIES ABUNDANCE IN A FORESTED LANDSCAPE. Ecology, 84(7), 1862–1878. https://doi.org/10.1890/0012-9658(2003)084%255B1862:FCSVOT%255D2.0.CO;2

Seidl, R., Rammer, W., & Spies, T. A. (2014). Disturbance legacies increase the resilience of forest ecosystem structure, composition, and functioning. Ecological Applications, 24(8), 2063–2077. https://doi.org/10.1890/14-0255.1

Stokes, M., & Smiley, T. L. (1996). An Introduction to Tree-Ring Dating. University of Arizona Press.

Turner, M. G., Pearson, S. M., Bolstad, P., & Wear, D. N. (2003). Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). Landscape Ecology, 18(5), 449–464. https://doi.org/10.1023/A:1026033116193

Vatandaslar, C., Lee, T., Bettinger, P., Ucar, Z., Stober, J., & Peduzzi, A. (2024). Mapping percent canopy cover using individual tree- and area-based procedures that are based on airborne LiDAR data: Case study from an oak-hickory-pine forest in the USA. Ecological Indicators, 167, 112710. https://doi.org/10.1016/j.ecolind.2024.112710

Vose, J. M., Swank, W. T., Clinton, B. D., Knoepp, J. D., & Swift, L. W. (1999). Using stand replacement fires to restore southern Appalachian pine–hardwood ecosystems: Effects on mass, carbon, and nutrient pools. Forest Ecology and Management, 114(2), 215–226. https://doi.org/10.1016/S0378-1127(98)00352-1

Waldron, J. D., Lafon, C. W., Coulson, R. N., Cairns, D. M., Tchakerian, M. D., Birt, A., & Klepzig, K. D. (2007). Simulating the impacts of southern pine beetle and fire on the dynamics of xerophytic pine landscapes in the southern Appalachians. Applied Vegetation Science, 10(1), 53–64. https://doi.org/10.1111/j.1654-109X.2007.tb00503.x

Web Soil Survey. (n.d.). Retrieved January 24, 2025, from https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx

Whittaker, R. H. (1956). Vegetation of the great smoky mountains. *Ecological monographs*, *26*(1), 2-80.

Wickham, J. D., Riitters, K. H., Wade, T. G., Coan, M., & Homer, C. (2007). The effect of Appalachian mountaintop mining on interior forest. Landscape Ecology, 22(2), 179–187. https://doi.org/10.1007/s10980-006-9040-z

Williams, C. E. (1998). History and Status of Table Mountain Pine–Pitch Pine Forests of the Southern Appalachian Mountains (USA). Natural Areas Association, 18(1), 81–90.

CHAPTER 5 - The influence of understory vegetation density on the accuracy of modeling tree structure using terrestrial and handheld mobile laser scanning

5.1 Abstract

Understory vegetation poses significant challenges for accurate tree structure modeling using ground-based laser scanning technologies in forest inventory applications. This study compares the performance of terrestrial laser scanning (TLS) and handheld mobile laser scanning (HMLS) in estimating tree diameter at breast height (DBH) and stem detection rates across varying levels of understory density in a mature mixed pine-hardwood forest in North Carolina. Nine fixed-area plots were stratified by understory density (open, medium, dense), and both TLS and HMLS data were collected and analyzed using filtering and segmentation techniques to isolate tree stems from LiDAR point clouds. Our results indicate that stem detection rates and DBH estimation accuracy for both TLS and HMLS decline as understory density increase, with HMLS achieving higher detection rates than TLS in dense vegetation. TLS demonstrated more consistent DBH accuracy across vegetation types but is more time consuming. Both technologies underestimated DBH as tree size increased, although this effect was only significant for TLS. The findings highlight trade-offs between detection efficiency, and DBH accuracy, highlighting the importance of evaluating the site conditions and project objectives before selecting appropriate laser scanning technologies for forest inventory.

5.2 Introduction

Although understory vegetation is a critical component of the forest ecosystem for nutrient cycling and biodiversity (Suchar & Crookston, 2010), when abundant it creates significant challenges in modeling tree stems for forest inventory. Forest understory vegetation is

influenced by factors such as composition and structure of the overstory forest floor, disturbance or management history, forest age (Barbier et al., 2008), soil nutrients (Moir, 1966), light availability (Légaré et al., 2002), and forest canopy gaps (Dupuy & Chazdon, 2008; Runkle, 1982). Areas with abundant sunlight and nutrient-rich soils often support dense understories, while shaded or nutrient-poor sites may have sparser vegetation (Aiba & Kitayama, 2020). This variability in understory density is desirable as it adds structural heterogeneity important for wildlife species and vegetation regeneration dynamics, yet it complicates the efficiency of forest inventory methods.

Laser scanning (LS) technology is becoming widely used for forest inventory (Bienert et al., 2021; Chen et al., 2019; Maas et al., 2008; Vauhkonen et al., 2014), due to its ability to quickly and non-destructively capture the 3D structure of individual trees as well as entire forest stands at numerous spatial scales (Åkerblom & Kaitaniemi, 2021). Among the available LS platforms, aerial laser scanning (ALS) is the most frequently used platform in forest research (Donager et al., 2021) to capture efficient and large scale forest inventory data (Wing et al., 2012). While ALS is capable of covering extensive areas, the resultant low density point cloud (< 1 – 200 points/m^2) is ill suited for detailed single tree measurements (Gollob et al., 2020), and can lead to increased cost for an individual tree or plot level assessment. Additionally, the vertical viewpoint of ALS technology poses challenges in areas with interlocking crowns for acquiring forest attributes situated beneath dense canopies (Murgoitio et al., 2013) such as diameter at breast height (DBH). This constraint of ALS has led researchers to utilize ground based platforms that operate beneath forest canopy such terrestrial laser scanning (TLS) and mobile laser scanning (MLS) when focused on measurements beneath the forest canopy (Bauwens et al., 2016; Bienert et al., 2006; Dassot et al., 2011; Guan et al., 2015; Liang et al., 2016).

TLS captures high-density (>1000 points/m^2) 3D data of forests at a millimeter-level detail, enabling precise measurements of tree structure (Li et al., 2021; Liang et al., 2016). TLS-derived tree height has been shown to be more accurate (RMSE = 0.55 m) than field-measured height (RMSE = 1.28 m) when compared with a reference height from destructive sampling (Calders et al., 2015). High accuracies for estimates of DBH are also commonly reported, for example RMSE ranged from 0.019–0.037 m in a dense Corsican pine stand (*Pinus nigra* subsp. *laricio*) (Tansey et al., 2009). Other forest parameters such as aboveground biomass also exhibit high accuracy (R^2 = 0.98, RMSE = 20.4 kg), when compared to values obtained via destructive sampling (Stovall et al., 2017). Additionally, TLS also enables the extraction of tree attributes such as stem volume, leaf area index, and leaf angle distribution, which are valuable in forest ecology but not easily measured during field inventory (Liang et al., 2016; Zhao et al., 2015). TLS has also demonstrated capabilities in determining tree position with high accuracy (standard error = 0.1 m) with its application in determining spatial distribution of forest canopy at a plot level (Vastaranta et al., 2009; Durrieu et al., 2008). Despite the benefits of TLS, a significant challenge and source of uncertainty is occlusion, especially when working within a dense forest setting (Schneider et al., 2019). Occlusion occurs when trees, understory vegetation, and branches block laser beams from reaching the intended target, and this interference increases with distance from the plot center (Liang et al., 2016). In a single scan approach a significant percentage of the trees ranging from 9% to 40% are undetectable from the plot center (Lovell et al., 2011; Murphy et al., 2010). To mitigate occlusion, adopting a multi-scan approach is recommended (Wilkes et al., 2017) though this may cause the data collection process to become more time consuming and inefficient (Bienert et al., 2012).

On the contrary, MLS significantly reduces the likelihood of occlusion and enables quicker data collection across larger areas with rates up to 12 times faster than TLS and few (less than 5%) trees missed (Bauwens et al., 2016; Liang, et al., 2014; Ryding et al., 2015). Initially, MLS systems were mounted on mobile platforms such as vehicles for forest surveys (Holopainen et al., 2013), utilizing inertial measuring units (IMUs) and global navigation satellite systems (GNSS) for real-time positioning (Chen et al., 2019). However, car- and van-mounted MLS systems are constrained to forest roads, resulting in commission errors and overestimations of DBH and stem volume within 0-10 meters of the sensor and omission errors and underestimations of forest attributes further from the sensor (50-60 m) (Pires et al., 2022).To mitigate the limitation of car- and van- mounted systems, ATVs have been used and proved more effective than van/car-mounted systems, but were also restricted to flat terrain and encountered significant challenges in dense canopies and difficult topography where GNSS signals were unavailable (Kukko et al., 2017). The introduction of Handheld Mobile Laser Scanners (HMLS), also known as Personal Laser Scanning (PLS) was a pivotal advancement. In HMLS, the operator's movement serves as the scanning platform (Bauwens et al., 2016), offering enhanced mobility across various terrain (Liang et al., 2014). Additionally, HMLS employs Simultaneous Localization and Mapping (SLAM), a technology where a mobile robot creates a map of its surroundings while simultaneously using that map to determine its own position (Bailey & Durrant-Whyte, 2006; Chen et al., 2019) making it particularly suitable for forest surveys with obstructed GNSS signals.

The ability to measure the structural characteristics of forest stands using LS has been found to be influenced by the intricacy and attributes of forest stands such as forest type and understory cover with the accuracy of forest visibility estimates reducing with dense understory (Zong et al., 2021). Studies comparing the accuracy and effectiveness of TLS and HMLS have

produced mixed results but suggest that HMLS may yield more accurate estimates of DBH and stem locations and have higher tree detection rate by 8%, primarily due to occlusion challenges (Donager et al., 2021). In contrast, HMLS underestimated DBH in an ash (*Fraxinus*) dominated woodland than TLS due to increased point noise of the data (Ryding et al., 2015). This discrepancy may be attributed to the different types of HMLS used and the varying forest stand structures, with one study focusing on pure ponderosa pine (*Pinus ponderosa*) open structured stands and the other on dense understory conditions. Additionally, performance evaluation of single and multi-scan TLS and HMLS in diverse forest types (broadleaved, coniferous, mixed), densities (ranging from 113 to 1344 trees per hectare), and terrain (ranging from flat to steep), revealed comparable result of HMLS and multi-scan TLS in tree detection and DBH estimates while the single scan resulted in lower accuracy (Bauwens et al., 2016). Multi-scan TLS and HMLS was able to detect all trees with DBH > 10 cm, while single-scan TLS missed 17% of trees due to occlusion. Additionally, multi-scan TLS and HMLS estimated DBH with RMSE < 5% compared to field measurements, whereas the accuracy of single-scan TLS reduced (RMSE = 15%) due to occlusion (Bauwens et al., 2016).

Given these challenges, the objective of this study was to compare the potential of TLS and HMLS in estimating forest parameters, including individual tree DBH, and locations, in a mature mixed pine-hardwood forest characterized by a patchy understory, at both the individual tree and plot scale. Here, we define the understory as all aboveground features below 2 m in height, including shrubs, herbaceous vegetation, saplings, and lower portions of tree stems as all of them influence the data collection from ground based LiDAR. Specifically, we asked:

1. How accurately can MLS and TLS estimate DBH, and detect individual stem across varying levels of understory density when compared to traditional manual field measurements?
2. How does tree size impact the accuracy of MLS and TLS in DBH estimation?

We sought to answer these questions in a naturally regenerated mixed oak-pine forest in which we stratified stands to differentiate understory vegetation complexity, ranging from sparse, single-stemmed vegetation to dense, multi-stemmed, and shrubby vegetation, including both evergreen and deciduous forms. We measured the vegetation using fixed area plots, TLS, and MLS in all strata, with the intention of using the manual field measurements as references for LS derived estimates. Dense understories can be difficult to physically measure, but also challenging to describe quantitatively. Therefore, another goal was to use and compare a LS derived understory density with lifeform specific field-based measurements to assess if interference caused by occlusion is the same whether it be from overstory trees, vines, shrubs, or saplings.

5.3 Methods

5.3.1 Study Site

The DuPont State Recreational Forest is located in Cedar Mountain, North Carolina and encompasses more than 4000 hectares. The site lies in the upland plateau of the Little River Valley within Transylvania and Henderson counties, NC. Elevation ranges from 701 to 1097 meters. The study area is 26 hectares in size and consists of mixed oak (Quercus spp.) and pine (Pinus spp.), including shortleaf pine (Pinus echinata), pitch pine (Pinus rigida), and eastern white pine (P. strobus). From 1981-2022, the mean annual temperature ranged from 12.98 to

15.68 °C, mean annual precipitation from 843.75 to 1792.97 mm, and mean relative humidity ranged from 69 to 81.38% (NASA Langley Research Center, 2022). This area encompasses a diverse range of soil types which are characterized by varying slopes, drainage capacities, and compositions. Predominantly, the soils in the region are well-drained with 8 to 15 percent slopes. Soils present in both moderate and steep slopes (8 to 95 percent), are somewhat excessively drained, and feature rocky and very rocky textures. Additionally, areas with fine sandy loam are notable for its frequent flooding at gentle slopes, 0 to 2 percent (Web Soil Survey, n.d.).

5.3.2 Field Methods

Experimental design. A stratified random sampling design was created to ensure we would capture the range of understory vegetation density throughout the site. Aerial light detection and ranging (LiDAR) data was sourced from 3DEP LiDAR explorer and the canopy layer was removed (Dahal, Chapter 1). Strata were designated based on a visual assessment of the point density. Potential plot centers were placed randomly within each stratum and inspected in the field to confirm suitability for manual measurements.

Manual measurements: Nine fixed area plots were established to quantify species composition and structure, with varied sizes per lifeform. Overstory trees (individuals ≥ 10 cm DBH), were measured using 0.04 ha plot (11.35 m radius) and species, crown class, DBH, distance and azimuth from the plot center was recorded. Saplings (woody stems ≥ 1.4 m height, < 10 cm DBH), were measured on 0.004 ha circular plots (3.6 m radius), located 7.6 m north of the plot center. For each sapling, species, DBH and height class (1.4–3 m, 3-5m, 5-10 m, >10 m) was recorded. Additionally, the percentage cover of shrub species was visually estimated and recorded.

HMLS scans were taken in each of the nine plots by walking 5 m transect, both up and down the slope, using Leica BLK2GO device in May, 2022. Each scan approximately took about 10-15 minutes depending on the understory complexity. The BLK2GO device is a compact dual-axis LiDAR sensor with scan velocity of 420,000 points per second (Limongiello et al., 2020). It is based on SLAM technology and has three panoramic cameras that continuously capture point clouds with XYZ coordinates of the environment you walk in. Data captured by Leica BLK2GO comes in its proprietary extension .b2g and Cyclone Register 360 (BLK edition) (Leica Geosystems, 2022) was used to convert the data from .b2g to .las file.

For TLS, Faro Focus 3D X330 was placed at the plot center and scans were taken in August, 2023. In plots with dense understory, two scans were taken to minimize occlusion and ensure full coverage. Reflective targets were strategically placed within the plot to facilitate the alignment of multiple scans. These targets were placed such that they were clearly visible from all scanning position. The Faro Focus 3D has a working range of up to 120 m and collected points at a rate of 9.76×10^5 per second. Each point provides XYZ coordinates and an intensity value (Ryding et al., 2015).

5.3.3 Data Analysis

To verify and better quantify the understory vegetation complexity, we classified understory density into three categories open, medium, and dense based on the percentage of shrub cover, sapling, and tree density as all of these influence the range of the laser scanner (Figure 5.1). Firstly, we calculated the mean shrub cover and mean sapling and tree density for each plot. Then we assign them a score of 1 (open), 2 (medium), and 3 (dense). Since high shrub cover did not always correspond with high sapling or tree density, we used a majority rule: plots scoring 3

in most categories were classified as dense, those with a majority score of 2 as medium, and those with a majority score of 1 as open.

Figure 5.2: Visual representation of the three strata differentiated based on understory density. The top image is plot with open understory, middle image is representing medium understory, and the bottom image is representing dense understory. The images are extracted from Faro SCENE.

All LiDAR datasets were first classified into ground and non-ground points using the cloth simulation filter algorithm (Zhang et al., 2016) in the lidR package (Roussel et al., 2020) using R. In this approach, a LiDAR point cloud is inverted, and a rigid cloth is draped over the inverted surface. By examining the interactions between the cloth nodes and the corresponding LiDAR points, the positions of the cloth nodes are determined, creating an approximation of the ground surface (Zhang et al., 2016). Following ground classification, the point clouds were normalized using the normalize_height() function in the lidR package which substitutes the vertical position of each point with its respective distance from the ground (Khosravipour et al., 2014). To further refine the dataset, an isolated voxels filter was applied using the classify_noise() function to segment and remove noise and outliers from the point clouds.

A multi-step filtering and segmentation process was applied to isolate tree stems from the point cloud. First, point-wise structural metrics such as verticality and curvature were computed using the fastPointMetrics() function from the TreeLS package (Conto et al., 2019). Points were filtered based on the following criteria: verticality between 80 and 100, and curvature between 0 and 0.2 (Figure 5.2). This filtering step effectively removed most understory vegetation while retaining stem structures.

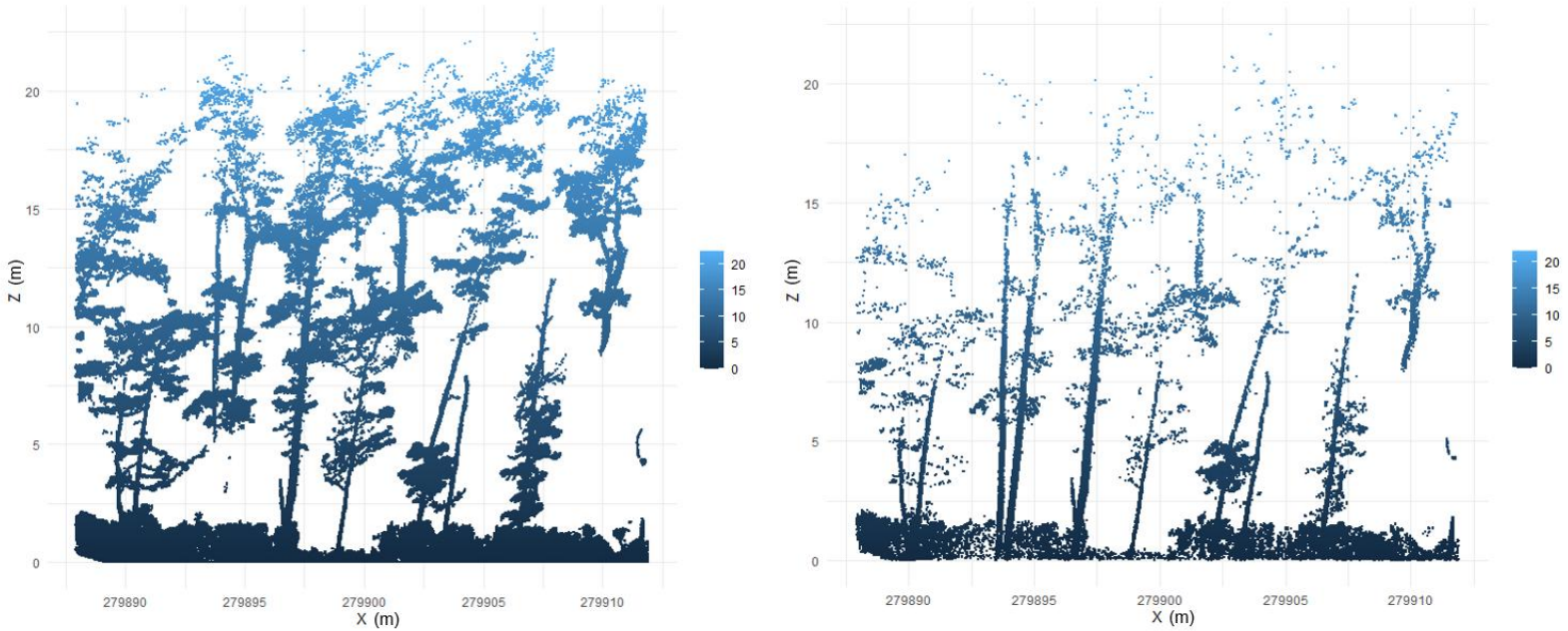

Figure 5.2: Original HMLS data (left) vs data after filtering with verticality and curvature (right)

To identify the trunk within the point cloud, we applied a density-based spatial clustering approach (DBSCAN) to segment the point cloud into spatially coherent clusters using the epsilon as 0.1 and minimum points as 15 (Figure 5.3). Each resulting cluster was then evaluated based on the structural criteria. Specifically, we calculated the number of points, vertical extent, and angle between the principal direction of the cluster and the vertical axis. Clusters were retained if they contained more than 0.1% of the total points or a minimum of 20 points, had a height greater than 2 meters, and exhibited orientation within 30 degrees of vertical (Figure 5.3). This multi-criteria filtering ensured the extraction of clusters consistent with vertical cylindrical structures, effectively isolating potential tree stems from the broader point cloud.

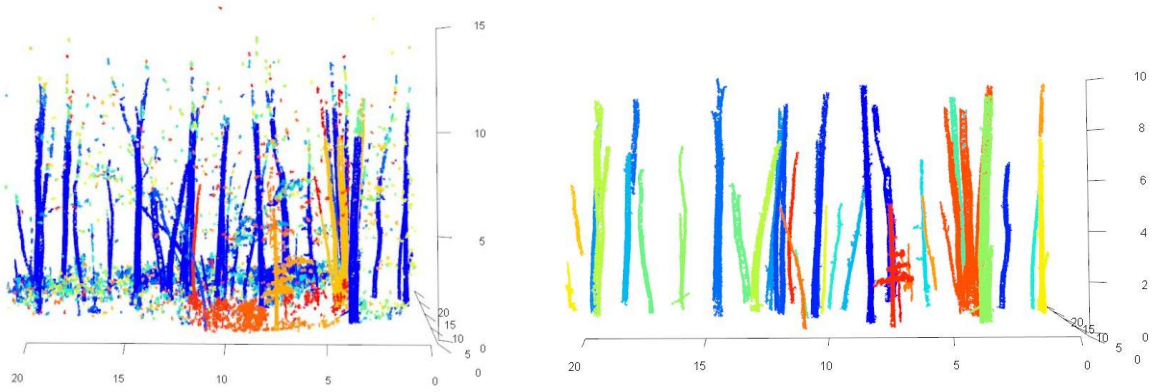

Figure 5.3: DBSCAN clustering (left) vs filtered data using the structural criteria (right)

The filtered tree stems were used to fit the random sample consensus (RANSAC) cylinder fitting algorithm from the TreeLS package to identify the tree locations and estimate the DBH. RANSAC algorithm is particularly well-suited for this task because the tree stems can often be represented as cylindrical shapes in a 3D space, and this algorithm fits geometric models like cylinders to the data that may contain noise or outlier.

RANSAC is an iterative algorithm that distinguishes between data points that fit the model (inliers) and those that do not fit (outliers). In our case the model is a cylinder, representing the tree stem. RANSAC begins by randomly selecting a subset of points from the point cloud at breast height (1.37 m). Using the selected subset of points, the algorithm fits a cylinder model. Once the model is fitted, RANSAC evaluates how well this model fits the rest of the points in the dataset. There exists a chance that all selected points belong to the model which are the inliers. This step is repeated for multiple times and each iterations provides a new randomly selected subset of points and potentially a new cylinder model. After all iterations are complete, the model with the highest number of inliers is selected as the best representation of the tree bole. The number of iterations required to ensure that the algorithm finds at least one good model with high probability is given by:

$$N = \frac{\log(1 - P)}{\log(1 - p^n)}$$

Where N represents the number of iterations needed to achieve a probability P of identifying at least one valid model. The variable n denotes the number of points selected, while p is the probability that a single chosen point is an inlier. This probability p can be expressed as:

$$p = \frac{M}{D}$$

Here, M is the count of inliers in the model, and D is the total number of points in the dataset. Typically, the ratio p is unknown and must be estimated from the dataset being analyzed (Olofsson et al., 2014).

Once the stem centers and DBH were extracted, they were matched to the corresponding field-measured fixed-area plot trees manually. For each plot, the detection rate was determined by dividing the number of matched trees by the total number of field-measured trees. Omission rates was calculated as the number of trees measured in the field that were not detected with the LiDAR dataset. Comparisons were made at the plot level, and the accuracy of DBH was determined.

To assess accuracy at the tree level, the estimated individual tree DBH estimates from both HMLS and TLS datasets were compared to the field measurements. To determine the accuracy of the estimates, mean absolute error (MAE) and root mean squared error (RMSE) was calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\frac{|d_i - D_i|}{D_i}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i - D_i)^2}$$

Where d_i is the estimated DBH from the point cloud, D_i is the field measured DBH and n is the number of trees. MAE and RMSE will be plotted across DBH classes to compare the LiDAR technologies and identify any trends.

Kruskal-Wallis rank sum test was performed to assess the degree to which understory density influenced DBH estimation for both TLS and MLS. Additionally, to examine how tree size influenced the accuracy of DBH estimation, we firstly calculated the DBH estimation error as the difference between the field measured DBH and LiDAR derived DBH. Tree size was

represented by the field measured tree and was treated as a continuous variable. We then fitted a linear regression model for MLS and TLS separately using field measured DBH as the independent variable and DBH error as the dependent variable.

5.4 Results

The understory vegetation had categories including open, medium and dense showed clear differences in the structural characteristics (Table 5.1). Dense plots had the mean tree density of 933.3 stems/ha, mean sapling density of 7833.3 stems/ha and shrub cover of 80.83%. Compared to this, open plots had significantly lower values with a mean tree density of 641.6 stems/ha, mean sapling density of 3583.3 stems/ha, and mean shrub cover of 33.4%.

Table 5.1: Summary of the plots classified into open, medium and dense.

Understory Vegetation Density	**Mean Tree Density (stems/ha)**	**Mean Sapling Density (stems/ha)**	**Mean Shrub Cover (%)**
Open	641.6	3583.3	33.4
Medium	741.6	5583.3	65.4
Dense	933.3	7833.3	80.83

Stem detection rate was influenced by understory density for both ground based laser scanners (Figure 5.4). For MLS, the overall detection rate in open understory density was 66.53%, medium understory was 56.22%, and dense understory density was 48.40%. In contrast, the TLS

had even lower detection rate of 56.61% for open understory, 42.25% for medium understory, and for 40.33% dense understory.

The average omission rate for MLS was 33.47% for open understory, 43.78% for medium understory and 51.6% for dense understory. In case of TLS, the omission rate increased to 43.39% for open understory, 57.75% for medium understory, and 59.67% for dense understory.

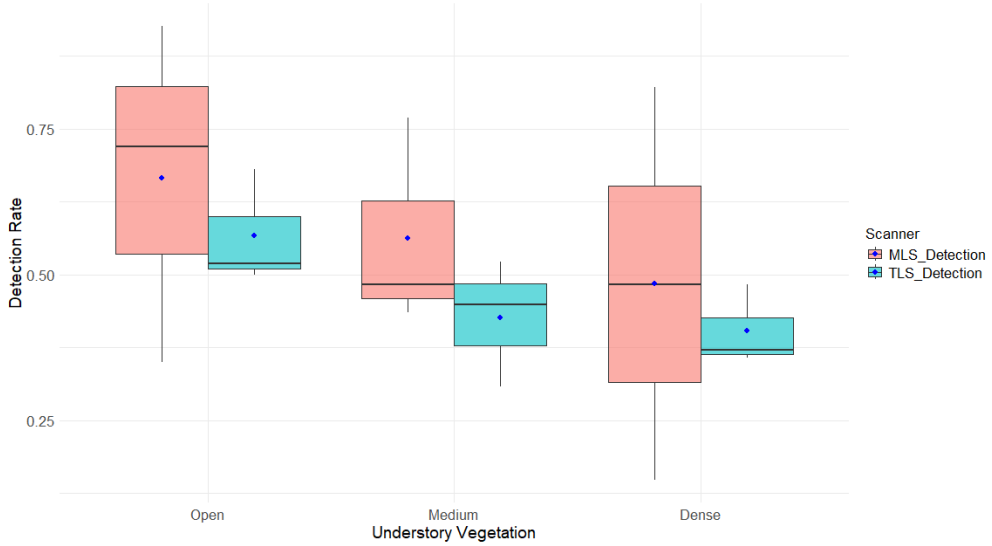

Figure 5.4: Detection rate by MLS and TLS under different understory density. Blue points represent mean detection rate.

DBH estimation accuracy was also influenced by understory density for both MLS and TLS (Figure 5.5). For MLS, the mean DBH error was low under open and dense understory density while it underestimated the DBH for the medium understory with a MAE of 3.38 (Table 5.2). In contrast, TLS had a higher mean DBH error in medium understory while it tended to underestimate DBH in open and dense understory density. Similarly, RMSE increased with the increase in understory density for MLS with the highest RMSE being for the medium understory.

For TLS, RMSE was the lowest for the dense understory density and highest for the medium understory density.

Additionally, Kruskal-Wallis test showed significant effect of understory density in the accuracy of DBH estimation by MLS, however such effect was not significant for TLS. For MLS, that significant difference only existed between DBH estimates in open and medium understory density (p.adj = 0.002).

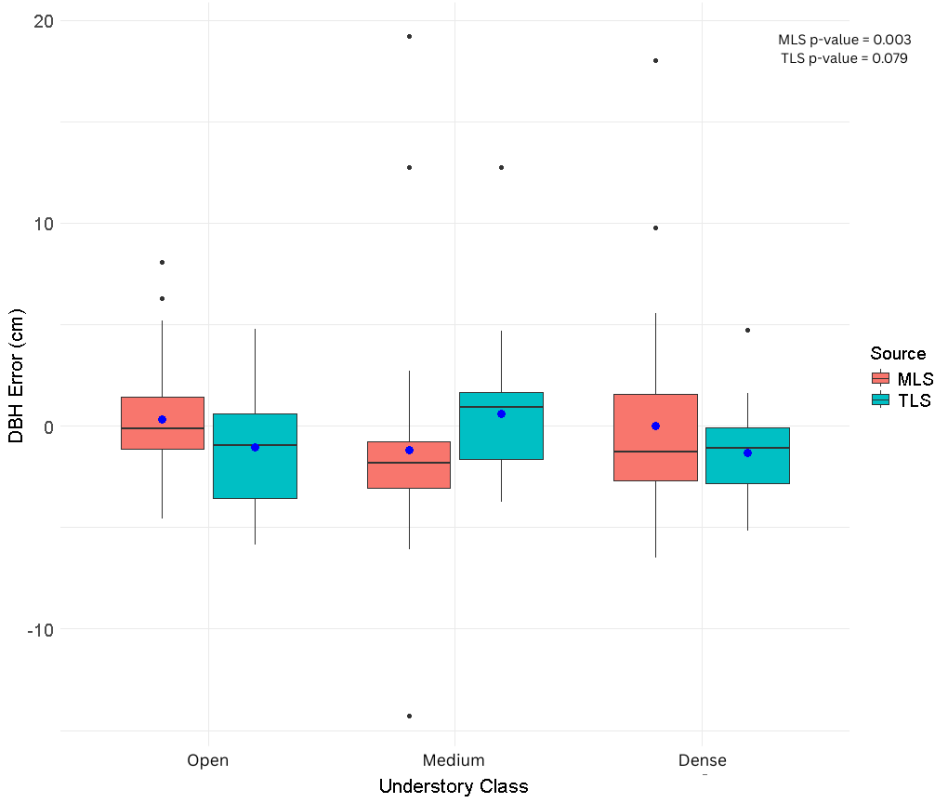

Figure 5.5: DBH estimation error under different understory density for MLS and TLS. The mean DBH error per understory density is shown as blue point.

Table 5.2: Summary statistics for the DBH estimation by MLS and TLS under different understory density.

Laser Scanner	Understory Vegetation	RMSE	MAE
MLS	Open	2.53	1.76
	Medium	5.31	3.38
	Dense	4.40	3.05
TLS	Open	3.25	2.74
	Medium	3.44	2.31
	Dense	2.58	2.02

With the increase in tree size, TLS significantly underestimated the DBH (estimate = -0.07337, p-value = 0.0082). Although MLS also underestimated the DBH with increase in tree size such effect were not significant (estimate = -0.05301, p-value = 0.117).

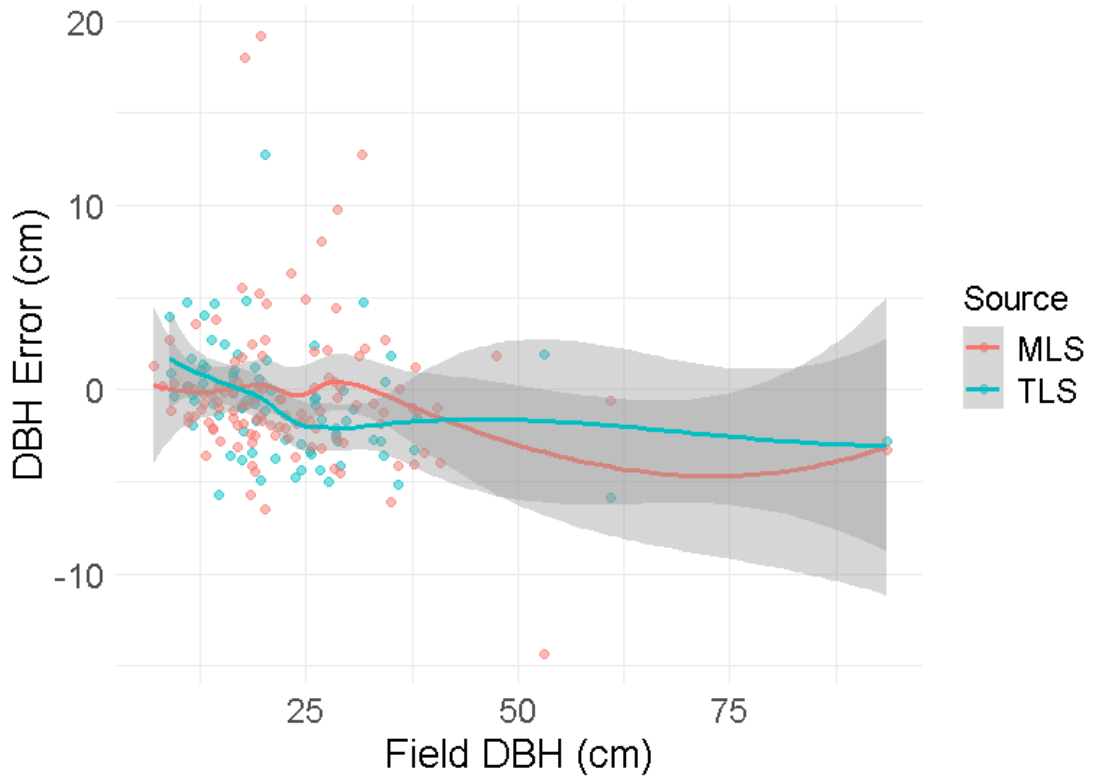

Figure 5.6: Relationship between field measured DBH and DBH estimation error across MLS and TLS. Each point represents an individual trees. The loess smoothing curve with shaded confidence interval shows the underlying trend in DBH estimation error relative to tree size by fitting a localized regressions across the data, allowing for visualization of non-linear patterns in error as tree size increases.

5.5 Discussion

While laser scanning technology such as MLS and TLS was initially developed for engineering and geospatial applications such as surveying, infrastructure monitoring, and urban mapping these technologies have been increasingly adapted for use in forest inventory for measuring DBH, height, estimation of aboveground biomass, and/or forest structure (Calders et al., 2015; Donager et al., 2021; Liang, Hyyppa, et al., 2014; Olofsson et al., 2014). For example,

in a pure ponderosa pine forest, DBH and tree height was estimated with a MAE of 5 cm and 2.2 cm respectively for TLS and 4.8 cm and 1.3 cm respectively for ALS (Donager et al., 2021). TLS was used to estimate above ground biomass in a native Eucalypt Open Forest with a coefficient of variation of the RMSE of 16.1% (Calders et al., 2015). MLS was used for forest inventory in an area with flat terrain and sparse understory achieving 87.5% stem mapping accuracy and RMSE of 2.36 cm and 0.28m for DBH and location estimations respectively (Liang et al., 2014). While more studies exists that highlight the potential of laser scanning for forest measurements, majority of them have been conducted in either plantations or mature closed canopy forests with minimal understory vegetation. Understory vegetation including shrub and herbs layers, lower branches and foliage significantly affects the range of the ground based laser scanner device (Bienert et al., 2018).

Our results showed that understory vegetation density influences both the stem detection rate and the DBH estimation accuracy for both ground bases laser scanners. The stem detection rate was higher for MLS than TLS across all the understory density. Both MLS and TLS achieved highest stem detection rate of 72.01% and 50.21% for plots with open understory vegetation. Interestingly, within each scanning method, the mean detection rates for medium and dense understory vegetation plots were relatively similar suggesting that detection performance may reach a threshold beyond which additional vegetation does not further reduce detection efficient. This trend is consistent with an earlier study reporting detection decreasing from 75% in low density stands (600 stems/ha) to 30% in high density (2000 stems/ha) (Liang et al. 2018).

A similar pattern was observed for the DBH estimation accuracy. The RMSE for MLS ranged more widely from 2.53 to 5.31, depending on the level of understory density. In contrast, TLS showed a narrower RMSE range, from 2.58 to 3.25 indicating more consistent performance

across varying vegetation densities. This could be attributed to the adaptive scanning strategy used during the TLS data acquisition where the number of scans was increased in denser plots to mitigate the effects of occlusion. DBH estimation accuracy is known to improve when multiple TLS scans are used instead of a single scan per plot (Bauwens et al., (2016). Moreover, this particular study reported that the best DBH estimation was achieved using HMLS (RMSE 1.11 cm) under the open understory density. This aligns with our result, where the lowest RMSE across both MLS and TLS was also obtained under open understory density, specifically 2.58 cm for MLS, highlighting the advantage of mobile scanning systems in open forest conditions. Additionally, both TLS and MLS underestimated the tree DBH with increase in tree size although such effect was only significant for TLS. TLS has been found to underestimate DBH and tree height with the increase in the tree size (Kükenbrink et al., 2022; Liu et al., 2018).

When comparing MLS and TLS performance, it is also important to consider the data acquisition and processing logistics. MLS scanning with the Leica BLK2GO took approximately 10–15 minutes per plot, depending on understory density. TLS scans, in contrast, took 12–13 minutes per scan, with additional time required for multi-scan setups in denser plots. TLS setup involved transporting and stabilizing a tripod and scanner which is heavier than the MLS equipment used in this study. The MLS was relatively lightweight and because it was a handheld device it allowed for broader coverage than the TLS scans.

However, the MLS device produced data in a proprietary .b2g format incompatible with open-source platforms like CloudCompare. Consequently, we had to purchase a one-year license for Cyclone REGISTER 360 (approximately $3,000 USD) to convert the data to LAS format. In contrast, TLS data could be directly processed in CloudCompare without additional software investment. In areas with dense understory, we encountered another MLS-specific issue of

double stems that appeared in the point cloud. Options for managing this issue include the exclusion of any plots where double stems occur (Bauwens et al., 2016) or manually cleaning the point clouds in software such as CloudCompare before proceeding with analysis.

These additional costs and processing demand raise the question of whether the adoption of MLS and TLS offers meaningful advantages over the traditional field-based inventories. While ground based inventories remain the standard in forest inventory, they are labor intensive, time consuming, and prone to observer bias. In contrast, MLS and TLS can offer significant benefit in terms of faster data collection and reduction of human bias in the DBH estimation. Although ground based LiDAR also involve considerable setup time and financial investment, they provide added benefit of capturing multiple forest structure parameters which may be a difficult to obtain from the traditional field based methods.

5.6 Conclusion

Our study highlights the importance of considering understory vegetation when evaluating the performance of ground based laser scanning technology in natural forest conditions. Understory vegetation is an important and inevitable part of forest ecosystems. While both TLS and MLS show promising results in stem detection and DBH estimation, the presence of both understory vegetation and tree density greatly influenced their performance in extracting essential forest attributes. MLS has advantages of higher detection rate and faster data acquisition time but requires both costly proprietary software and additional pre-processing in dense understory vegetation. On the other hand, TLS was time intensive but showed consistent DBH accuracy across vegetation types. Our findings showed that there are tradeoffs between

both ground based laser scanning software and the selection of either technology should be based on specific research objectives and the potential forest types where the scanners will be applied.

References

Aiba, S., & Kitayama, K. (2020). Light and nutrient limitations for tree growth on young versus old soils in a Bornean tropical montane forest. *Journal of Plant Research*, *133*(5), 665–679. https://doi.org/10.1007/s10265-020-01217-9

Åkerblom, M., & Kaitaniemi, P. (2021). Terrestrial laser scanning: A new standard of forest measuring and modelling? *Annals of Botany*, *128*(6), 653–662. https://doi.org/10.1093/aob/mcab111

Bailey, T., & Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics & Automation Magazine*, *13*(3), 108–117. https://doi.org/10.1109/MRA.2006.1678144

Barbier, S., Gosselin, F., & Balandier, P. (2008). Influence of tree species on understory vegetation diversity and mechanisms involved—A critical review for temperate and boreal forests. *Forest Ecology and Management*, *254*(1), 1–15. https://doi.org/10.1016/j.foreco.2007.09.038

Bauwens, S., Bartholomeus, H., Calders, K., & Lejeune, P. (2016). Forest Inventory with Terrestrial LiDAR: A Comparison of Static and Hand-Held Mobile Laser Scanning. *Forests*, *7*(12), 127. https://doi.org/10.3390/f7060127

Bienert, A., Georgi, L., Kunz, M., Maas, H.-G., & Von Oheimb, G. (2018). Comparison and Combination of Mobile and Terrestrial Laser Scanning for Natural Forest Inventories. *Forests*, *9*(7), 395. https://doi.org/10.3390/f9070395

Bienert, A., Georgi, L., Kunz, M., Von Oheimb, G., & Maas, H.-G. (2021). Automatic extraction and measurement of individual trees from mobile laser scanning point clouds of forests. *Annals of Botany*, *128*(6), 787–804. https://doi.org/10.1093/aob/mcab087

Bienert, A., Scheller, S., Keane, E., Mohan, F., & Nugent, C. (2006). *Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds*. *36*(In ISPRS workshop on laser scanning), 50–55.

Bienert, A., Scheller, S., Keane, E., Mullooly, G., & Mohan, F. (2012). Application of Terrestrial Laser Scanners for the Determination of Forest Inventory Parameters. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*(5), 1–5.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., & Kaasalainen, M. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution*, *6*(2), 198–208. https://doi.org/10.1111/2041-210X.12301

Chen, S., Liu, H., Feng, Z., Shen, C., & Chen, P. (2019). Applicability of personal laser scanning in forestry inventory. *PLOS ONE*, *14*(2), e0211392. https://doi.org/10.1371/journal.pone.0211392

Conto, T. de, Romain, J., Hamamura, C., & Marcozzi, A. (2019). *TreeLS: High performance R functions for forest data processing based on Terrestrial Laser Scanning (but not only) point clouds.* https://github.com/tiagodc/TreeLS

Dassot, M., Constant, T., & Fournier, M. (2011). The use of terrestrial LiDAR technology in forest science: Application fields, benefits and challenges. *Annals of Forest Science*, *68*(5), 959–974. https://doi.org/10.1007/s13595-011-0102-2

Donager, J. J., Sánchez Meador, A. J., & Blackburn, R. C. (2021). Adjudicating Perspectives on Forest Structure: How Do Airborne, Terrestrial, and Mobile Lidar-Derived Estimates Compare? *Remote Sensing*, *13*(12), 2297. https://doi.org/10.3390/rs13122297

Dupuy, J. M., & Chazdon, R. L. (2008). Interacting effects of canopy gap, understory vegetation and leaf litter on tree seedling recruitment and composition in tropical secondary forests. *Forest Ecology and Management*, *255*(11), 3716–3725. https://doi.org/10.1016/j.foreco.2008.03.021

Durrieu, S., Allouis, T., Fournier, R. A., Véga, C., & Albrech, L. (2008). *Spatial quantification of vegetation density from terrestrial laser scanner data for characterization of 3D forest structure at plot level*. SilviLaser.

Gollob, C., Ritter, T., & Nothdurft, A. (2020). Forest Inventory with Long Range and High-Speed Personal Laser Scanning (PLS) and Simultaneous Localization and Mapping (SLAM) Technology. *Remote Sensing*, *12*(9), 1509. https://doi.org/10.3390/rs12091509

Guan, H., Yu, Y., Ji, Z., Li, J., & Zhang, Q. (2015). Deep learning-based tree classification using mobile LiDAR data. *Remote Sensing Letters*, *6*(11), 864–873. https://doi.org/10.1080/2150704X.2015.1088668

Holopainen, M., Kankare, V., Vastaranta, M., Liang, X., Lin, Y., Vaaja, M., Yu, X., Hyyppä, J., Hyyppä, H., Kaartinen, H., Kukko, A., Tanhuanpää, T., & Alho, P. (2013). Tree mapping

using airborne, terrestrial and mobile laser scanning – A case study in a heterogeneous urban forest. *Urban Forestry & Urban Greening*, *12*(4), 546–553. https://doi.org/10.1016/j.ufug.2013.06.002

Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. (2014). Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing*, *80*(9), 863–872. https://doi.org/10.14358/PERS.80.9.863

Kükenbrink, D., Marty, M., Bösch, R., & Ginzler, C. (2022). Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. *International Journal of Applied Earth Observation and Geoinformation*, *113*, 102999. https://doi.org/10.1016/j.jag.2022.102999

Kukko, A., Kaijaluoto, R., Kaartinen, H., Lehtola, V. V., Jaakkola, A., & Hyyppä, J. (2017). Graph SLAM correction for single scanner MLS forest data under boreal forest canopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 199–209. https://doi.org/10.1016/j.isprsjprs.2017.09.006

Légaré, S., Bergeron, Y., & Paré, D. (2002). Influence of forest composition on understory cover in boreal mixedwood forests of western Quebec. *Silva Fennica*, *36*(1). https://doi.org/10.14214/sf.567

Li, S., Wang, T., Hou, Z., Gong, Y., Feng, L., & Ge, J. (2021). Harnessing terrestrial laser scanning to predict understory biomass in temperate mixed forests. *Ecological Indicators*, *121*, 107011. https://doi.org/10.1016/j.ecolind.2020.107011

Liang, X., Hyyppä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., Holopainen, M., Brolly, G., Francesco, P., Hackenberg, J., Huang, H., Jo, H.-W., Katoh, M., Liu, L., Mokroš, M., Morel, J., Olofsson, K., Poveda-Lopez, J., Trochta, J., … Wang, Y. (2018). International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, *144*, 137–179. https://doi.org/10.1016/j.isprsjprs.2018.06.021

Liang, X., Hyyppa, J., Kukko, A., Kaartinen, H., Jaakkola, A., & Yu, X. (2014). The Use of a Mobile Laser Scanning System for Mapping Large Forest Plots. *IEEE Geoscience and Remote Sensing Letters*, *11*(9), 1504–1508. https://doi.org/10.1109/LGRS.2013.2297418

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., & Vastaranta, M. (2016). Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 63–77. https://doi.org/10.1016/j.isprsjprs.2016.01.006

Liang, X., Kukko, A., Kaartinen, H., Hyyppä, J., Yu, X., Jaakkola, A., & Wang, Y. (2014). Possibilities of a Personal Laser Scanning System for Forest Mapping and Ecosystem Services. *Sensors*, *14*(1), 1228–1248. https://doi.org/10.3390/s140101228

Limongiello, M., Ronchi, D., & Albano, V. (2020). *BLK2GO for DTM generation in highly vegetated area for detecting and documenting archaeological earthwork anomalies*.

Liu, G., Wang, J., Dong, P., Chen, Y., & Liu, Z. (2018). Estimating Individual Tree Height and Diameter at Breast Height (DBH) from Terrestrial Laser Scanning (TLS) Data at Plot Level. *Forests*, *9*(7), 398. https://doi.org/10.3390/f9070398

Lovell, J. L., Jupp, D. L. B., Newnham, G. J., & Culvenor, D. S. (2011). Measuring tree stem diameters using intensity profiles from ground-based scanning lidar from a fixed viewpoint. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 46–55. https://doi.org/10.1016/j.isprsjprs.2010.08.006

Maas, H. -G., Bienert, A., Scheller, S., & Keane, E. (2008). Automatic forest inventory parameter determination from terrestrial laser scanner data. *International Journal of Remote Sensing*, *29*(5), 1579–1593. https://doi.org/10.1080/01431160701736406

Moir, W. H. (1966). Influence of Ponderosa Pine on Herbaceous Vegetation. *Ecology*, *47*(6), 1045–1048. https://doi.org/10.2307/1935654

Murgoitio, J. J., Shrestha, R., Glenn, N. F., & Spaete, L. P. (2013). Improved visibility calculations with tree trunk obstruction modeling from aerial LiDAR. *International Journal of Geographical Information Science*, *27*(10), 1865–1883. https://doi.org/10.1080/13658816.2013.767460

Murphy, G. E., Acuna, M. A., & Dumbrell, I. (2010). Tree value and log product yield determination in radiata pine (*Pinus radiata*) plantations in Australia: Comparisons of terrestrial laser scanning with a forest inventory system and manual measurements. *Canadian Journal of Forest Research*, *40*(11), 2223–2233. https://doi.org/10.1139/X10-171

Olofsson, K., Holmgren, J., & Olsson, H. (2014). Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm. *Remote Sensing*, *6*(5), 4323–4344. https://doi.org/10.3390/rs6054323

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., Boissieu, F. de, & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, *251*. https://doi.org/10.1016/j.rse.2020.112061

Runkle, J. R. (1982). Patterns of Disturbance in Some Old-Growth Mesic Forests of Eastern North America. *Ecology*, *63*(5), 1533–1546. https://doi.org/10.2307/1938878

Ryding, J., Williams, E., Smith, M., & Eichhorn, M. (2015). Assessing Handheld Mobile Laser Scanners for Forest Surveys. *Remote Sensing*, *7*(1), 1095–1111. https://doi.org/10.3390/rs70101095

Schneider, F. D., Kükenbrink, D., Schaepman, M. E., Schimel, D. S., & Morsdorf, F. (2019). Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR. *Agricultural and Forest Meteorology*, *268*, 249–257. https://doi.org/10.1016/j.agrformet.2019.01.033

Stovall, A. E. L., Vorster, A. G., Anderson, R. S., Evangelista, P. H., & Shugart, H. H. (2017). Non-destructive aboveground biomass estimation of coniferous trees using terrestrial LiDAR. *Remote Sensing of Environment*, *200*, 31–42. https://doi.org/10.1016/j.rse.2017.08.013

Suchar, V. A., & Crookston, N. L. (2010). Understory cover and biomass indices predictions for forest ecosystems of the Northwestern United States. *Ecological Indicators*, *10*(3), 602–609. https://doi.org/10.1016/j.ecolind.2009.10.004

Tansey, K., Selmes, N., Anstee, A., Tate, N. J., & Denniss, A. (2009). Estimating tree and stand variables in a Corsican Pine woodland from terrestrial laser scanner data. *International Journal of Remote Sensing*, *30*(19), 5195–5209. https://doi.org/10.1080/01431160902882587

Vauhkonen, J., Maltamo, M., McRoberts, R. E., & Næsset, E. (2014). Introduction to Forestry Applications of Airborne Laser Scanning. In M. Maltamo, E. Næsset, & J. Vauhkonen (Eds.), *Forestry Applications of Airborne Laser Scanning* (Vol. 27, pp. 1–16). Springer Netherlands. https://doi.org/10.1007/978-94-017-8663-8_1

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., Bartholomeus, H., Brede, B., & Herold, M. (2017). Data acquisition considerations for Terrestrial Laser Scanning of forest plots. *Remote Sensing of Environment*, *196*, 140–153. https://doi.org/10.1016/j.rse.2017.04.030

Wing, B. M., Ritchie, M. W., Boston, K., Cohen, W. B., Gitelman, A., & Olsen, M. J. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment*, *124*, 730–741. https://doi.org/10.1016/j.rse.2012.06.024

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G. (2016). An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sensing*, *8*(6), 501. https://doi.org/10.3390/rs8060501

Zhao, K., García, M., Liu, S., Guo, Q., Chen, G., Zhang, X., Zhou, Y., & Meng, X. (2015). Terrestrial lidar remote sensing of forests: Maximum likelihood estimates of canopy profile,

leaf area index, and leaf angle distribution. *Agricultural and Forest Meteorology*, *209–210*, 100–113. https://doi.org/10.1016/j.agrformet.2015.03.008

Zong, X., Wang, T., Skidmore, A. K., & Heurich, M. (2021). The impact of voxel size, forest type, and understory cover on visibility estimation in forests using terrestrial laser scanning. *GIScience & Remote Sensing*, *58*(3), 323–339. https://doi.org/10.1080/15481603.2021.1873588

CHAPTER 6 - Synthesis

In this dissertation, I examined the spatial distribution, regeneration dynamics, and environmental influence on pine communities in the southern Appalachian region using three different forms of LiDAR data including aerial, terrestrial and mobile LiDAR with field based ecological measurements. In Chapter 2, I developed and validated a machine learning framework to detect and classify pine species from publicly available low-density aerial LiDAR data (sourced from USGS 3DEP LiDAR explorer). I applied three classification algorithms including 2D-CNN, RF, and SVM to distinguish pine and non-pine trees. Among these, I found that 2D-CNN achieved the highest accuracy across all performance metrics (precision, recall, and F1-score) effectively classifying both pine and non-pine species. This work demonstrates that, despite inherent limitations of low density LiDAR, it is still possible to derive coarser level classification when paired with appropriate machine learning techniques.

Building upon the classified pine locations from Chapter 2, I analyzed the spatial arrangement of pine and examined the role of environmental gradients in shaping the pine distribution in Chapter 3. I showed that environmental variables including elevation, slope, aspect and TPI had significant influence on the heterogeneous distribution of pine in the study area, while the effects differ between the variables. Even after accounting for these environmental effects, pines were still clustered at small spatial scale suggesting that microsite conditions, limited seed dispersal, and potential facilitative interactions among conspecifics may contribute to aggregation. When stratified by height class (short, intermediate, and tall), all classes exhibited small scale clustering however, only short and intermediate trees showed a positive spatial association with each other, while all other pairwise combinations displayed negative associations.

In Chapter 4, I showed the mismatch between overstory dominance and understory regeneration by analyzing sapling and seedling distributions alongside canopy structure and environmental variables. While pine species including shortleaf pine, pitch pine, and eastern white pine dominate the overstory, their regeneration was virtually absent. In contrast mountain laurel dominated the understory reducing light availability and inhibiting pine seedling establishment. I also found that while the overstory species composition was more influenced by the abiotic variables such as slope, TPI, and aspect, the understory composition was more related to overstory and understory canopy openness. Importantly, I also examined tree establishment, release events and decadal growth trends. Pitch pine, shortleaf pine, oak, and maple exhibited peak establishment in 1950 corresponding with the time when Dupont Corporation acquired 4000 hectares for recreational purposes for their employees. I found that the release frequency increased and was the highest for 1990 where over 30% of trees exhibited some form of release. Additionally, there was a significant difference among growth of all species. All the pine species had higher growth in first or second decade meaning they established in open conditions. However, in the later decade their growth rate reduced indicating increased in competition and reduced disturbance. In contrast, while red maple and oaks had lower growth during the first and second decade, their growth rate were higher than of shortleaf and pitch pine in the later decades. This indicated that, these species grows consistently over time and are better able to persist and compete under closed-canopy conditions.

In Chapter 5, I evaluated how ground based laser scanner including both terrestrial and mobile platforms, perform under varying understory vegetation density with a focus on stem detection and DBH estimation. I showed that the understory vegetation density strongly influenced the stem detection rate where the mean detection rate reduced with the increase in

understory density for both MLS and TLS. While the understory density significantly affected the DBH estimated for the terrestrial LiDAR such effect was not significant for mobile LiDAR

The overarching goal of this research was not only to assess the current status and drivers of pine regeneration but also to develop scalable and cost effective methodologies for mapping pine populations in forested landscapes where low density aerial LiDAR are available. By augmenting this with ground-based LiDAR to assess the understory complexity, this approach provides comprehensive and replicable toolset for forest managers aiming to assess pine status across large, structurally complex landscapes.

From an ecological standpoint, the result affirms that the pine ecosystem in this region is disturbance dependent. Without fire, reduction of stand density or some other forms of disturbances, pine recruitment is likely to continue declining which not only threatens the sustainability of these ecosystems but also the ecological functions and biodiversity associated with it. Forest managers should implement restoration strategies such as prescribed burning, thinning, and planting in areas where pine presence is declining, ensuring that the planting designs reflects the natural spatial patterns observed in this study characterized by clustered and random distributions to better restore the historical stand structure and spatial heterogeneity.

This dissertation enhances our understanding of pine regeneration dynamics by integrating remote sensing, field ecology, and spatial analysis, thus providing valuable tools for monitoring and management. Observations of spatial clustering in pine, the disconnection between overstory and understory composition, and the reduction in growth rates over decades all suggest a transitioning system. If proactive measures are not taken, the southern Appalachian pine forests will continue to shift compositionally toward mesophytic hardwoods and thick shrub

layers, leading to significant ecological repercussions. The combined methods and findings discussed here present a strategy for both comprehending and altering this trend.

APPENDICES

S.2 CHAPTER2 – SUPPLEMENTARY MATERIAL

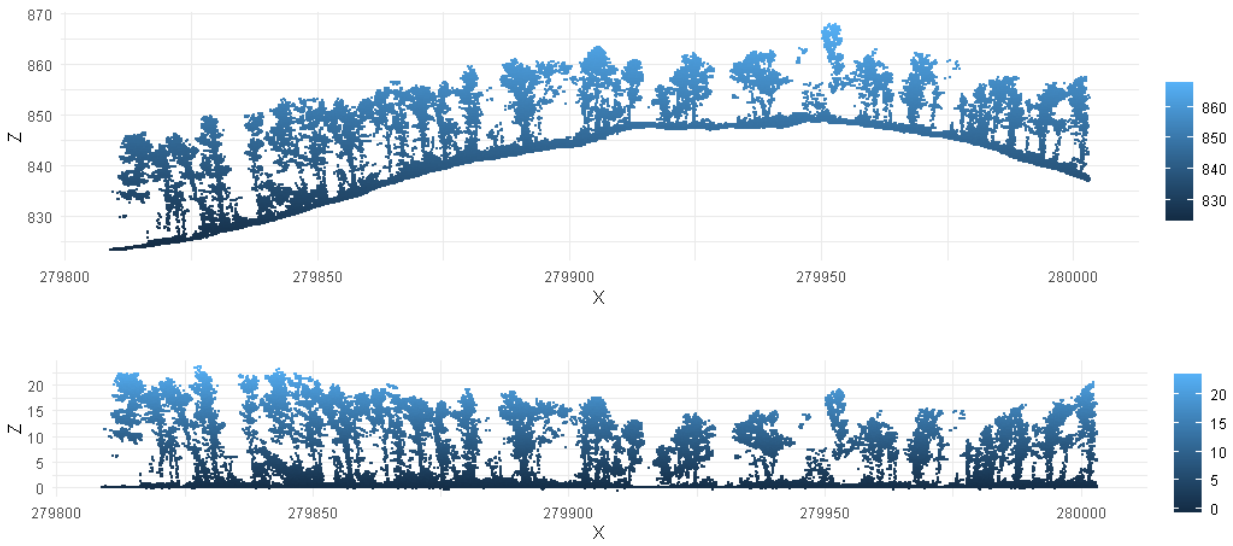

Figure S.2.1: Raw aerial LiDAR data (top) vs height normalized data (bottom).

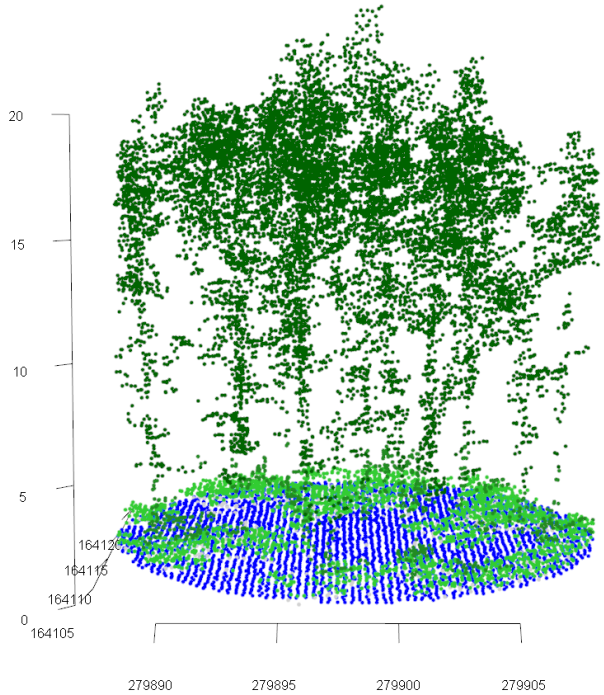

Figure S.2.2: Height normalized aerial LiDAR data. Blue represents ground points.

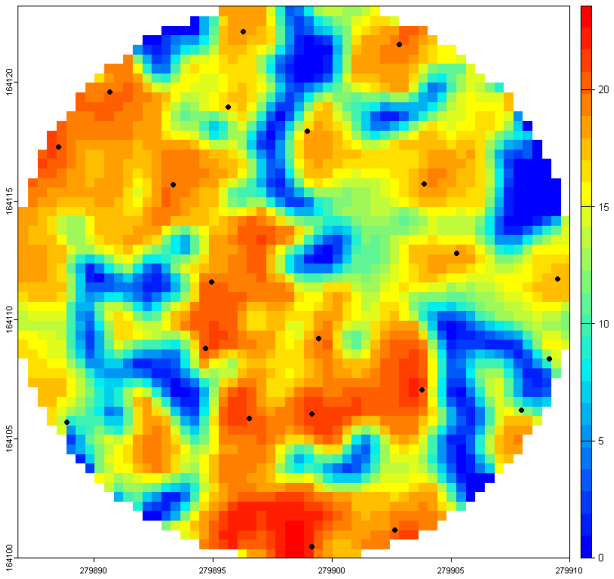

Figure S.2.3: CHM derived from the height normalized aerial LiDAR data. Black dots represent individual tree tops.

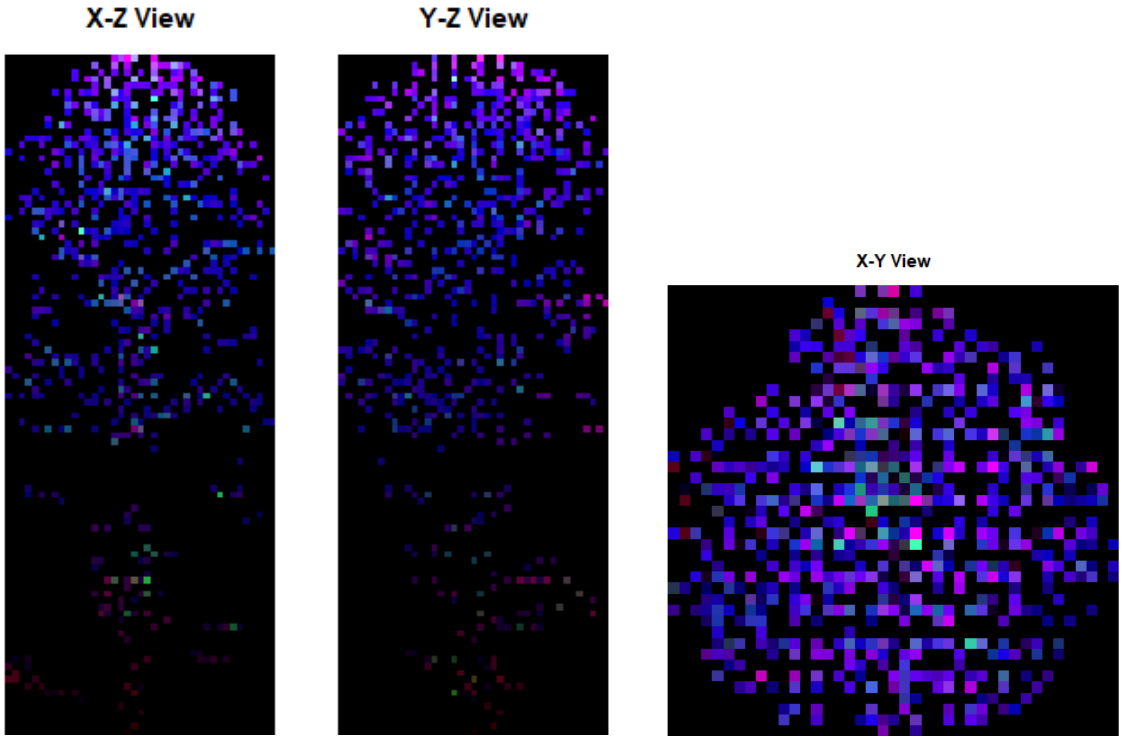

Figure S.2.4: LiDAR data converted to 2D rasters with different view perspective. Each pixel (0.1m resolution) represents mean point density, mean intensity, and mean elevation values.

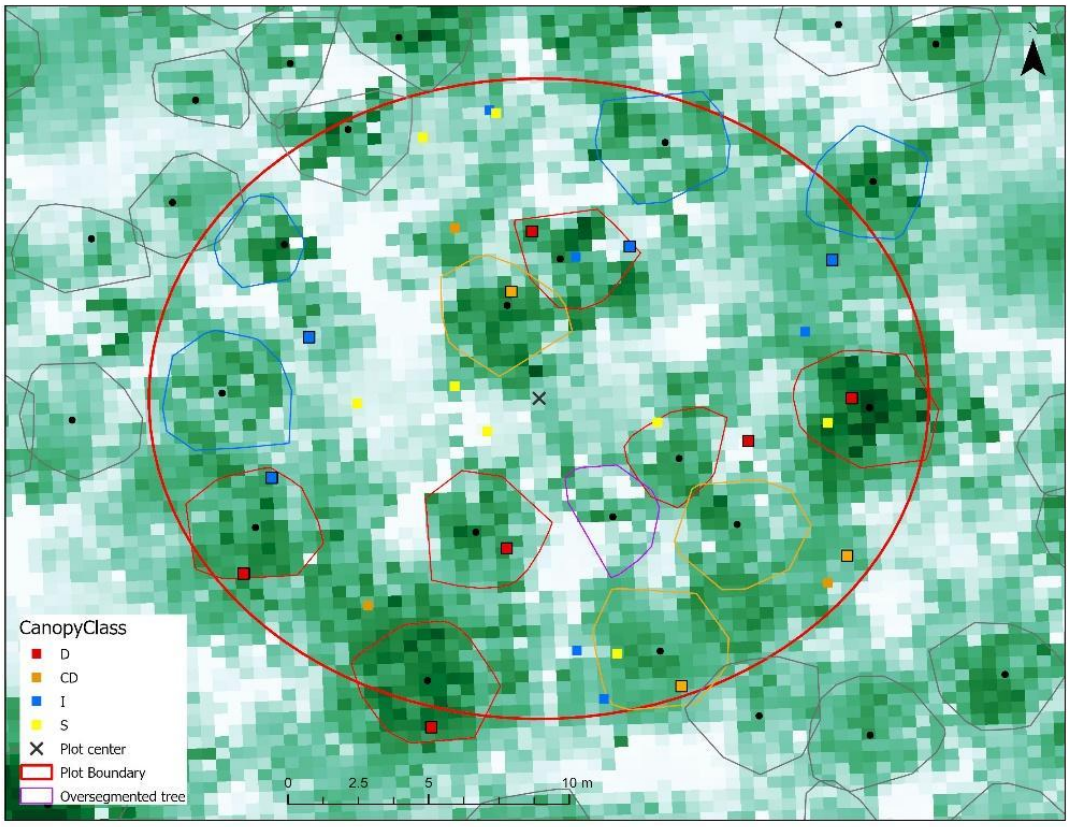

Figure S.2.5: Comparison of individual tree segmentation from aerial LiDAR data with field plot data. Black points represent individual tree locations derived from aerial LiDAR. Field mapped trees are color-coded by canopy class, with black-bordered points representing trees successfully matched to the LiDAR tree locations, while points without the border indicate unmatched trees.

S.3 CHAPTER3 – SUPPLEMENTARY MATERIAL

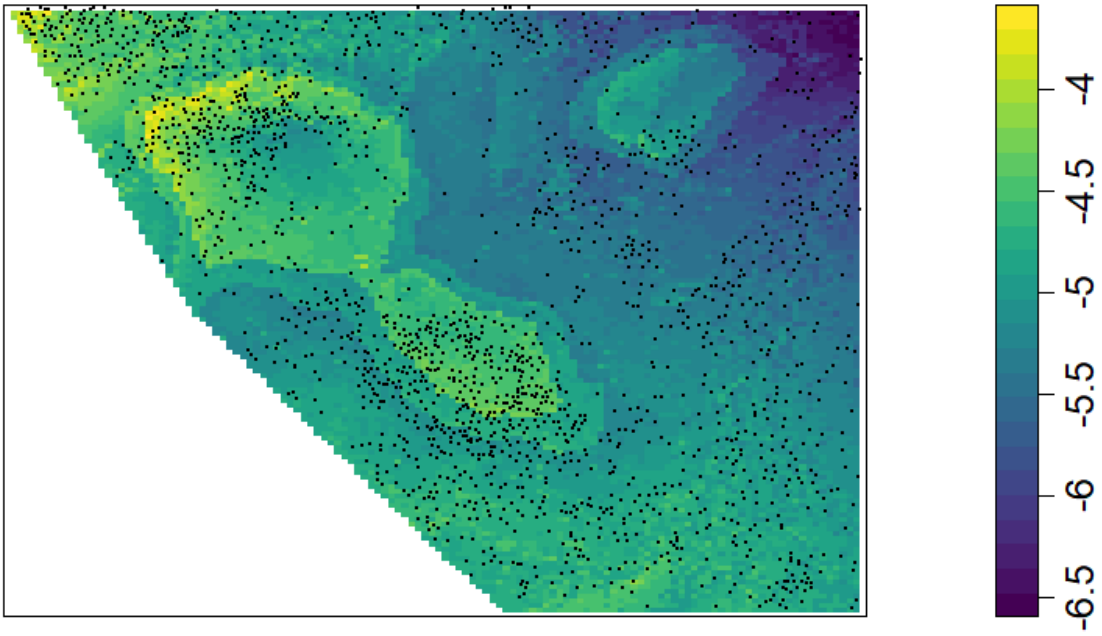

Figure S.3.1: Predicted pine intensity with observed pine locations. The area with brighter color represents the area where the IPP model predicted to have higher pine density, while the darker color areas are the ones where the model predicted to have lower pine density.

S.4 CHAPTER 4 – SUPPLEMENTARY MATERIAL

Table S.4.1: Summary statistics of the environmental variables used in the ordinations.

Variables	**Mean ± SE**
Elevation (m)	817.76 ± 8.69
Slope (degree)	10.58 ± 0.92
OS canopy openness (%)	11.87 ± 1.36
US canopy openness (%)	33.53 ± 3.55
Total Exchange Capacity (meq/100g)	3.98 ± 0.75
Soil pH	4.3 ± 0.07
Organic Matter (%)	8.67 ± 1.20
S (ppm)	20.56 ± 2.43
P (mg/kg)	6.19 ± 0.76
Ca (mg/kg)	192.06 ± 51.76
Mg(mg/kg)	23.69 ± 2.00
K (mg/kg)	33.06 ± 2.00
Na (mg/kg)	11.63 ± 0.36

Table S.4.1(continued):

Fe (mg/kg)	168.19 ± 12.18
Mn (mg/kg)	9.5 ± 2.35
Zn (mg/kg)	1.43 ± 0.11
AI (mg/kg)	1282.06 ± 72.56

Table S.4.2: Tukey HSD result for each species. In the table, FD represents the first decade, SD is the second decade, PD is the penultimate decade, and LD is the last decade. Diff is the difference in growth between two decades in mm.

Decade	**Species/Genus**									
	Pitch pine		**Shortleaf pine**		**E. white pine**		**Oaks**		**Red maple**	
	Diff	p-value	Diff	p-value	Diff	p-value	Diff	p-value	Diff	p-value
SD-FD	-0.94	< 0.001	-0.69	< 0.001	-0.67	0.41	-0.27	0.33	0.25	0.75
PD-FD	-1.89	< 0.001	-1.71	< 0.001	-1.46	0.01	-1.07	< 0.001	-0.27	0.70
LD-FD	-1.90	< 0.001	-2.03	< 0.001	-2.03	< 0.001	-1.09	< 0.001	-0.59	0.10
PD-SD	-0.95	< 0.001	-1.01	< 0.001	-0.80	0.25	-0.81	< 0.001	-0.52	0.16
LD-SD	-0.96	< 0.001	-1.33	< 0.001	-1.37	0.01	-0.82	< 0.001	-0.84	0.005
LD-PD	-0.01	0.99	-0.32	0.16	-0.57	0.55	-0.01	0.99	-0.31	0.57

S.5 CHAPTER 5 – SUPPLEMENTARY MATERIAL

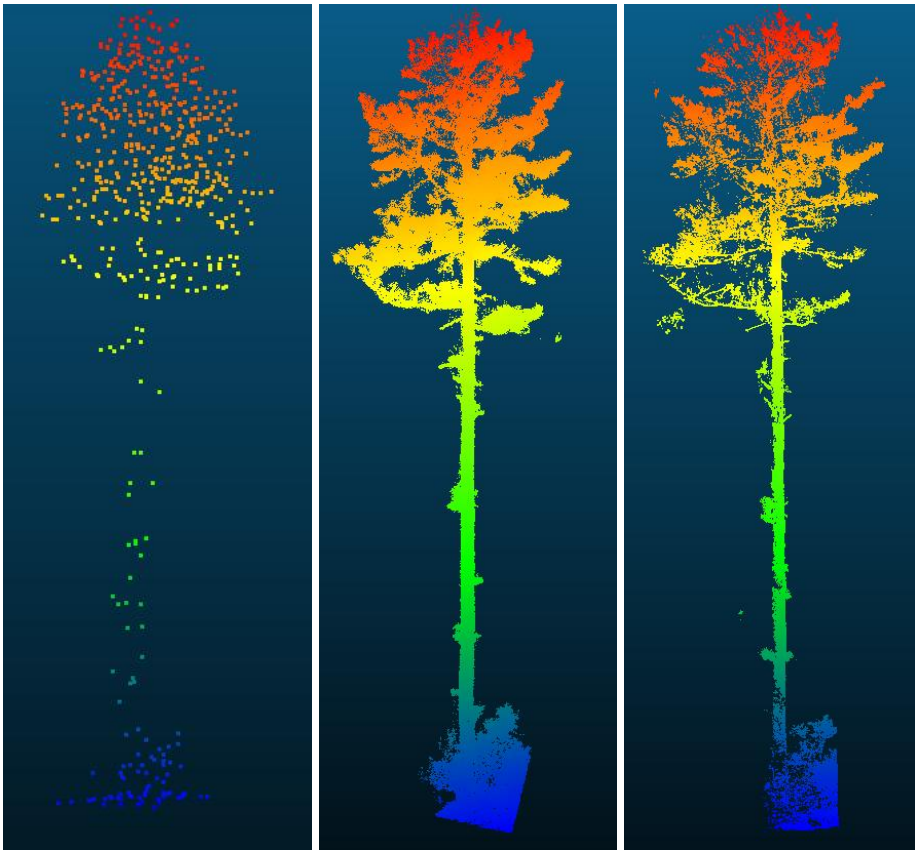

Figure S.5.1**:** Comparison between data collected by left) aerial, middle) mobile & right) terrestrial LiDAR