8-8-2023

Modeling the effect of neighborhood competition on tree diameter growth in the Pacific Northwest Coast Range

Htet Lin Naing
Mississippi State University, hn326@msstate.edu

Modeling the effect of neighborhood competition on tree diameter growth in the Pacific Northwest Coast Range

Htet Lin Naing

Approved by:

Austin Himes (Major Professor)
Krishna P. Poudel
Heidi J. Renninger (Committee Member/Graduate Coordinator)
L. Wes Burger (Dean, College of Forest Resources)

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Forestry
in the Department of Forestry

Mississippi State, Mississippi

August 2023

Name: Htet Lin Naing

Date of Degree: August 8, 2023

Institution: Mississippi State University

Major Field: Forestry

Major Professor: Austin Himes

Title of Study: Modeling the effect of neighborhood competition on tree diameter growth in the Pacific Northwest Coast Range

Pages in Study 44

Candidate for Degree of Master of Science

Abstract

Trees compete for various resources such as sunlight, water, and nutrients, which can be expressed as numerical terms, called competition indices (CI). Competition between individual trees is correlated with their growth and mortality. Therefore, CIs are used as independent variables to develop, improve and modify growth and yield models. This study was conducted to test the effect of neighborhood competition on tree diameter growth among Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), western hemlock (*Tsuga heterophylla* (Raf.) Sarg) and red alder (*Alnus rubra* Bong.), in the Pacific Northwest Coast Range, USA. After testing seven distance-independent CIs and three distance-dependent CIs, only the distance-independent CIs were found to significantly affect the diameter growth model. Among them, CIs with basal area and diameter information were the most impactful. As a result, a simple CI was very effective in a model that accounts for the basal area information of different tree species.

ACKNOWLEDGEMENTS

First of all, I am grateful to my mother for her continuous support and for believing in me whatever I do. Being a first-generation forester, forestry and environmental subjects were new to her. However, since my first degree which was also forestry, she trusted in me without knowing what I would become after earning a forestry degree. When I explained what I learned and the research I was working on in graduate school, she was way happier than me. As a Buddhist who believes in “life-after-death”, she said that working for our environment is noble, and I would be in a better life after death because of this contribution.

I would like to extend my sincerest gratitude to the United States Agency for International Development (USDA) – Lincoln Scholarship Program for their financial provision to study this research-based master’s degree program in Forestry. I also thank to my English teacher and mentor, Daw Zune Zune Pwint Phyu, for encouraging me to apply for international scholarships and study in the US. Her support was very instrumental not only in academic but also in my lifelong pathways. Then, I would like to thank my referees to apply for this scholarship, Dr. Htun Thein, Dr. Idd Idd Shwe Zin and Dr. Daniel Spencer.

Most importantly, this research work could not have been done without the immense support of my advisor, Dr. Austin Himes, and the constructive suggestions from my research committee members, Dr. Heidi Renninger and Dr. Krishna Poudel. After changing from my first research to the current project for some reason, my advisor introduced me to an interesting data set that he had, and it led to this research topic. Weekly advising meeting with him and steady work progress have turned into the completion of my thesis.

Last but not least, my life in the graduate school at Mississippi State University was very joyous and less stressful because of my colleagues. Both international and domestic students were very warm and welcoming to me. I especially thank to Dr. Thu Ya Kyaw, Casey Iwamoto, Mahesh Tiwari, Nitant Rai, Elizabeth Baach, Jianing Liang, and Macy Gosselaar.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Introduction

Trees compete for resources, like light and water, as they become limiting on a site. The impact of competition between trees can be represented in growth and yield models by competition indices, which are numerical expressions of the degree of growth potential that is limited by other trees (Kahriman et al., 2018) that compete for resources such as water, sunlight and nutrients (Kocher & Harris, 2007). Competition indices are correlated with annual diameter and height growth, as well as with mortality due to competition with neighboring trees, which is why simple competition indices are well suited to tree growth simulation models and field applications (Daniels, 1976). Performance of multiple linear regression models to predict growth of individual trees work well when different forms of competition indices are used as independent variables in combination with stand-level variables (Tomé & Burkhart, 1989). Many competition indices have been proposed in the literature. However, the effectiveness of different competition indices depends on stand conditions (Kahriman et al., 2018). Therefore, it is a challenge to find the optimal competition index that satisfies every tree species in different yield studies (Pretzsch, 2009).

Competition is an interaction between different individuals, which results in a decline in the survival, growth increment and reproduction among them (Begon et al., 1996). Competition is a fundamental ecological process that plays a major role in population dynamics, survival, growth and species replacement (Peet & Christensen, 1987). In the field of ecology and forestry, there are several studies to develop, improve or modify various models of stand dynamics by including competition indices as an independent variable in projecting individual-tree growth (Kahriman et al., 2018; Maleki et al., 2015). Competition among different trees plays a significant role in determining diameter growth, and the contribution of competition indices in diameter growth models is meant to account for the competition status of the target tree relative to its neighboring trees (Kahriman et al., 2018).

The effectiveness of a competition index in growth models depends on tree species, accessible data, and the structure of the selected model (Biging & Dobbertin, 1995; Kahriman et al., 2018; Tomé & Burkhart, 1989). Competition indices are categorized into two groups: distance-independent indices (Biging & Dobbertin, 1995; Crookston, 1982; Lorimer, 1983a; Schröder & Gadow, 1999) and distance-dependent indices (Alemdag, 1978; Bella, 1971; Hegyi, 1974; Kahriman et al., 2018; Martin & Ek, 1984).

Distance-independent competition indices are functions of stand-level variables, or of the initial dimensions of the trees, but do not require information of the spatial arrangement of trees or their proximity to each other (Kahriman et al., 2018; Maleki et al., 2015). A direct relationship of a tree with its neighbors can be determined by comparing the target tree with the sizes of all other trees within the stand without knowing its location within the stand or relative to its neighbors (Kahriman et al., 2018). Therefore, distance-independent competition indices are effective when the exact location of the target tree, relative to neighbors, is unknown.

Distance-dependent competition indices are usually computed as functions of the initial dimensions of the target tree, as well as the distance, dimensions, numbers, and sometimes spatial arrangement of the neighboring competitor trees (Kahriman et al., 2018). The approaches to quantify the level of competition in distance-dependent indices include influence-zone overlap indices or crown area overlap indices, and distance-weighted, size ratio approaches (Kahriman et al., 2018). Size-ratio indices are formulated as the sum of the ratios of the dimensions (e.g., diameter at breast height (dbh), tree height, and basal area) of the target tree to the dimensions of the nearby trees. Distance-weighted, size-ratio approaches are similar, but weight size ratios of neighboring trees by their distance from the target tree when summing (Kahriman et al., 2018). These types of distance-dependent indices are based on the hypothesis that neighboring trees with larger dimensions at closer distances to the target tree are more competitive (Kahriman et al., 2018). In this case, competitor trees should fall in the boundary of a circle with a fixed radius or a certain number of trees that are closest to the target tree (Kahriman et al., 2018).

When comparing effectiveness of distance-dependent competition indices relative to distance-independent indices for the individual-tree basal area growth models of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) and western hemlock (*Tsuga heterophylla* (Raf.) Sarg), the addition of spatial index only gave a small (less than 0.01) increase in adjusted coefficient of determination (R^2_{Adj}) compared with models that only included distance-independent competition indices (Wimberly & Bare, 1996). Therefore, developing growth models for these type of forests by using distance-related stand data is inefficient because of the additional effort and expense needed to collect spatially referenced information (Wimberly & Bare, 1996). In other research, it was found that spatial information does not correspond to the basal area growth when the maximum growing space is reached (Radtke et al., 2003). Therefore, distance information is not often deemed worth the effort of collecting and distance-independent competition indices are considered more practical for growth modeling.

Fundamentally, competition among forest trees can be categorized into two groups: competition among *conspecific* individuals, plants of the same species, and *heterospecific* individuals, plants of different species (Maleki et al., 2015). The former is also called intraspecific competition, whereas the latter is termed interspecific competition (Maleki et al., 2015). Productivity of mixed-species forests has frequently been reported to differ from the growth of monoculture stands, suggesting different levels of intraspecific or interspecific competition among them (Brunner & Forrester, 2020). The differences in growth between monocultures and mixed-species stands depends on the tree species involved and other factors such as site quality and stand structure, and can change over time (Brunner & Forrester, 2020; Forrester, 2014; Forrester & Bauhus, 2016). However, most forest growth and yield models are based on dynamics in monoculture forests. Therefore, there is a substantial need to expand the quantitative empirical knowledge of competition and stand density effects to mixed-species forests in order to manage them more effectively (Brunner & Forrester, 2020).

In the Pacific Northwest region, Douglas-fir is the predominant commercial timber species (Cole & Newton, 2023). Red alder (*Alnus rubra* Bong.) is one of the most common species to appear after timber harvest (Trappe et al., 1968); it is usually considered less commercially desirable than Douglas-fir, but does have substantial market value. Western hemlock is another important commercial species across the Oregon Coast Range. It is more shade tolerant, compared to Douglas-fir, and is a preferred option in the western Oregon Coast Range where Douglas-fir growth is drastically reduced by Swiss needle cast (SNC) (*Phaeocryptopus gaeumannii*) (Cole & Newton, 2023; Zhao et al., 2014). Due to its economic profitability, Douglas-fir was the focus species in the coastal fog belt in the latter half of the twentieth century, and it led to the current epidemic of SNC, which is a foliar disease of

Douglas-fir caused by the endemic ascomycete (*Phaeocryptopus gaeumanii* (T. Rohde) Petr) (Hansen et al., 2000). Largely because of the impacts of SNC in the Oregon Coast Range, it is common that a mixture of conifer and hardwood species is maintained by forest managers (Cole & Newton, 2023).

This region in the Pacific Northwest is very suitable for almost any vegetation type, including conifers, hardwoods, shrubs, and herbaceous plants (Cole & Newton, 2023). Average annual precipitation in this region ranged from 180-320 cm $year^{-1}$, temperature is between 7 and 11 °C, and the soil is very well drained (Himes & Puettmann, 2020). Among the popular species in this region, alder species gained attention as sources of nitrogen, however competitive interactions between red alder and Douglas-fir often result in a growth decline of Douglas-fir (Cole & Newton, 2023). Regarding western hemlock, it is more tolerant to competition from a partial cover of shrub or hardwood, compared to Douglas-fir (Cole & Newton, 2023).

Few studies have investigated interactions among western hemlock, Douglas-fir and red alder, regarding how competition among them affects their growth (Binkley, 2003; Briggs et al., 1978; Cortini & Comeau, 2008a; Fang et al., 2019; Radosevich et al., 2006; Shainsky & Radosevich, 1992; Zhao et al., 2014). A study by Zhao et al. (2014) found that the decline in Douglas-fir trees due to SNC resulted in higher growth increment for western hemlock in mixed forest stands within 29 km of the Pacific Ocean, compared to Douglas-fir monocultures. In research about the effects of competition from red alder and paper birch (*Betula papyrifera* Marsh.) on the growth of western hemlock, Douglas-fir and western redcedar (*Thuja plicata* Donn), significant reductions in stem volume of Douglas Fir and western redcedar occurred after red alder's density exceeded 500 stems per hectare, however when the red alder density was lower, the growth of intermixed conifers was higher (Cortini & Comeau, 2008b). Different responses between pure stands and mixed stands of red alder and conifer trees, in low nitrogen soil (Wind River, Washington) and in rich nitrogen soil (Cascade Head, Oregon), were also observed by Binkley (2003). Binkley (2003) showed that in low nitrogen soil, the pure conifer stand had lower stem biomass and net increment compared to the conifer stem mass in mixed plantations according to remeasurement after 70 years of stand development. But, contrasting results were observed in nitrogen-rich soil because the stem mass of the pure conifer plot was greater than the total amount of both conifer and alder in the mixture (Binkley, 2003). Considering multiple ecosystem services, including aboveground biomass, as a result of the combination among these three species, mixtures of different tree species, Douglas-fir, western hemlock and red alder, produce more ecosystem services compared to single-species plantations, and crown length and understory vegetation diversity are also greater in the former (Himes and Puettmann, 2020; Himes et al., 2020).

Objectives, and hypotheses

There are three objectives of this study and I aim to achieve these objectives by testing a specific hypothesis for each of them. The objectives and their associated hypothesis are as follows:

Objective (1): Determine the contribution of competition indices towards predicting tree diameter growth in even-aged forests across a range of different species compositions in the Coast Range of Oregon and Washington.

> Hypothesis (1): Diameter growth models will perform better when they include a competition index as an independent variable.

Objective (2): Compare the performance of distance-dependent and distance-independent competition indices and select the best competition index.

Hypothesis (2): Distance-independent competition indices will contribute to better predictions of target tree diameter growth than distance-dependent competition indices.

Objective (3): Compare observed growth rate and model predicted growth rate.

Hypothesis (3): The best model with the best competition index will be a good predictor of observed diameter growth.

CHAPTER II

MATERIALS AND METHODS

Study area and data

The study area is in the Coast Range mountains of Northwest Oregon and Southwest Washington, USA, close to the mouth of the Columbia River, and specifically within ~70,000-ha industrial plantations on Lewis and Clark Timberlands (LCT), which were managed by GreenWood Resources (GWR) (Himes & Puettmann, 2020). Elevation varies between sea level and 1,000 m above sea level (Himes & Puettmann, 2020). Average annual temperature and rainfall are 7-11 °C and 180-320 cm year^{-1}, respectively (Himes & Puettmann, 2020). Regarding the soil class, it possesses igneous and sedimentary soil types in origin, which are very well drained with high water retaining capacity (USDA, 2018).

The majority of LCT lands have been managed with the purpose of commercial timber production for two rotations at minimum (Himes & Puettmann, 2020). Although ownership of the property has changed repeatedly, management practices at the time of this study included chemical weed control during the initial two years after harvest and planting at a density of 890-1075 trees ha^{-1}, pre-commercial thinning at the approximate age of 15, and clear-cutting around age 45 (Himes & Puettmann, 2020). In the study site, western hemlock, Sitka spruce (*Picea stichensis* (Bong.) Carriere), western redcedar, and Douglas-fir were planted, while natural regeneration of western hemlock and red alder from windblown seeds were also added to the area (Himes & Puettmann, 2020). Consequently, the area includes a landscape of even-aged stands found in a mosaic of species mixtures and monocultures among western hemlock, Douglas-fir, Sitka spruce, red alder, and western redcedar (Himes & Puettmann, 2020).

In this study, I leverage information gathered as part of a previous project investigating tree species diversity and composition related to biomass, understory community, and crown architecture in production plantations of the coastal Pacific Northwest, USA (see Himes & Puettmann, 2020 and Himes et al., 2020). The previous study specifically focused on three species: western hemlock, Douglas-fir, and red alder. Sampling was conducted in seven species compositions: 1. Western hemlock monoculture (WH), 2. Douglas-fir monoculture (DF), 3. red alder monoculture (RA), 4. western hemlock and Douglas-fir mixture (WH & DF), 5. western hemlock and red alder mixture (WH & RA), 6. Douglas-fir and red alder mixture (DF&RA), and 7. western hemlock, Douglas-fir and red alder mixture (WH, DF & RA). In this trial, monoculture means plots with more than 90% of all tree stems of the same species (Himes & Puettmann, 2020). Two-species plots had an approximate proportion of 30-70% for each target species, and three-species plots were composed of 25-50% each, whereas a proportion of non-target species is allowed up to 5% (Himes & Puettmann, 2020). Details of all possible combinations of the three selected tree species and their monocultures are described in Table 1.

Table 1 Species composition criteria of plots, with proportions based on stem counts.

Composition criteria	DF	WH	RA	DFWH	DFRA	WHRA	DFWHRA
Tree species in plot	DF only	WH only	RA only	DF and WH	DF and RA	WH and RA	DF, WH and RA
Maximum % single species	100	100	100	70	70	70	50
Minimum % single species	90	90	90	30	30	30	25

Note: DF, Douglas-fir; WH, western hemlock; RA, red alder; NA, not applicable

Data were collected in the summer of 2017 (Himes & Puettmann, 2020). As per Table 1, combinations of the three desired tree species resulted in seven different species compositions, including monocultures and all possible species mixtures. As this study was carried out in existing mature, operational stands, the species composition covers varieties of management options and environmental conditions that produced the mixtures of tree species included (Himes & Puettmann, 2020). Nine replicates were accounted for in this study, totaling 63 plots (9 x 7). The shape of the plots was circular, with radii of 10 m each (area of 314 m^2) (Himes & Puettmann, 2020).

There are two categories of trees: target trees and neighboring trees, with the assumption that the overall influence of neighboring trees has an impact on the growth of target trees. One tree of each species that was closest to the plot center was identified as a target tree whereas all other trees with crowns touching or overlapping with the crown of the target tree were considered neighboring trees. Since crown closure had occurred in all stands, and trees with crowns directly connected to each other would have a direct competition, this approach to identifying neighboring trees was applied. This approach helped to quickly assess whether trees were interacting or not in the field. Distance between two trees alone could not be relied upon because trees were not distributed geometrically (e.g., in easily identifiable lines or consistently spaced), and precise stem mapping was too time intensive. From each target tree, two 5mm tree cores were taken at breast height (dbh, 1.4m) on the uphill side of the stem and at 90 degrees from one tree of each target species, and the tree number and distance to all trees whose canopies overlapped or touched the target tree were also recorded. All tree cores of the target trees were mounted, sanded, and viewed under a microscope. To calculate annual growth, annual tree rings were identified and the width, which is the distance between rings, was measured. The average ring width of the two cores per target tree was used to estimate diameter growth. In this study, 72 target trees and 399 neighboring trees were included in total.

Exploratory data analysis

Average annual growth over the previous ten years, determined from the tree rings, was the primary response variable (eq.1). Information of eight other independent variables was also collected. They are as follows: species identity of target trees, the number of neighboring trees per plot, dbh of target trees, basal area of all neighboring Douglas-fir trees per target tree, basal area of all neighboring western hemlock trees per target tree, basal area of all neighboring red alder trees per target tree, site index, and total basal area per plot. "Y" (eq. 2) is the average annual growth rate expressed as a proportion of the starting diameter. I have applied two major types of regression models, ordinary least squares (OLS) and beta. To this end, "Y" will be used as the response variable in both models because the beta regression model requires a response variable in the form of a proportion with values between 0 and 1. However, in presenting results, the response will be back transformed into growth rate in cm year^{-1} (eq. 3). In both equations (2 and 3), d0 is the diameter growth increment in the most recent year (2017) of the study, and d10 is the diameter growth increment 10 years before that (2007). A descriptive summary of all dependent and independent variables is presented in Table 2.

$$Average\ growth\ rate = \frac{d0 - d10}{10} = \frac{\Delta d}{10} \tag{1}$$

$$Y\ (Dependent\ variable) = \frac{d0 - d10}{d10} = \frac{\Delta d}{d10} \tag{2}$$

$$Back\ transformaiton\ of\ Y\ to\ aveage\ growth\ rate = Y\ x\ \frac{d10}{10} \tag{3}$$

Table 2 Descriptive statistics of major variables

Variables	Minimum	Mean	Maximum
Independent variable			
Target species	-	-	-
Number of neighboring trees	2	6	10
dbh of target trees (cm)	8.89	24.36	44.65
Basal area of all neighboring DF (m^2 ha^{-1})	0	3.59	19.39
Basal area of all neighboring WH (m^2 ha^{-1})	0	3.08	20.39
Basal area of all neighboring RA (m^2 ha^{-1})	0	2.41	13.21
Site Index of DF (base age 50)	107	131	154
Total basal area per plot (m^2 ha^{-1})	24.13	53.79	76.11
Dependent variable			
Average annual growth rate (cm $year^{-1}$)	0.14	0.44	0.93
Average annual growth rate as a proportion of initial dbh	0.05	0.20	0.53

Note: dbh = diameter at breast height

Before selecting the main predictors for modeling, I checked for multicollinearity between the independent variables. Multicollinearity, sometimes referred to as collinearity, occurs if a certain independent variable has a relatively high correlation with one or many other independent variables in the model (Watson & Nimmo-Smith, 2003). The problems with multicollinearity are (1) the regression coefficients are very sensitive to small changes in the dataset, (2) the regression coefficients produce large standard errors, resulting in poor power for the independent variables, and (3) in the most extreme cases, singularity, perfect linear relationship with a correlation coefficient exactly 1.0 or -1.0, can occur, affecting the interpretation of the explanatory variables' effect on the response variable because variance of the coefficients happen to be too wide (Bhandari, 2020; Watson & Nimmo-Smith, 2003).

I calculated the variance inflation factor (VIF) to check the multicollinearity between the independent variables. I used the “car package” in R to compute the VIF values (Fox & Weisberg, 2018). The VIF values can be used to interpret the degree to which the standard error of the independent variable is increased as a result of its correlation with the other independent variables in a regression model (Watson & Nimmo-Smith, 2003). A high VIF value on an independent variable means that there is a high colinear relationship to the other independent variables, which should be considered or adjusted in the model structure and variable selection (Potters & Li, 2023). The formula of VIF is provided in equation 4 where R^2_i is the unadjusted coefficient of determination for regressing the i^{th} independent variable on the remaining ones (Potters & Li, 2023)

$$VIFi = \frac{1}{1 - R_i^2} \tag{4}$$

Table 3 describes the VIF values of all major independent variables which will be included in the model. VIF values were calculated by using equation 4. In theory, the higher the VIF, the higher the possibility that multicollinearity exists, and when it is higher than 10, there is significant multicollinearity that needs to be corrected (Potters & Li, 2023). VIF values of all eight independent variables are lower than the threshold, 10. Therefore, it was assumed that there was no multicollinearity among the independent variables and all of them were used as explanatory variables for the diameter growth rate models.

Table 3 Values of variance inflation factor (VIF) to check multicollinearity

Independent Variable names	VIF	DF
Target species	2.659602	2
Number of neighboring trees	2.568644	1
dbh of target trees (cm)	7.889852	1
Basal area of all neighboring DF (m^2 ha^{-1})	5.697102	1
Basal area of all neighboring WH (m^2 ha^{-1})	7.363678	1
Basal area of all neighboring RA (m^2 ha^{-1})	3.955586	1
Site Index	1.067131	1
Total basal area per plot	1.859846	1

Note: VIF = variance inflation factor; DF = degree of freedom; dbh = diameter at breast height.

Diameter growth models

Individual-tree diameter growth depends on factors such as genetic traits, dbh, height, age, crown size, competition indices, site indices, stand density, stand age, and the number of stems in the plot, etc. (Sterba et al., 2002). Therefore, the model of growth rate was considered to be a function of the target tree species, neighboring tree species, dbh of the neighboring trees, distance between target and nearby trees, and the number of nearby trees, in the base model (eq. 5 - 9). I selected those five base model forms by using different transformations of basal area and site index. In the past, successful tree growth models used transformed versions of those two variables such as natural logarithmic basal area and inverse of site index (Kahriman et al., 2018). Among the base models, I compared five diameter growth models: four ordinary least squares

(OLS) models (eq. 5-8) and one beta regression model (eq. 9). Some of the commonly-used variables applied in diameter growth models are dbh, basal area, height, site index, distance between the target and neighboring trees, crown length and stand age (Kahriman et al., 2018; Weiskittel et al., 2007). In the model structure of this research, I included the explanatory variables of target species identity, number of neighboring trees, initial dbh of target trees ten years ago, total basal area of all Douglas-fir trees per target tree, total basal area of all western hemlock trees per target tree, total basal area of all red alder trees per target tree, site index and total basal area of trees per plot. I then added different distance-independent and distance-dependent competition indices as a new variable, one at a time, to the base model (eq. 10). There is established growth modeling research in which a reduced model is developed as a control to determine the change in diameter growth of individual trees without accounting for competition among trees (Kahriman et al., 2018). This means that eq. 10 can be interpreted as a model that reveals the impact of competition on the diameter growth of the trees.

$$M1: Y = b0 + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 + \mathrm{b6X6} + \mathrm{b7X7} + \mathrm{b8X8} \quad (5)$$

$$M2: Y = b0 + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 + b6X6 + \frac{b7}{X7} + b8X8 \quad (6)$$

$$M3: Y = b0 + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 + \mathrm{b6X6} + \mathrm{b7X7} + \mathrm{b8.}\log(\mathrm{X8}) \quad (7)$$

$$M4: Y = b0 + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 + b6X6 + \frac{b7}{X7} + b8.\log(X8) \quad (8)$$

$$B: Logit(Y) = b0 + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 + b6X6 + \frac{b7}{X7} + b8X8ion \quad (9)$$

$$Y = base\ model + b9X9 \tag{10}$$

Where Y = Growth change relative to 10-year growth (eq. 2 and 3)

X1 = Target tree species (Douglas-fir, western hemlock, red alder)

X2 = Number of neighboring trees

X3 = dbh of target trees

X4 = Basal area of all neighboring DF

X5 = Basal area of all neighboring WH

X 6 = Basal area of all neighboring RA

X 7 = Site Index

X 8 = Total basal area

X 9 = Competition indices (CI1 – CI10)

b0, b1…, b9 = Intercept and slope values of parameter estimates

The goodness of fit of those regression models was determined by using the adjusted coefficient of determination (R^2_{adj}) (eq. 11), root mean square error (RMSE) (eq. 12), Akaike's information criterion (AIC) (eq. 14), Schwarz's Bayesian information criterion (BIC) (eq. 15), and bias (eq. 16). The best base models were selected by using all of these criteria. Competition indices were added to those base models to test the effectiveness of each index.

R^2 is one of the most commonly used criteria to judge the performance of a regression model. It is a measure of the variation of a regression model (Muralidhar, 2023). On the other hand, adjusted R^2 measures the variation of the models with multiple independent variables, which helps determine the goodness of fit of a model (Muralidhar, 2023). However, there is no minimum threshold for R^2 and adjusted R^2 values because it depends on the context. In simple linear regression where there is only one independent variable, the values of adjusted R^2 and R^2 are the same. But when the number of independent variables is more than one, adjusted R^2 is used to test the goodness of fit of the model, which is smaller than the R^2 (Muralidhar, 2023). The difference between them is that adjusted R^2 only adds new predictors to its model if it improves the model's performance, unlike R^2 (Muralidhar, 2023). With R^2, it is assumed that all independent variables considered have impact on the result of the model, while adjusted R^2 accounts for only those independent variables that actually have an impact on the model's performance (ProjectPro, 2022).

Bias is the difference between the observed value of the population and predicted value from a model which is why it is related to the accuracy of an estimator (Grossmann, 2019). On the other hand, RMSE is the square root of the mean square error (MSE) which is the predicted value of the square of the difference between the estimator and the parameter (Grossmann, 2019). They both are useful criteria to decide the performance of a statistical model. The lower the values of these two criteria, the better the model (Grossmann, 2019). RMSE% and bias% are the percentage of root mean square error and bias, respectively.

Akaike information criterion (AIC) (Akaike et al., 1973) is a measure based on in-sample fit to estimate the likelihood of a model (Mohammed et al., 2015). A model with the minimum AIC value among its counterparts can be considered as a good model. Another commonly used criterion is Bayesian information criterion (BIC) (Schwarz, 1978) which measures the trade-off between model fit and complexity of the model. Lower values of both criteria indicate a better fit (Mohammed et al., 2015).

The best base model form(s) were selected, and I added them to CIs to test if any CI improved the model fit for diameter growth predictions and to see which CI would be best.

$$R\ adj^2 = 1 - \frac{(n-1)\sum_{i=1}^{n}(yi - \widehat{yi})^2}{(n-p)\sum_{i=1}^{n}(yi - \overline{yi})^2} \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(yi - \widehat{yi})^2}{n}} \tag{12}$$

$$RMSE\ \% = \frac{RMSE}{mean\ (yi)} \times 100\% \tag{13}$$

$$AIC = -2\ln(L) + 2k \tag{14}$$

$$BIC = -2\ln(L) + k * \ln(n) \tag{15}$$

$$Bias = \frac{Yi - \widehat{Yi}}{n} \tag{16}$$

$$Bias\ \% = \frac{bias}{mean\ (Yi)} \times 100\ \% \tag{17}$$

$$MSER = \left(1 - \frac{MSE_{Full\ model}}{MSE_{Base\ model}}\right) \times 100 \tag{18}$$

Where R_{adj}^2 = Adjusted coefficient of determination

RMSE = Root mean squared error

AIC = Akaike information criterion

BIC = Bayesian information criterion

MSER = Mean squared error reduction

Yi = Observed value

$\widehat{Yi}$ = Predicted value from the model n = Sample size

p = Number of independent variables

$L =$ Maximum value of the likelihood function for the model

k = number of estimated parameters in the model

MSE = Mean squared error

Competition indices and their evaluation

There are several competition indices (CI) used in modeling the growth increment of trees (Kahriman et al., 2018). Among the 18 most commonly used CIs (Kahriman et al., 2018), a total of ten CIs were selected based on the data available. Since I have some variables such as dbh of both target and neighboring trees, the distance between target and neighboring trees, the number of trees, and the plot size, they can be easily transformed into basal area, distance summation, etc. that are required for different CI calculations. I have selected seven distance-independent and three distance-dependent competition indices for this research. CI1 is the sum of the basal areas of the trees which have diameters that are greater than the target tree's diameter (Wykoff et al., 1982). CI2 is the ratio of the sum of the dbh of nearby trees to the dbh of the target tree (Lorimer, 1983b). CI3 is the ratio of the dbh of the target tree to the quadratic mean diameter of the trees in the plot (Hamilton, 1986). CI4 is the ratio of the dbh of the subject tree to the largest tree's dbh inside the plot (Tomé & Burkhart, 1989). CI5 is the ratio of the basal area of the target tree to the average basal area of the trees in the plot (Tomé & Burkhart, 1989). CI6 is the ratio of the basal area of the target tree to the basal area of the largest tree with the largest diameter in the plot (Tomé & Burkhart, 1989). CI7 is the ratio of the summation of the basal area of the nearby trees in the plot to the basal area of the target tree (Corona & Ferrara, 1989). Mathematical equations of these competition indices are described in Table 4.

Smaller values of CI1 and CI7 indicates that the target tree has a competitive advantage i.e. it has approached free growth (Kahriman et al., 2018). Increasing the values of other distance-independent competition indices (CI2 – CI6) means that the target tree has approached free growth (Kahriman et al., 2018). In this case, the target tree can reach the highest potential yield (Kahriman et al., 2018). CI8, CI9 and CI10 are size-ratio indices which were weighted by the distance of the subject tree to its neighboring trees. These CIs assume that the competitive influence of a neighboring tree increases when the trees' dimensions are greater and the distance is shorter (Kahriman et al., 2018).

Table 4 Mathematical formula of ten competition indices for tree diameter growth models

Source	Competition Indices
Distance-independent CI's	
(Wykoff et al., 1982)	$CI1 = BALi = \left(\sum_{di<dj}^{n} gi \right)$
(Lorimer, 1983b)	$CI2 = \left[\left(\sum_{j \neq i}^{n} dj \right) / di \right]$
(Hamilton, 1986)	$CI3 = \frac{di}{dg}$
(Tomé & Burkhart, 1989)	$CI4 = \frac{di}{dmax}$
(Tomé & Burkhart, 1989)	$CI5 = \frac{gi}{\bar{g}}$
(Tomé & Burkhart, 1989)	$CI6 = \frac{gi}{gmax}$
(Corona & Ferrara, 1989)	$CI7 = \left[\left(\sum_{j \neq i}^{n} gj \right) / gi \right]$

Table 4 (continued)

Distance-dependent CI's	
(Hegyi, 1974)	$$CI8 = BALi = \sum_{j=1}^{n}\left(\frac{dj}{di}\times\frac{1}{Lij}\right)$$
(Alemdag, 1978)	$$CI9 = \sum_{j=1}^{n}\left\{\pi\left(\frac{Lij \times di}{di+dj}\right)\left[\frac{dj/_{Lij}}{\sum\left(dj/_{Lij}\right)}\right]^2\right\}$$
(Martin & Ek, 1984)	$$CI10 = \sum_{J=1}^{n}\left(\frac{dj}{di}\right) \times e^{(16\times Lij)(di+dj)}$$

Note: CI1, CI2,…, CI10, competition indices; i, target tree; j, neighboring tree; di, dbh of target tree (cm), dj, dbh of nearby tree (cm); gi, basal area of target tree (m^2); BALi, basal area of trees larger than the target tree (m^2); dg, quadratic mean diameter (cm); dmax, maximum dbh of the tree with the largest diameter in the sample plot; $\bar{g}$, mean basal area of sample plot (m^2ha^{-1}); gmax, basal area of the thickest diameter in the sample plot (m^2); gj, basal area of nearby trees, Lij, distance between target and the neighboring trees (m)

CHAPTER III

RESULTS

Base model selection

Values of R^2_{Adj}, RMSE, RMSE%, AIC, BIC, Bias, and Bias% were compiled for each base model (Table 5). R^2_{Adj} values are better when they are higher, however, it is not always the single best factor to critique the strength of the model. All the models have relatively similar R^2_{Adj} values (Table 5). Smaller values of AIC and BIC indicate that the model is better. According to all performance criteria, all ordinary least squares models (M1, M2, M3, M4) have similar values. Only M2 has a slightly better R^2_{Adj} value. Therefore, transformation of independent variables did not have a significant impact on the performance of the model. But the beta regression model (B) stood out as the model explaining the most variation. Therefore, model M2 (eq. 5) and model B (eq. 9) were chosen as the best base models. Those two models were used as the base models to test the contribution of the ten competition indices.

Table 5 Selection criteria to test the goodness of fit of the regression models

M	R^2_{Adj}	RMSE	RMSE (%)	AIC	BIC	Bias	Bias%
M1	0.6425	0.0557	28.3049	-189.4985	-164.4551	$1.35x10^{-18}$	$6.86 x10^{-16}$
M2	0.6431	0.0557	28.3049	-189.6165	-164.5732	$1.16 x10^{-18}$	$5.89 x10^{-16}$
M3	0.6409	0.0558	28.3557	-189.1654	-164.1221	$5.8 x10^{-19}$	$2.95 x10^{-16}$
M4	0.6415	0.0558	28.3557	-189.2935	-164.2501	$-3.9 x10^{-19}$	$-1.98 x10^{-16}$
B	0.7398	0.0490	24.9002	-220.5776	-195.5343	-0.0005	-0.2541

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

Full models with ten competition indices

Based on the assessed model statistics, the diameter growth models including CIs generally performed better than their respective base models. As one of the important performance criteria, RMSE values were found to be lower in all full models, compared to their specific base models (Tables 6 and 7). Among the ordinary least square models, RMSE of the base model was around 28% whereas the errors were reduced up to 12% in the full model with CI5 and other competition indices also contributed to decrease the RMSE% (Table 6). Similarly, the beta regression model shows lower RMSE values when competition indices were added. The original RMSE of the beta model was around 25%, but the full models with the competition indices had better RMSE values (up to 10%) (Table 7). Looking at the AIC and BIC values, lower values of those criteria are found when competition indices are added to both the least-squares and beta models (Tables 6 and 7). Models with lower values of AIC and BIC indicate that they perform better. In ordinary least squares models, AIC values decreased from -190in the base model to -3010in the full model, and BIC values were reduced from -165 in the base model to -282in the full model, as the lowest values. The same trend of lower AIC and BIC values were also found in the full beta regression models, compared to their base model. In addition, when a significance test was performed, most of the full models have a significant contribution towards the diameter growth models at significant levels 0.01 and 0.05, in the ordinary least squares and beta regression models respectively.

Table 6 Contribution of competition indices to the diameter growth model (Ordinary Least Squares Model)

CI	R^2_{Adj}	RMSE	RMSE (%)	AIC	BIC	Bias	Bias%	MSER	Partial F test
No CI	0.6431	0.0557	28.3049	-189.6165	-164.5732	1.16×10^{-18}	5.89×10^{-16}		
CI1	0.6798	0.0523	26.5771	-196.584	-169.264	-1.9×10^{-19}	-9.66×10^{-17}	12.90323	8.0909**
CI2	0.7865	0.0427	21.6987	-225.7799	-198.4599	1×10^{-19}	5.08×10^{-17}	41.93548	42.64***
CI3	0.6804	0.0522	26.5263	-196.728	-169.408	-7.7×10^{-19}	-3.91×10^{-16}	12.90323	8.2292**
CI4	0.8848	0.0314	15.9564	-270.1849	-242.8649	1.35×10^{-18}	6.86×10^{-16}	67.74194	131.03***
CI5	0.9334	0.0238	12.0944	-309.7057	-282.3857	9.6×10^{-19}	4.88×10^{-16}	80.64516	271.47***
CI6	0.8842	0.0315	16.0073	-269.7968	-242.4768	3.08×10^{-18}	1.57×10^{-15}	67.74194	130***
CI7	0.9123	0.0274	13.9238	-289.8492	-262.5292	2.55×10^{-18}	1.3×10^{-15}	77.41935	191.34***
CI8	0.6563	0.0542	27.5426	-191.494	-164.174	1.93×10^{-18}	9.81×10^{-16}	6.451613	3.3752 .
CI9	0.6411	0.0554	28.1524	-188.3832	-161.0632	1.9×10^{-19}	9.66×10^{-17}	0	0.653
CI10	0.6403	0.0554	28.1524	-188.2184	-160.8984	7.7×10^{-19}	3.91×10^{-16}	0	0.5121

Note: ., significant at 0.10 level; *, significant at 0.05 level; **, significant at 0.01 level; ***, significant at 0.001 level

Table 7 Contribution of competition indices to diameter growth model (Beta Regression Model)

CI	R^2_{Adj}	RMSE	RMSE (%)	AIC	BIC	Bias	Bias%	MSER	Chi Sq test
No CI	0.7398	0.0490	24.9002	-220.5776	-195.5343	-0.0005	-0.2545		
CI1	0.7871	0.0434	22.0544	-235.6791	-208.3591	-0.0004	-0.2334	20.83333	17.102***
CI2	0.8429	0.0342	17.3793	-262.5932	-235.2732	0.0001	0.0362	50	44.016***
CI3	0.7792	0.0451	22.9183	-230.2025	-202.8825	-0.0005	-0.2287	16.66667	11.625***
CI4	0.8569	0.0328	16.6679	-267.609	-240.289	0.0001	0.0271	54.16667	49.031***
CI5	0. 9537	0.0189	9.6044	-347.7711	-320.4511	0.00001	0.0088	83.33333	129.19***
CI6	0.8603	0.0324	16.4646	-269.6005	-242.2805	0.0001	0.0284	58.33333	51.023***
CI7	0.8957	0.0264	13.4156	-294.8871	-267.5671	0.0001	0.0531	70.83333	76.31***
CI8	0.7504	0.0469	23.8330	-223.1191	-195.7991	-0.0004	-0.2100	8.333333	4.5415*
CI9	0.7411	0.0484	24.5953	-219.9557	-192.6357	-0.0004	-0.2521	4.166667	1.3781
CI10	0.7408	0.048	24.392	-220.6313	-193.3113	-0.0005	-0.2573	4.166667	2.0537

Note: *, significant at 0.05 level; **, significant at 0.01 level; ***, significant at 0.001 level

CI5, which is the ratio of target tree's basal area to the average basal area, is considered the best competition index for the diameter growth models of ordinary least squares and beta regression. The ordinary least squares model with CI5 and the beta regression model with CI5 had the highest R^2_{Adj} values, 0.9334 and 0.9537 respectively (Tables 6 and 7). This CI improved the models a substantial amount since the original R^2_{Adj} values were only 0.6431 and 0.7398 in ordinary least squares and beta regression models respectively. However, the adjusted R^2 value alone cannot decide the strength of a model. Looking at the RMSE% values, the same model gave the lowest values; 12.0944% and 9.6044% respectively, whereas the base models had

RMSE% of 28.3049% and 24.9002% in ordinary least squares and beta regression models respectively. Another criterion applied in selecting the best competition index is the mean square error reduction (MSER). CI5 had the highest MSER values for both models and gives the best performance amongst all. They both are also significant at the 0.001 level.

In both tables (Tables 6 and 7), CI 1-7 are distance-independent competition indices and CI 8-10 are distance-dependent indices. Among the full forms of both the ordinary least squares and beta models, higher values of R^2_{Adj} were found when the distance-independent competition indices were added. Regarding the RMSE values, models with distance-independent competition indices show better results i.e., lower RMSE values. Similarly, AIC and BIC values were lower in the first group. Among the three distance-dependent competition indices (CI8-10), only CI8 was significant at the 95% confidence level in the beta regression model and 90% confidence level in the ordinary least squares, whereas the other two were not significant. Therefore, distance-independent competition indices contribute better than distance-dependent indices towards the tree diameter growth models of both ordinary least squares and beta regression.

After finding out which competition index performs better for my chosen models, comparison between the full ordinary least squares model and the full beta regression model was also made. According to my model selection criteria, the pseudo R^2 value of the beta model was slightly better than the R^2_{Adj} value of the ordinary least squares, 0.9537 and 0.9334. This means that the beta model can explain 95.37% of the variation in the data, while the ordinary least squares model can only explain 93.34%. In the other criteria such as RMSE% and bias%, the beta model had lower values compared to its counterpart. Therefore, application of beta regression can slightly improve the performance of tree diameter growth models, compared to ordinary least squares regression.

Table 8 Regression models developed to predict diameter growth

Parameters	Ordinary Least Squares Full Model		Beta Regression Full Model	
	Estimates	SE	Estimates	SE
b0 (targ_DF)	0.9273***	4.709×10^{-2}	3.005***	2.379×10^{-1}
b1 (targ_RA)	6.773×10^{-3}	9.484×10^{-3}	9.773×10^{-2}*	4.942×10^{-2}
b2 (targ_WH)	1.909×10^{-2}.	1.036×10^{-2}	0.2535***	5.383×10^{-2}
b3 (n_nei)	-3.683×10^{-3}	3.525×10^{-3}	-3.385×10^{-2}*	1.607×10^{-2}
b4 (targ_d11)	-4.383×10^{-3}*	1.671×10^{-3}	-4.290×10^{-2}***	8.196×10^{-3}
b5 (df_ba_sum)	-4.339×10^{-4}	2.216×10^{-7}	2.421×10^{-6}*	1.078×10^{-6}
b6 (wh_ba_sum)	5.689×10^{-3}*	2.191×10^{-7}	4.972×10^{-6}***	1.044×10^{-6}
b7 (ra_ba_sum)	3.026×10^{-3}	2.763×10^{-7}	4.436×10^{-6}***	1.329×10^{-6}
b8 (inv_site_index)	3.464	3.943	-9.689	2.063×10^{1}
b9(ba_sum)	6.863×10^{-5}	3.538×10^{-8}	-9.537×10^{-8}	1.755×10^{-7}
b10 (CI 5)	-0.6650***	4.036×10^{-2}	-3.616***	1.906×10^{-1}
R^2	0.9334		0.9537	
RMSE%	12.0944		9.6044	
AIC	-309.7057		-347.7711	
BIC	-282.3857		-320.4511	

Note: targ_DF = Target tree's species identity, Douglas-fir; targ_RA = Target tree's species identity, red; targ_WH = Target tree's species identity, western hemlock; n_nei = Number of neighboring trees; targ_d11 = Target tree's initial diameter; df_ba_sum = Total basal area of all Douglas-fir per plot; wh_ba_sum = Total basal area of all western hemlock per plot; ra_ba_sum = Total basal area of all red alder per plot; inv_site_index = Inverse of site index; ba_sum = Total basal area of all trees per plot.

Table 8 describes the model coefficients of the full ordinary least squares model and the full beta regression model with CI5.

Figure 1 shows the histograms of residuals from the two diameter growth models and the scatterplots between the residuals and the predicted values, to check for homogeneity of variance. It can be seen that the histograms are parabola shaped, and the residuals are scattered randomly, which indicates that the variance of the residuals is constant across fitted values.

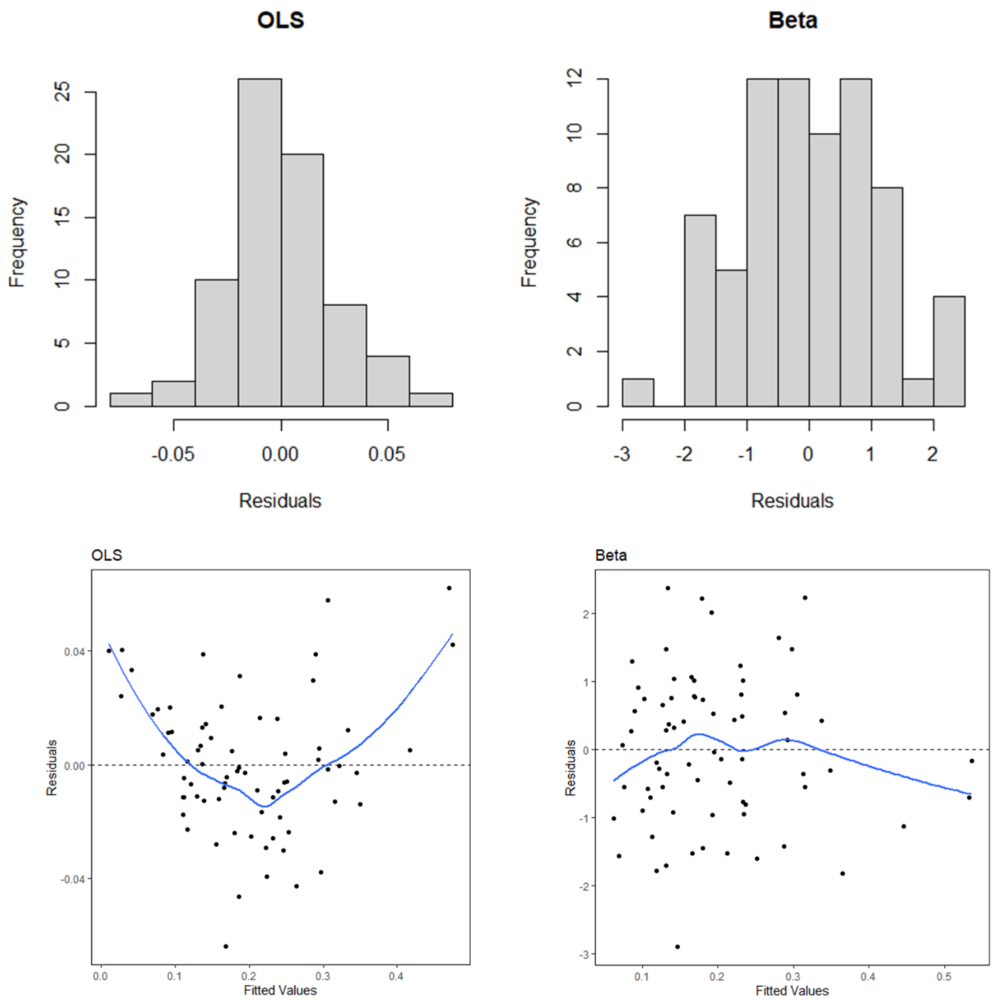

Figure 1 Histograms of residuals and the scatterplots between residuals and fitted values, of both ordinary least squares (OLS) and beta regression models

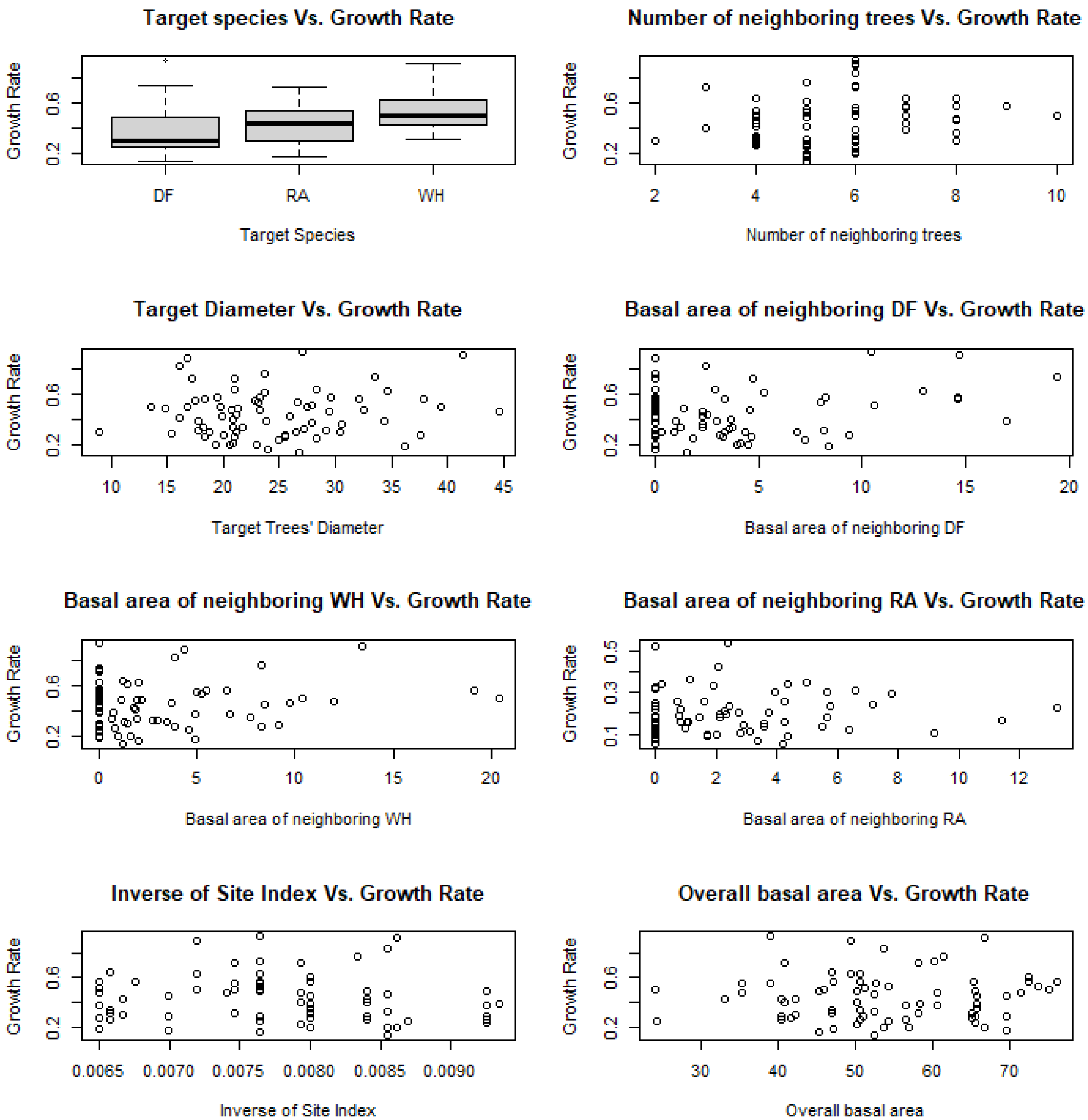

Figure 2 The relationship between tree diameter growth rate and independent variables

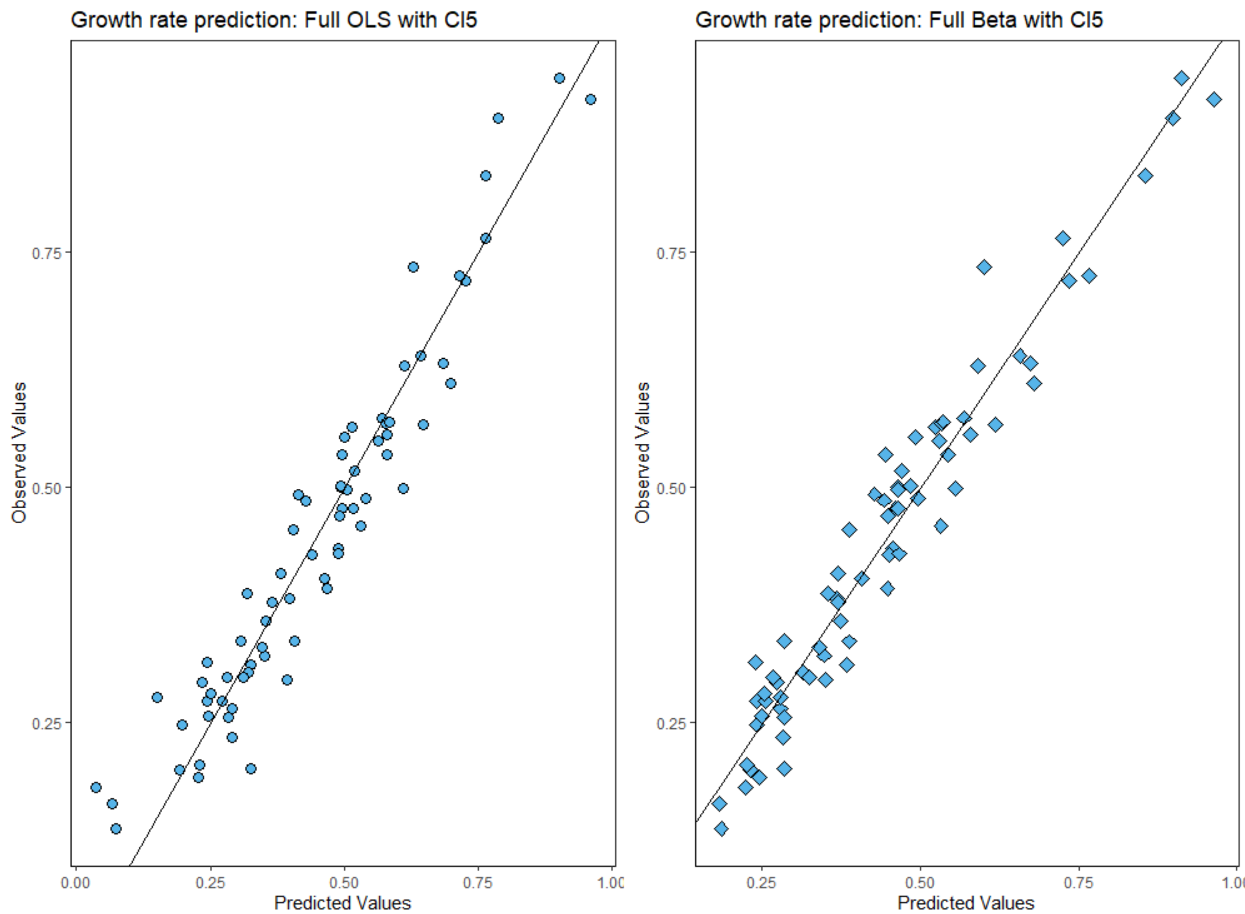

Figure 3 Predicted diameter growth rate vs. observed diameter growth rate for the full ordinary least squares model (left) and the full beta regression model (right)

The relationship between the tree diameter growth rate and the main predictors such as target species, the number of neighboring trees, initial dbh of target trees, basal area of all neighboring Douglas-fir, basal area of all neighboring western hemlock, basal area of all neighboring red alder, site index, and total basal area per plot are illustrated in Figure 2. The predicted diameter growth rate according to the best two models containing CI5 versus observed diameter growth rate is shown in Figure 3. The diagonal lines represent the best fit line. Figures 4 and 5 depict the relationship between growth rate and ten competition indices. According to the trends, some of the competition indices show a linear relationship with growth rate. Specifically, the CI5 competition index has a negative correlation with diameter growth rate. CI5 seems to be strongly linear for the full range of CI values, compared to other competition indices.

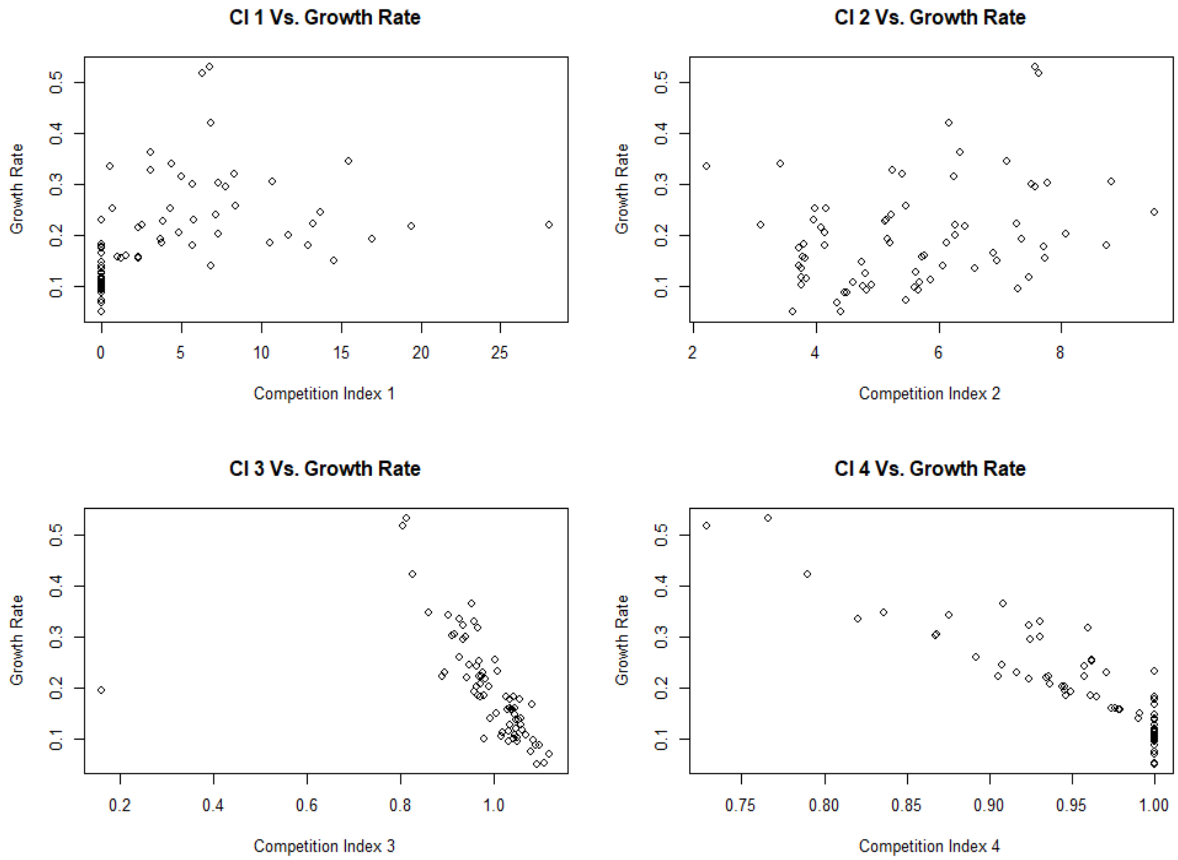

Figure 4 Dependent variable versus competition indices (CI1-CI4)

Note: CI = competition index.

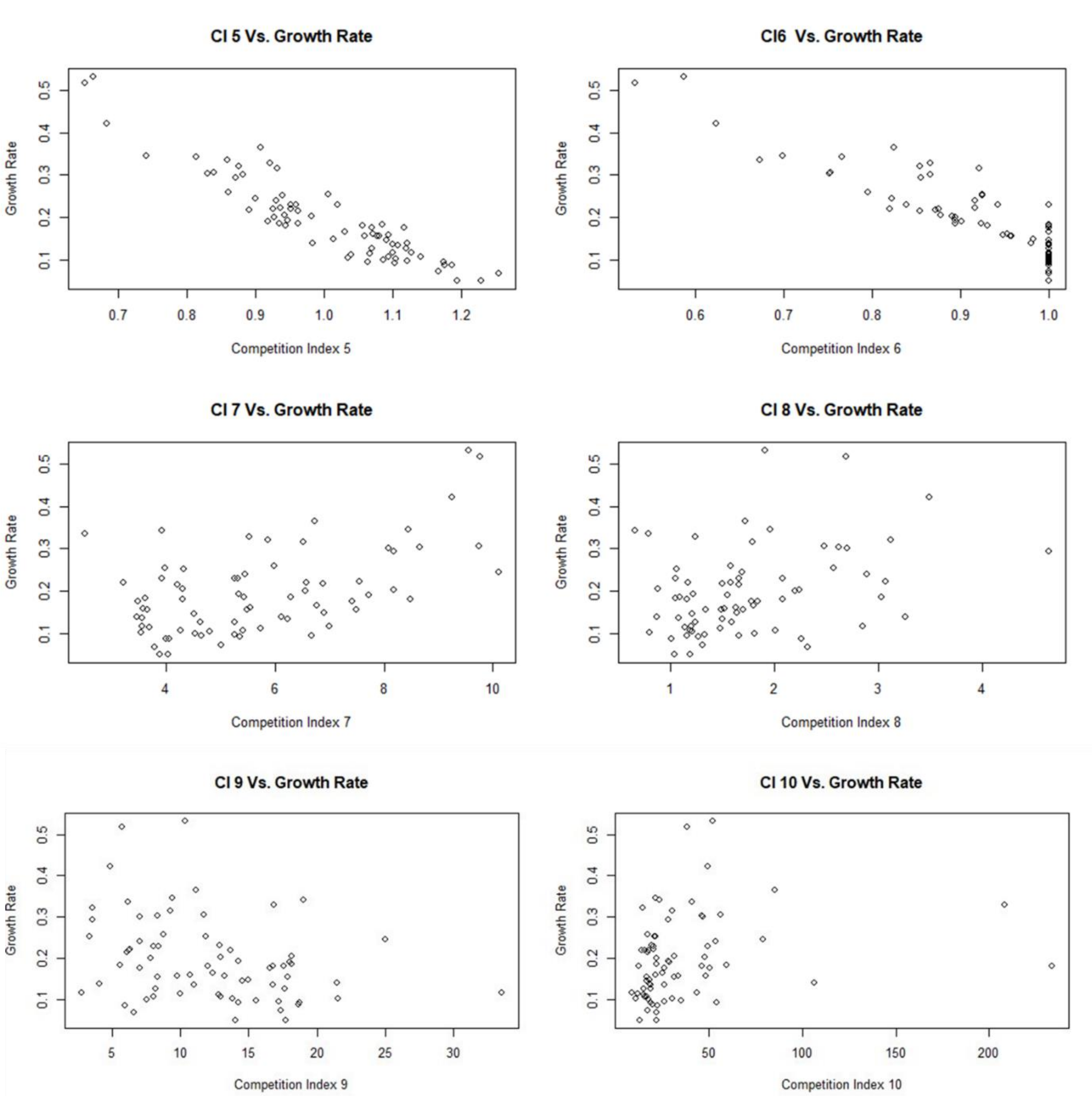

Figure 5 Dependent variable versus competition indices (CI5-CI10)

Note: CI = competition index.

Model Validation

Among the different types of model validation techniques, cross-validation is a resampling procedure which is applied to evaluate statistical models on a limited data sample size. In the k-fold cross-validation method, the data sample is split into k groups (Brownlee, 2018). The value for k is usually chosen as either 5 or 10, but there is no fixed rule for it (James et al., 2013). The larger the k value is, the difference in size between the training set and the resampling subsets gets smaller, and the bias of the technique also becomes smaller (Kuhn & Johnson, 2013). In this research, the value for k is fixed to 10 to generally result in a model estimate with low bias and modest variance. The procedure is very straightforward: splitting the dataset into 10 groups, taking out each group as a test data set and taking the remaining groups as a training data set, fitting a model with the training set and evaluating it on the test data set, and deciding to retain or discard the model based on the performance criteria (Brownlee, 2018). I did not scale the data in any way before fitting the models. Table 9 describes three metrics provided as the output of the validation test: root mean square error (RMSE), R^2 and mean absolute error (MAE). MAE is the average absolute difference between the predictions made by the model and the actual observations (Zach, 2020). The smaller the MAE, the more accurately a model can predict the actual observations (Zach, 2020). These three metrics provided information of the models' performance measure. RMSE and MAE values of the final models were relatively low and even close to zero in both models. In terms of percentage, the RMSE was approximately 13% and the MAE was approximately 10%. R^2 values of both models were considerably high. Hence, the performance of both models fits with the dataset.

Table 9 Results of the two full models after being validated with the K-fold method

Model	RMSE	RMSE%	R^2	MAE	MAE%
OLS with CI5	0.0274	13.7	0.9388	0.0219	10.95
Beta with CI5	0.0263	13.15	0.9392	0.0210	10.5

Note: RMSE = root mean squared error, R^2 = coefficient of determination, MAE = mean absolute error.

CHAPTER IV

DISCUSSION AND CONCLUSION

Among the ten competition indices applied in this research, all seven distance-independent indices contributed significantly towards the diameter growth rate in both ordinary least squares and in beta regression models, whereas only one distance-dependent index, CI8, the sum of distance-weight ratio of the neighboring tree, improved the growth rate model. However, CI8 only made a significant contribution to the beta model, but not the ordinary least squares model. The cause of relatively poor performance of the distance-dependent competition indices is that I had already selected neighboring trees which had the crowns touching with the target tree's crown in the base models, and that direct interaction was based on stem proximity. If I had not already made that selection of neighbors, the role of the distance between stems might be more important. This finding from the research also matches the results of previous studies that support the idea that distance-dependent competition indices tend to be less important than distance-independent indices in diameter growth models (Radtke et al., 2003; Wimberly & Bare, 1996). Therefore, it can be interpreted that the use of proximity of canopies is an effective way to identify the competing neighborhood of trees, and it is also relatively easy to carry out in the field. It is also something that can likely be determined from the spatial resolution of remote sensing, e.g. multi-spectral images (Richter et al., 2021; Tu et al., 2019)

The CI5 competition index was the best based on all model assessment criteria, but CI7, CI6, CI4 and CI2 also substantially improved the amount of variation in target tree diameter growth that could be explained. In the full ordinary least squares models with these four competition indices, R^2_{adj} values were between 0.7865 and 0.9123 for the ordinary least squares model and between 0.8429 and 0.8957 for the beta regression model. On top of that, other performance criteria such as RMSE, RMSE%, bias, bias%, AIC and BIC also improved. Therefore, although this research highlighted CI5 as the best competition index for diameter growth models, the other CIs could also be used in future modeling because the effectiveness of different competition indices depends on several stand factors (Kahriman et al., 2018).

The best resulting competition index, CI5, is the ratio of target tree's basal area to the average basal area of the plot. Therefore, it can be easily derived using easy to acquire variables such as the dbh of the trees in each plot. But it needs information on all trees to get the average basal area of the plot. The second and third best resulting competition indices, CI7 and CI6, are also related to the ratio of different basal areas: the ratio of the sum of neighboring trees' basal area to the target tree's basal area, and the ratio of target tree's basal area to the largest tree's basal area, of each plot. Therefore, competition indices using basal area information had a significant effect in improving the diameter growth rate models in both ordinary least squares and beta regression models. Another interesting point to note is that all five best competition indices are distance-independent competition indices. Therefore, distance information does not make a significant contribution to the diameter growth models. This might also be because the crowns of the neighboring trees selected for this study overlap with the target trees already.

There are some common characteristics among the best five CIs. Whereas CI5, CI7, and CI6, are the values of the proportion between different basal areas, the fourth and fifth best CIs, CI4 and CI2 are related to the ratio of different diameters: CI4 being the ratio of target tree's dbh to the maximum dbh per plot, and CI2 being the ratio of the sum of all neighboring trees' dbh to the dbh of the target tree. Even though basal area is derived from dbh, they have different impacts in diameter growth models. Hence, basal area and diameter information are useful parameters in diameter growth models. However, the relative performance of these CIs indicates that it is best to incorporate the basal area information rather than the simple dbh in growth and yield models.

In addition to R^2_{Adj}, RMSE, bias, AIC and BIC which are the commonly found criteria in statistical modelling, another interesting criterion applied in this research is MSER. It is used to assess if the performance of the model increased, or the errors decreased, when a particular competition index was added to the reduced model (Kahriman et al., 2018). The significance of the MSER statistic is tested by the partial F test for the ordinary least squares model and by the chi-square test for the beta regression model. Like other performance criteria, the best MSER value is produced by the CI5 competition index, 81 and 83 for the ordinary least squares and beta regression models, respectively, which are followed by the same four CI competition indices, CI7, CI6, CI4 and CI2. Their values range from 42 and 77 for the ordinary least squares model, and from 50 to 71 for the beta regression model.

Looking at the growth rate of each species according to the best resulting models, both ordinary least squares and beta regression, growth rate changes across different species. The growth rate of western hemlock was the highest, which was followed by red alder and Douglas-fir in descending order. Another point of interest is that the models considered the effect of basal area information of each species, and all of them were also significant at 95% significance level. For practical purpose, deriving models should be simple but not too simple (Otto & Day, 2007). Therefore, it was found that incorporation of a simple competition index was very effective in a growth model that accounts for the basal area of different species. Different species involved in this research have different interactions with each other, resulting in differing growth rate depending on the species.

For future research directions, improvements can be made by adding more variables such as the distance of crown projection overlap between the target and the neighboring tree, crown radius of the target tree, crown overlap between the target and the neighboring tree, and other distance-related information, to test more distance-dependent competition indices (Arney, 1973; Bella, 1971; Staebler, 1951). This method can also be applied to test different species and in different localities, although my study focuses on Douglas-fir, red alder, and western hemlock.

One of the primary purposes of building diameter growth models by using competition indices is to determine the compromise between conspecific and heterospecific tree individuals with a few easy measurements for feasibility (Kahriman et al., 2018). Instead of attempting to collect several variables that are costly and time-consuming, using simple, non-spatially referenced information, and easy-to-calculate competition indices are very efficient. However, as per a famous quote of George E.P. Box, a British statistician, “Every model is wrong, but some are useful.” Therefore, future growth and yield studies which apply either the methodology of this research or the results of it, should modify the structure of the model and the choice of potential CIs based on the stand structure and species composition.

REFERENCES

Akaike, H., Petrov, B., & Csaki, F. (1973). Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, 276–281.

Alemdag, I. S. (1978). Evaluation of some competition indexes for the prediction of diameter increment in planted white spruce. *Information Report Forest Management Institute (Canada). No. FMR-X-108.*

Arney, J. D. (1973). *Tables for quantifying competitive stress on individual trees*. *078*. http://cfs.nrcan.gc.ca/publications?id=26500

Begon, M., Harper, J. L., & Townsend, C. R. (1996). *Ecology: Individuals, Populations and Communities* (3rd edition). Wiley-Blackwell.

Bella, I. E. (1971). A new competition model for individual trees. *Forest Science*, *17*(3), 364–371.

Bhandari, A. (2020, March 19). Multicollinearity | Causes, Effects and Detection Using VIF (Updated 2023). *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2020/03/what-is-multicollinearity/

Biging, G. S., & Dobbertin, M. (1995). Evaluation of competition indices in individual tree growth models. *Forest Science (USA)*.

Binkley, D. (2003). Seven decades of stand development in mixed and pure stands of conifers and nitrogen-fixing red alder. *Canadian Journal of Forest Research*, *33*(11), 2274–2279. https://doi.org/10.1139/x03-158

Briggs, D. G. comp, DeBell, D. S. comp, & Atkinson, W. A. comp. (1978). Utilization and management of Alder. Proceedings of a symposium held at Ocean Shores, Washington, April 25-27, 1977. *USDA For Serv Gen Tech Rep PNW Pac Northwest For Rangeexp Stn*.

Brownlee, J. (2018, May 22). A Gentle Introduction to k-fold Cross-Validation. *MachineLearningMastery.Com*. https://machinelearningmastery.com/k-fold-cross-validation/

Brunner, A., & Forrester, D. I. (2020). Tree species mixture effects on stem growth vary with stand density – An analysis based on individual tree responses. *Forest Ecology and Management*, *473*, 118334. https://doi.org/10.1016/j.foreco.2020.118334

Cole, E., & Newton, M. (2023). Density effects on growth and dominance in pure and mixed stands of red alder and western hemlock in western Oregon. *Forest Ecology and Management*, *537*, 120950. https://doi.org/10.1016/j.foreco.2023.120950

Corona, P., & Ferrara, A. (1989). Individual competition indices for conifer plantations. *Agriculture, Ecosystems & Environment*, *27*(1), 429–437. https://doi.org/10.1016/0167-8809(89)90103-5

Cortini, F., & Comeau, P. G. (2008a). Effects of red alder and paper birch competition on juvenile growth of three conifer species in southwestern British Columbia. *Forest Ecology and Management*, *256*(10), 1795–1803. https://doi.org/10.1016/j.foreco.2008.03.022

Cortini, F., & Comeau, P. G. (2008b). Effects of red alder and paper birch competition on juvenile growth of three conifer species in southwestern British Columbia. *Forest Ecology and Management*, *256*(10), 1795–1803. https://doi.org/10.1016/j.foreco.2008.03.022

Crookston, N. L. (1982). User's Guide to the Stand Prognosis Model. *Gen. Tech. Rep. INT-133. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station; 1982. 112 p.*, *133*. https://doi.org/10.2737/INT-GTR-133

Daniels, R. F. (1976). Simple Competition Indices and Their Correlation with Annual Loblolly Pine Tree Growth. *Forest Science*, *22*(4), 454–456. https://doi.org/10.1093/forestscience/22.4.454

Fang, C., Comeau, P. G., & Harper, G. J. (2019). Effects of red alder on growth of Douglas-fir and western redcedar in southwestern British Columbia. *Forest Ecology and Management*, *434*, 244–254. https://doi.org/10.1016/j.foreco.2018.12.023

Forrester, D. (2014). The spatial and temporal dynamics of species interactions in mixed-species forests: From pattern to process. *Forest Ecology and Management*, *312*, 282–292. https://doi.org/10.1016/j.foreco.2013.10.003

Forrester, D., & Bauhus, J. (2016). A Review of Processes Behind Diversity—Productivity Relationships in Forests. *Current Forestry Reports*, *2*(1), 45–61. https://doi.org/10.1007/s40725-016-0031-2

Fox, J., & Weisberg, S. (2018). *An R Companion to Applied Regression* (3rd edition). SAGE Publications, Inc.

Grossmann, W. (2019, May 9). *Root mean square error (RMSE)* [Text]. CROS - European Commission. https://cros-legacy.ec.europa.eu/content/root-mean-square-error-rmse_en

Hamilton, D. A. (1986). A Logistic Model of Mortality in Thinned and Unthinned Mixed Conifer Stands of Northern Idaho. *Forest Science*, *32*(4), 989–1000. https://doi.org/10.1093/forestscience/32.4.989

Hansen, E. M., Stone, J. K., Capitano, B. R., Rosso, P., Sutton, W., Winton, L., Kanaskie, A., & McWilliams, M. G. (2000). Incidence and Impact of Swiss Needle Cast in Forest Plantations of Douglas-fir in Coastal Oregon. *Plant Disease*, *84*(7), 773–778. https://doi.org/10.1094/PDIS.2000.84.7.773

Hegyi, F. (1974). A simulation model for managing jack-pine stands. *RoyalColl. For, Res. Notes*, *30*, 74–90.

Himes, A., & Puettmann, K. (2020). Tree species diversity and composition relationship to biomass, understory community, and crown architecture in intensively managed plantations of the coastal Pacific Northwest, USA. *Canadian Journal of Forest Research*, *50*(1), 1–12. https://doi.org/10.1139/cjfr-2019-0236

Himes, A., Puettmann, K., & Muraca, B. (2020). Trade-offs between ecosystem services along gradients of tree species diversity and values. *Ecosystem Services*, *44*, 101133. https://doi.org/10.1016/j.ecoser.2020.101133

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R* (1st ed. 2013, Corr. 7th printing 2017 edition). Springer.

Kahriman, A., Şahin, A., Sönmez, T., & Yavuz, M. (2018). A novel approach to selecting a competition index: The effect of competition on individual-tree diameter growth of Calabrian pine. *Canadian Journal of Forest Research*, *48*(10), 1217–1226. https://doi.org/10.1139/cjfr-2018-0092

Kocher, S. D., & Harris, R. (2007). *Forest Stewardship Series 5: Tree Growth and Competition*. University of California, Agriculture and Natural Resources. https://doi.org/10.3733/ucanr.8235

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling* (1st ed. 2013, Corr. 2nd printing 2018 edition). Springer.

Lorimer, C. G. (1983a). Tests of age-independent competition indices for individual trees in natural hardwood stands. *Forest Ecology and Management*, *6*(4), 343–360. https://doi.org/10.1016/0378-1127(83)90042-7

Lorimer, C. G. (1983b). Tests of age-independent competition indices for individual trees in natural hardwood stands. *Forest Ecology and Management*, *6*(4), 343–360. https://doi.org/10.1016/0378-1127(83)90042-7

Maleki, K., Kiviste, A., & Korjus, H. (2015). Analysis of individual tree competition on diameter growth of silver birch in Estonia. *Forest Systems*, *24*(2), e023.

Martin, G. L., & Ek, A. R. (1984). A comparison of competition measures and growth models for predicting plantation red pine diameter and height growth. *Forest Science*, *30*(3), 731–743. https://doi.org/10.1093/forestscience/30.3.731

Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Chapter 32—Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In Q. N. Tran & H. Arabnia (Eds.), *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology* (pp. 577–602). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-802508-6.00032-6

Muralidhar, K. (2023, February 7). *Demystifying R-Squared and Adjusted R-Squared | Built In* [Data Science]. Builtin. https://builtin.com/data-science/adjusted-r-squared

P. Otto, S., & Day, T. (2007). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. https://press.princeton.edu/books/hardcover/9780691123448/a-biologists-guide-to-mathematical-modeling-in-ecology-and-evolution

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

Potters, C., & Li, T. (2023, February 12). *Variance Inflation Factor (VIF)*. Investopedia. https://www.investopedia.com/terms/v/variance-inflation-factor.asp

Pretzsch, H. (2009). Forest Dynamics, Growth, and Yield. In H. Pretzsch (Ed.), *Forest Dynamics, Growth and Yield: From Measurement to Model* (pp. 1–39). Springer. https://doi.org/10.1007/978-3-540-88307-4_1

ProjectPro. (2022, December). *What is the difference between R squared and adjusted R squared?* ProjectPro. https://www.projectpro.io/recipes/what-is-difference-between-r-squared-and-adjusted-r-squared

Radosevich, S. R., Hibbs, D. E., & Ghersa, C. M. (2006). Effects of species mixtures on growth and stand development of Douglas-fir and red alder. *Canadian Journal of Forest Research*, *36*(3), 768–782. https://doi.org/10.1139/x05-280

Radtke, P. J., Westfall, J. A., & Burkhart, H. E. (2003). Conditioning a distance-dependent competition index to indicate the onset of inter-tree competition. *Forest Ecology and Management*, *175*(1), 17–30. https://doi.org/10.1016/S0378-1127(02)00118-4

Richter, R., Hutengs, C., Wirth, C., Bannehr, L., & Vohland, M. (2021). Detecting tree species effects on forest canopy temperatures with thermal remote sensing: the role of spatial resolution. *Remote Sensing*, *13*(1), Article 1. https://doi.org/10.3390/rs13010135

Schröder, J., & Gadow, K. von. (1999). Testing a new competition index for Maritime pine in northwestern Spain. *Canadian Journal of Forest Research*, *29*(2), 280–283. https://doi.org/10.1139/x98-199

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shainsky, L. J., & Radosevich, S. R. (1992). Mechanisms of competition between Douglas-fir and red alder seedlings. *Ecology*, *73*(1), 30–45. https://doi.org/10.2307/1938718

Staebler, G. R. (George R. (1951). *Growth and spacing in an even-aged stand of Douglas-fir* [Thesis]. http://deepblue.lib.umich.edu/handle/2027.42/114216

Sterba, H., Blab, A., & Katzensteiner, K. (2002). Adapting an individual tree growth model for Norway spruce (Picea abies L. Karst.) in pure and mixed species stands. *Forest Ecology and Management*, *159*(1), 101–110. https://doi.org/10.1016/S0378-1127(01)00713-7

Tomé, M., & Burkhart, H. E. (1989). Distance-dependent competition measures for predicting growth of individual trees. *Forest Science*, *35*(3), 816–831. https://doi.org/10.1093/forestscience/35.3.816

Trappe, J. M., Franklin, J. F., Tarrant, R. F., & Hansen, G. M. (1968). Biology of alder: Proceedings of a Symposium. *Fortieth Northwest Scientific Association Meeting. USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, OR*.

Tu, Y.-H., Johansen, K., Phinn, S., & Robson, A. (2019). Measuring Canopy Structure and Condition Using Multi-Spectral UAS Imagery in a Horticultural Environment. *Remote Sensing*, *11*(3), Article 3. https://doi.org/10.3390/rs11030269

USDA. (2018). *Web Soil Survey [online]*. Natural Resources Conservation Service, United States Department of Agriculture. https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx

Watson, P., & Nimmo-Smith, I. (2003, November 27). *Collinearity*. https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/Collinearity

Weiskittel, A. R., Garber, S. M., Johnson, G. P., Maguire, D. A., & Monserud, R. A. (2007). Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. *Forest Ecology and Management*, *250*(3), 266–278. https://doi.org/10.1016/j.foreco.2007.05.026

Wimberly, M., & Bare, B. B. (1996). Distance-dependent and distance-independent models of Douglas-fir and western hemlock basal area growth following silvicultural treatment. *Forest Ecology and Management*, *89*, 1–11. https://doi.org/10.1016/S0378-1127(96)03870-4

Wykoff, W. R., Crookston, N. L., & Stage, A. R. (1982). User's Guide to the Stand Prognosis Model. *Gen. Tech. Rep. INT-133. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station; 1982. 112 p.*, *133*. https://doi.org/10.2737/INT-GTR-133

Zach. (2020, November 4). K-Fold Cross Validation in R (Step-by-Step). *Statology*. https://www.statology.org/k-fold-cross-validation-in-r/

Zhao, J., Maguire, D. A., Mainwaring, D. B., & Kanaskie, A. (2014). Western hemlock growth response to increasing intensity of Swiss needle cast on Douglas-fir: Changes in the dynamics of mixed-species stands. *Forestry: An International Journal of Forest Research*, *87*(5), 697–704. https://doi.org/10.1093/forestry/cpu030