A Morphometric Approach to Modelling Coast Redwood Productivity at Hundalee Forest

A dissertation submitted in partial fulfilment of the requirements for the degree of Bachelor of Forestry Science with Honours

Blake Rapley

New Zealand School of Forestry

University of Canterbury

Christchurch, New Zealand

2022

Abstract

Coast redwood (*Sequoia sempervirens*) is an evergreen species native to the coast of California and Southwestern Oregon in the United States. Making up less than 1% of New Zealand's plantation estate, it is a minor commercial species in New Zealand. However, the species' characteristics of valuable timber, high volume production, and its ability to regrow from stumps after harvesting mean it could become more common as a plantation species.

This dissertation aimed to improve the prediction accuracy of the coast redwood 300 Index model developed by Watt et al. (2021). The model fails to accurately predict 300 Index at Hundalee Forest, a redwood plantation owned by The New Zealand Redwood Company. Only 6.7% of the variation between the observed and predicted 300 Index values are explained by the model. Using an approach similar to that of Salekin (2019), it would investigate how site morphometry influenced redwood productivity.

A combination of site, climate, and morphometric variables were analysed using RStudio to develop a model that could more accurately predict 300 Index at Hundalee Forest. The final model produced was:

*Observed 300 Index ~ Morphometric protection index + Topographic wetness index * Cosine of aspect*

This model produced an adjusted R-squared value of 51.4%, an improvement of 44.7% over the original redwood 300 Index model. Each of the variables in the models had statistically significant P-values (<0.05), which means that each of the variables in the model were significant predictors of redwood productivity at Hundalee Forest and that methods used in Salekin's study of eucalyptu*s* productivity influences applies to coast redwood.

While this study has achieved what it set out to, there were limitations to it. The main limitations were having a small dataset (n=31), which likely reduced the model's explanatory power and prevented validation, as well as some variables not being useful due to high correlations between layers sourced from the internet. In the future, it is recommended that The New Zealand Redwood Company investigates the effectiveness of the model produced on their other plantations around New Zealand and that morphometry is taken into consideration in future coast redwood productivity modelling.

Acknowledgements

Firstly, I would like to acknowledge my supervisor for this project, Mark Bloomberg. Mark’s advice and guidance throughout this process has been invaluable and I would not have been able to complete this project without it.

I would also like to thank Simon Rapley and The New Zealand Redwood Company for giving me the opportunity to take on this project, sharing their expert knowledge, and for helping me to pursue a career in forestry.

I would also like to acknowledge Luis Apiolaza, who coordinated the course and provided feedback for presentations, the contributions of Justin Morgenroth, Dave Palmer and Professor Elena Moltchanova who shared their knowledge in statistics and GIS.

Lastly, thanks to my family and friends who have supported me throughout this project, which has been quite stressful at times!

Table of Contents

1. Introduction

Currently, around 90% of New Zealand's 1.7 million hectares of productive plantation forest estate is radiata pine (*Pinus radiata*) (Ministry for Primary Industries, 2022). There is considerable interest in alternative plantation species such as Douglas fir (*Pseudotsuga menziesii*), eucalyptus (*Eucalyptus spp.*), cypresses (*Cupressus spp.*), and coast redwood (*Sequoia sempervirens*). These species are not only suitable for timber production but have traits that make them more suitable for mitigating erosion, alternative silviculture systems, aesthetic purposes, and economic values (Satchell, 2018).

Coast redwood is an evergreen species of the cypress family native to the coast of California and Southern Oregon. It is currently only a minor plantation species in New Zealand, behind radiata pine and Douglas fir, making up less than 1% of our plantation area. However, its use could become more frequent because of its favourable plantation qualities such as high timber value, volume production, ability to control erosion, and potential as a carbon sequestration crop. Because of the promising characteristics of the species, it is becoming a species of interest to forest growers, meaning there will be curiosity as to how redwood performs in terms of volume growth around New Zealand.

Productivity in New Zealand's plantations can be measured with 300 Index, which is the mean annual increment (MAI) (m^3/ha/year), at age 30 for a stand with 300 stems per hectare. 300 Index was developed by Kimberley et al. (2005) to improve on Site Index, which at the time was the standard way of measuring site productivity based on dominant tree heights at age 20. It was found that Site Index was not suitable to predict site quality in terms of volume productivity on its own. Site Index was weakly related to basal area growth, and the relationship between Site Index and volume growth was primarily because of the effect of tree height on volume. Findings also showed that height and diameter growth were not strongly related. 300 Index is therefore a more suitable measure of productivity, because it is a direct measurement of volume production, works irrespective of stocking, and accounts for age differences.

From growth models, geospatial productivity layers can be created, which let a user estimate the productivity for any given site. Watt et al. (2021) developed a 300 Index model and productivity layer for coast redwood to directly compare its productivity in New Zealand to radiata pine. An initial sample of 31 inventory plots from The New Zealand Redwood Company's Hundalee Forest in North Canterbury has shown that the redwood 300 Index layer is very unreliable at a local scale. Using the productivity layer, it was found that predicted 300 Index values at the plot locations at Hundalee Forest ranged from 1.7 to 16.6 m^3/ha/year, whereas observed productivity (calculated from the 31 inventory plots) ranged from 2.2 to 31.6 m^3/ha/year. Regression analysis determined that based on these plots, the coefficient of determination (R-squared value) between the predicted and observed productivity values was 0.067, meaning that only 6.7% of the variation in observed 300 Index was explained by the model.

2. Problem Statement and Objectives

2.1 Problem Statement

Coast redwood is a species that is extremely site specific. The 300 Index productivity layer recently developed by Watt et al. does not account for this specificity, and in the case of Hundalee Forest in North Canterbury, has largely misrepresented the productivity of coast redwood. A growth model that inaccurately predicts redwood growth at the forest level could cause several issues for The New Zealand Redwood Company (NZRC), such as inaccurate yield and revenue predictions and site selection decisions.

2.2 Objectives

This dissertation aims to improve the accuracy of the redwood 300 Index growth model/productivity layer made by Watt et al. (2021) at a local level by adding and testing predictor variables that will account for local site variation. Having a more accurate productivity model than what is currently available will aid The NZRC with making forecasts, determining site limitations, and site selection when acquiring new land for redwood afforestation. Ideally, the growth model that is developed as a part of this study will be effective for not only Hundalee Forest, but also for other sites owned by The NZRC, and other redwood growers around the country.

3. Review of Literature

Coast redwood is an evergreen species of the cypress family Cupressaceae native to the West Coast of North America (Watt et al., 2021). It is an important species for producing timber in New Zealand, behind radiata pine and Douglas fir (Rapley, 2018), and could become more common because of its high timber value and volume production, opportunity in overseas markets, and capability as a carbon sequestrating crop (Hocking, 2003, & Watt & Kimberley, 2021). As with all plant species, the productivity of redwood is a product of multiple factors. These factors include soil types and characteristics, climate, and other site factors (Maclaren, 2004). Given that radiata pine is such a dominant part of New Zealand's plantation forest industry, accounting for 90% of the estate, little research has been carried out on the volume productivity of redwood (Rapley, 2018 & Watt et al., 2021).

Tree productivity and health can be influenced by the topography/morphometry of the landscape and the microsites within it. Maclaren (2004) states that redwood is site-specific, requiring deep soils and can be killed and stunted by severe out of season frosts and wind. To measure the effect of different morphometric features, several morphometric indices have been developed, including wind exposition index, morphometric protection index, and topographic wetness index. Harris & Baird (2018) describe these topographic variables as:

- Wind exposition index (WEI) – the average wind effect across all directions
- Morphometric protection index (MPI) – the degree to which a point is protected/exposed by the relief surrounding it
- Topographic wetness index (TWI) – an index that estimates where water will accumulate in different areas of a landscape

Salekin et al. (2019) found that the height growth and survival of *Eucalyptus bosistoana* and *Eucalyptus globoidea* were significantly influenced by micro-topographic variables. MPI was positively correlated with growth, meaning sites more protected by the surrounding relief produced greater height growth. WEI had a negative correlation with height growth, which indicates that trees in microsites protected from the wind will experience less growth suppression.

Soils are a resource vital for plant growth with properties that vary spatially and temporally (Armson, 1977). A source of spatial variation in soil characteristics is topography, which can lead to changes in soil depth, nutrition, and moisture. This change is described by the catena concept, which refers to the variation of soils with similar parent materials existing in similar climatic conditions caused by differences in relief (Schaetzl & Anderson, 2005, as cited in Rosemary et al., 2017). In 2021, Salekin et al. found that soil properties were influenced by topography. Total carbon, total nitrogen, total phosphorous, extractable potassium, and hot-water extractable boron were positively correlated with MPI and negatively correlated with WEI. These findings suggest that the relief protection of the site and its exposure to wind were correlated with soil properties important for tree growth. The study also found that the levels of boron, an essential micronutrient for plant growth, were correlated with TWI and profile curvature.

Growth models are widely used in forestry as they are an important tool that helps foresters to predict the growth rates of sites and yields of plantation forests. A common approach to productivity modelling in New Zealand is Site Index, a method that describes the mean top height of dominant trees at age 20. Height was selected as an appropriate measure of productivity as it is largely unaffected by stocking and thinning, whereas taking a measure of stand volume can be greatly influenced by silvicultural practices. This is generally true for New Zealand radiata pine, although research has shown that height growth has a slight tendency to increase with stocking (MacLaren et al., 1995). Regardless, estimating productivity based on tree height was justified as it was a common belief that tree height and stem volume growth were strongly correlated. In contrast to this belief, research has shown that Site Index is very weakly related to basal area and that the relationship between Site Index and volume is mainly because of the effect of stem height on volume instead of stem height and basal area. Therefore, it was determined that Site Index did not fully measure site quality and that an index that is more closely related to stem volume would be more effective (Kimberley et al., 2005).

These limitations of Site Index led to the development of the 300 Index in 2005. 300 Index, which is defined as the mean annual increment (MAI) of volume growth for a defined silvicultural regime of 300 stems/ha at age 30, which was pruned to 6m after 3 pruning lifts. By having a fixed regime, the effects of stocking on volume were accounted for and silvicultural practices would not affect the stand volume, which overcomes the main challenge of using stand volume as a productivity index (Kimberley et al., 2005). Given that the 300 Index model can account for differences in regimes, its direct volume measure means that it should be a more suitable predictor of site quality than Site Index. Since its development in 2005, 300 Index has become New Zealand's standard measure of forest productivity and has been used widely for assessing site productivity. Its wide use throughout the forest industry suggests that is a superior productivity modelling tool to its predecessor, Site Index.

In 2021, a 300 Index model for New Zealand grown coast redwood was developed by Watt et al. (2021) so direct volume productivity comparisons could be made with radiata pine. The model estimates productivity as a function of several variables describing the climate, soil, site, and tree crop. The final regression model used for making the model contained the following variables:

- Number of ground frosts in January
- Establishment year
- Minimum air temperature
- Autumn vapour pressure deficit (VPD)
- Sin of aspect
- Cosine of aspect
- Cation exchange capacity

This model performed well with an R-squared of 0.66. However, each of the environmental variables used in the model were derived from surfaces with 25m of spatial resolution, so the model may fail to predict productivity accurately on a more local level where site factors may vary over small distances.

4. Site description

Hundalee Forest is a 5,525ha mixed coast redwood and Douglas fir plantation owned by The New Zealand Redwood Company. The forest is in the Conway Hills in North Canterbury, 125km northeast of Christchurch and 31km southwest of Kaikoura (Figure 1). The site ranges in altitude from 40-806m asl, and has a mean annual rainfall of 1050mm, although the Northern region of the forest behind the main range is drier (S. Rapley, personal communication, May 17, 2022). The mean annual temperature of the site is between 9 and 10 degrees Celsius. Soil maps show minor variations in soil types at Hundalee Forest, with Orthic Brown soils being the most dominant type. According to Hewitt et al., (2021) Orthic Brown soils have few limitations for productive land uses.

Hundalee Forest is a diverse site in terms of its quality for growing trees and has produced stands with a range of stand characteristics. Target stockings for planting at Hundalee Forest were 500stems/ha, although stockings range from 100-1000 stems/ha. The large discrepancies from the nominal stocking can be attributed to mortality in some parts of the forest and poor quality control during establishment in others. There has been no thinning in the forest.

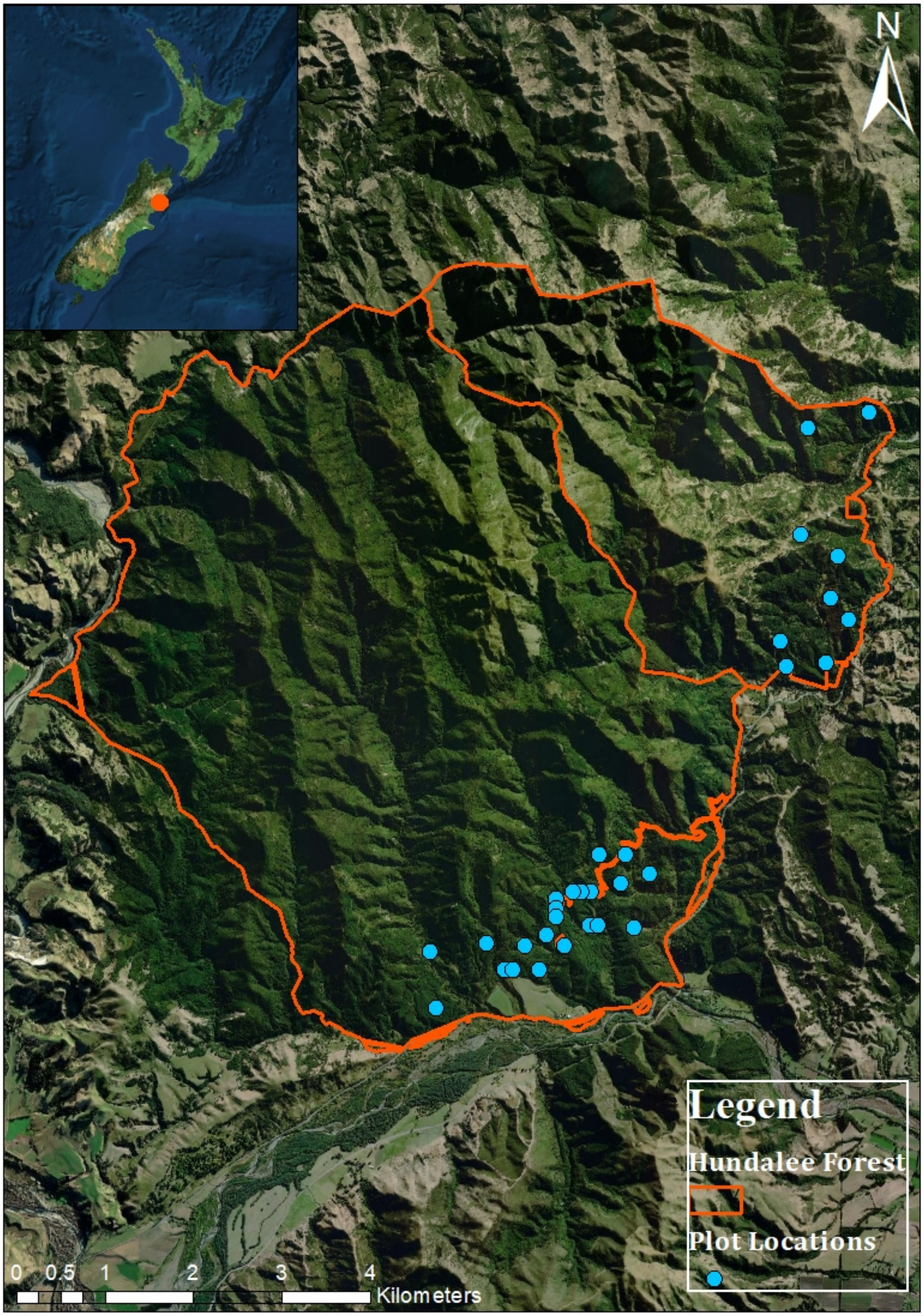

Figure 1: Hundalee Forest location

5. Methods

5.1 300 Index model performance

Evaluating the performance of the redwood 300 Index model produced by Watt et al. (2021) for Hundalee Forest was the first important step in this analysis. 300 Index values were calculated using plot data from Hundalee Forest collected by Buck Forestry in 2019 and 2021 for standard forest inventory purposes.

Table 1 shows the variation in stand characteristics across 31 inventory plots in the forest (plot locations are shown in Figure 1).

Table 1: Stand characteristics variation at Hundalee Forest

Stand characteristic	**Range**
Age (years)	12-17
Stocking (stems/ha)	167-940
Mean top height (m)	8.9-22.1
Basal area (m^2/ha)	2.4-58.7
Observed 300 Index (m^3/ha/year)	2.2-31.6
Observed Site Index (m)	17.2-37.8

Measurements for tree height and diameter from the 31 plots were input into the New Zealand Redwood Growth Model Excel sheet and the observed 300 Index was calculated for each plot.

To find the predicted 300 Index values for Hundalee Forest, the Redwood Productivity Surface for 300 Index was added to ArcMap (Esri, 2020). By entering the coordinates of each plot into ArcMap, the productivity predicted by the growth model for each plot was determined by taking the predicted 300 Index value from the redwood productivity layer.

With values for observed and predicted 300 Index at Hundalee Forest, a scatter plot and R-squared value was used to see the relationship between the two productivity estimates. There was a very weak correlation between the two estimates, and only 6.7% of the variation between the predicted and observed 300 Index values were explained by the model, which can be seen in Figure 2.

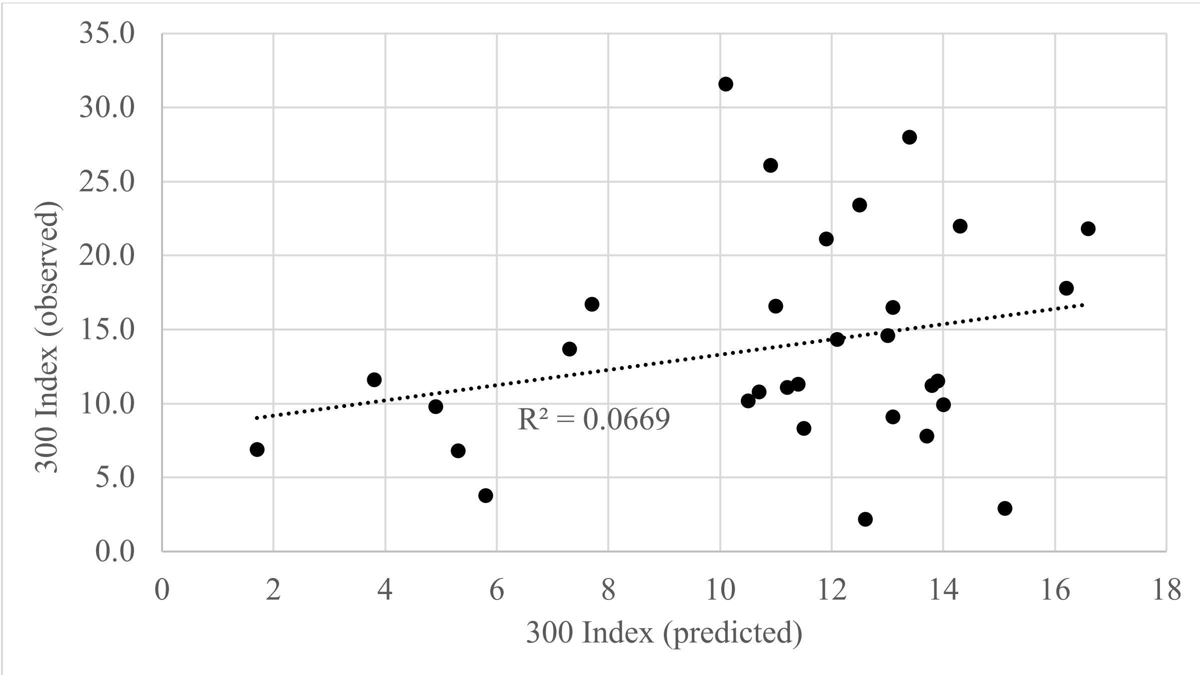

Figure 2: Predicted vs observed 300 Index at Hundalee Forest

5.2 Identifying candidate variables

The variables used in the original redwood 300 Index model made by Watt et al. in 2021 formed the basis of the variables used in the model, however, there were some changes to these variables. The layer for Autumn VPD was not accessible, so a more general annual VPD layer was used in its place.

To make improvements in the prediction accuracy of the model, variables that could be the source of this improvement needed to be identified. Research published by Salekin et al. in 2019 and 2021, as well as a visit to Hundalee Forest in 2021, suggested that morphometry would be a strong driver of productivity. Therefore, it was important to use some variables that would account for the changes in morphometry. Three morphometric indices were selected as candidate variables for this study. These were Wind Exposition Index (WEI), Morphometric Protection Index (MPI), and Topographic Wetness Index (TWI). These variables were used in the Salekin (2019) paper to determine how morphometry influenced height growth and survival of eucalyptus species, so it was assumed that these variables would also be useful to use in this study.

5.3 Data frame construction

With all candidate variables identified, layers of the variables were added into ArcMap. This included variables that were used in the original 300 Index model by Watt et al., as well as the morphometric indices identified by Salekin et al. that were created for Hundalee Forest using SAGA (Conrad et al., 2015). Making the layers for the morphometric indices (MPI, TWI & WEI) required using a digital elevation model (DEM) of Hundalee Forest which was supplied by The New Zealand Redwood Company. Using SAGA's geoprocessing tools, each morphometric index was generated for Hundalee Forest in the form of a layer with values for each 5 metre square.

Each surface was made with or converted to a spatial resolution of 5m in ArcMap. With all the variable surfaces and the plot locations on the map, the Extract Multi Values to Points tool was used to take the value from each surface of each plot and add them all into an attribute table. This attribute table was then converted into a CSV file for analysis in RStudio (R Core Team, 2021).

5.4 Identifying the best model

Identifying the best model was a process that required the use of RStudio and several statistical tools within it, such as variance inflation factor (VIF) and second-order Akaike's Information Criterion (AICc). The best model was one that increased the amount of variation explained by the model from the original Watt et al. model while accounting for the small (n=31) sample size, and not under or over-parameterizing the model. A loose rule for fitting models is that the number of inputs should be n/10, which suggests that the model for this dataset should contain no more than three independent variables (Professor E. Moltchanova, personal communication, September 7, 2022).

Correlations between variables were assessed as part of the model selection process. It is important that when fitting linear regression models the predictor variables are independent, and that there is a low correlation among them. Collinearity can cause problems as variables that should be independent become less so, leading to problems with fitting the model and interpreting its results. With high collinearity, it is hard for the model to estimate the relationship between each independent predictor variable and the dependent variable, as a change in one predictor will lead to changes in the others (Frost, n.d.). Collinearity was investigated with two methods. The first was a correlation plot produced in RStudio which provides a visual representation of the correlations between the variables in the dataset. The relationships between all the variables in the dataset are seen above in Figure 3. The circles provide a representation of the correlation strength, with a larger, bolder circle representing a stronger relationship, whereas the smaller and more faint circles show a weaker correlation. There was a range of correlations among the variables in the dataset, with some variables having almost no relationships and some having almost perfect correlations.

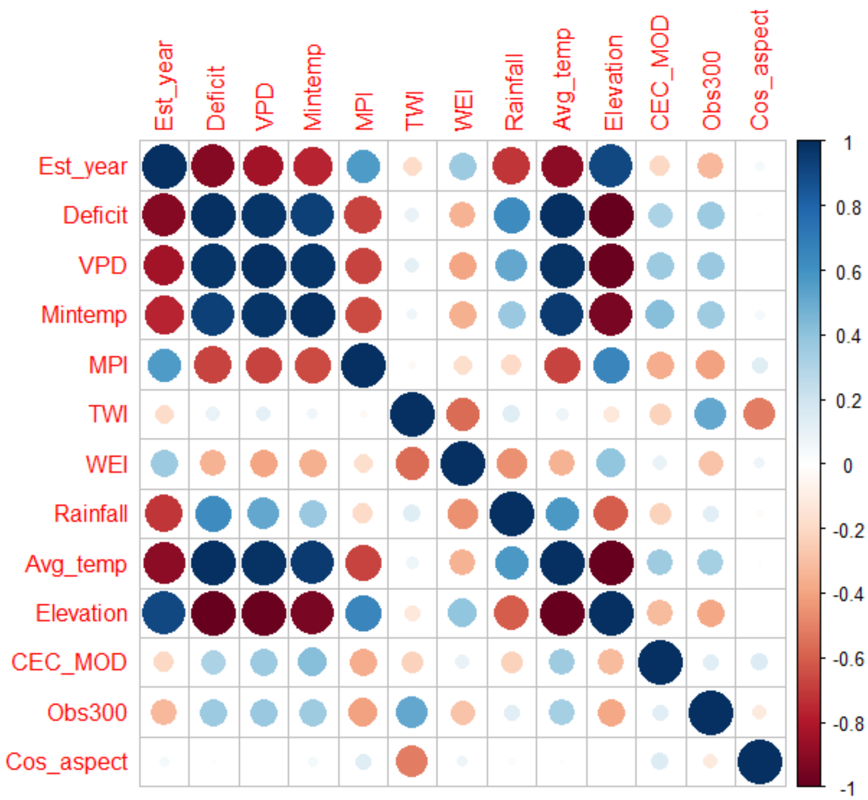

Figure 3: Correlations between variables

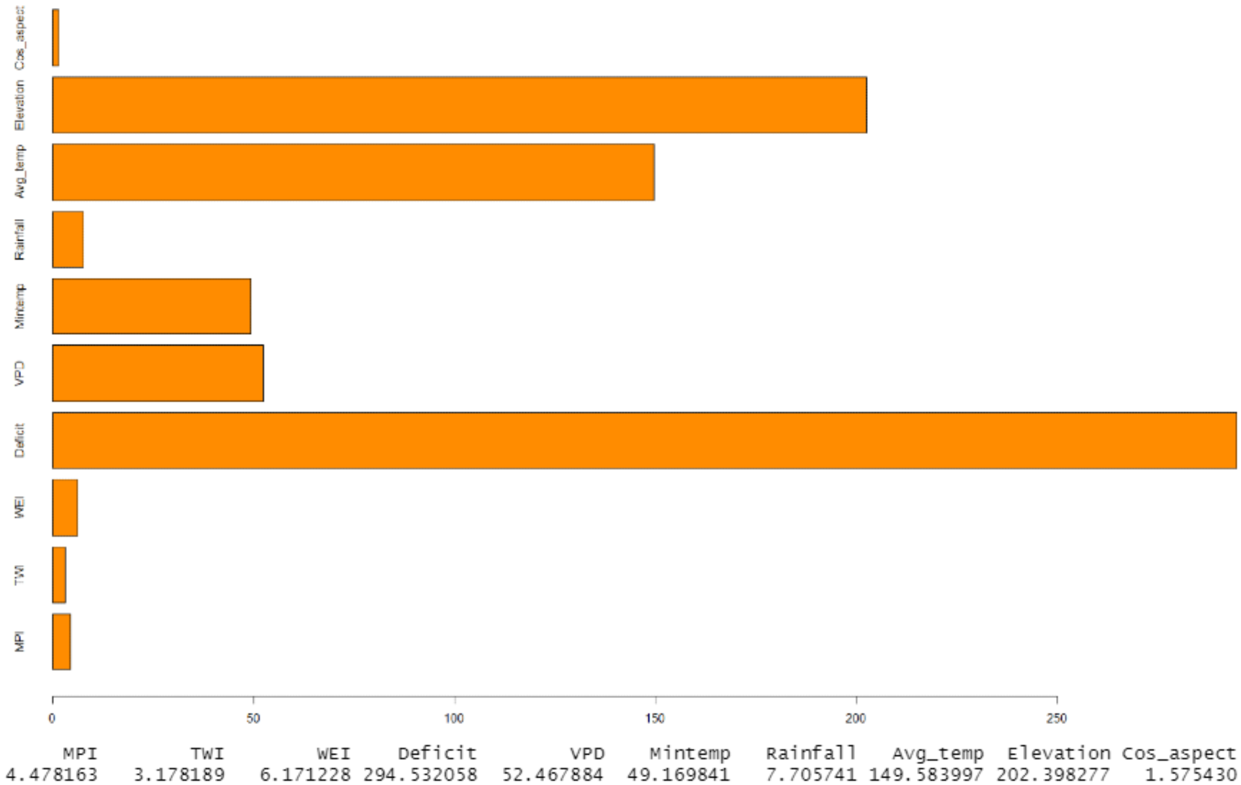

Figure 4: VIF values for continuous variables

Variance inflation factor (VIF) was used to determine the degree of multicollinearity (correlation between independent variables) in the dataset. Figure 4 depicts the results of the VIF analysis. The VIF analysis does not include the variables establishment year or cation exchange capacity because these variables are categorical and therefore, VIF cannot be used. According to James et al. (2013), VIF can be interpreted using a rule which suggests that:

<5 = low correlation between predictors

5-10 = moderate correlation between predictors

>10 = high correlation that is not tolerable for predictors

Based on this rule, some variables were extremely highly correlated, with VIF values up to 294.5. Therefore, annual water deficit, annual VPD, minimum annual temperature, and average annual temperature cannot be included in the model as the correlations could negatively affect the results of the regression. Although elevation has a VIF value of 202, it is reduced to a usable level when the aforementioned variables were removed from the dataset.

Another measure used to assess the performance of the model was Akaike Information Criterion corrected for small samples (AICc). Because the dataset used for this study is small (n=31), it was determined that the bias-corrected version of the standard AIC was a better way to analyse the model performance (Professor E. Moltchanova, personal communication, September 7, 2022). AICc is a way to interpret the fit of a model to the data that was used to make the model. Using AICc, comparisons can be made between models to determine which fits best, based on the amount of variation explained with the fewest independent variables. The output of AICc is a value, for which a smaller number indicates a better-fitting model. When a model has an AICc value that is less than another model by 2 or more, it is considered an improved model (Bevans, 2022).

When fitting the models, performance was based on the adjusted R-squared and AICc values. Initially, the variables with VIF values above 10 were removed. Following the initial removal of overly correlated variables, the P-values of variables were used to evaluate which variables should be taken out of the model. The P-value indicates how much of an effect the variable has on the dependent variable by testing the null hypothesis that there is no correlation between the two variables. A less significant P-value suggests that the variable is less effective in the model and was a basis for removing variables. Each model had its performance assessed with adjusted R-squared and AICc. A higher adjusted R-squared and lower AICc value showed improvement from one model compared to another.

In statistics, validation is an important step for confirming the performance of a model and making sure the model achieves its purpose. Model validation involves taking a different dataset from the one the model was generated with and testing the model on that dataset. While validating the model made for predicting observed 300 Index at Hundalee Forest would be a worthwhile process that would ensure the final model is effective across a range of datasets, constraints around time for collecting additional data and few available plot data means that validation was not feasible for this model.

6. Data

Each of the variables used in the analysis are shown in Table 2, along with a brief description of the variable. These variables include the dependent response variable, observed 300 Index, and the independent explanatory variables.

Table 2: Variables investigated for the analysis

Data	Abbreviation	Source	Description
Inventory data	-	NZRC	Plot measurements including DBH and height.
Observed 300 Index	Obs300	NZRC	Observed 300 Index at plot sites
Predicted 300 Index	Pre300	Watt et al. model	300 Index predicted by Watt et al. model
Wind Exposition Index	WEI	SAGA	A measure of how exposed a point is to wind. A value >1 means wind exposed and <1 means wind shadowed.
Morphometric Protection Index	MPI	SAGA	A measure of how protected a point is by the relief surrounding it. A value >0 means the point is protected and a value <0 means it is not.
Topographic Wetness Index	TWI	SAGA	Describes the propensity of an area to accumulate water. A higher value represents a wetter area.
Establishment year	Est_year	NZRC	The year that trees were planted.
Cosine of aspect	Cos_aspect	ArcGIS	Cosine of compass direction. The cosine of aspect is a

			measure of 'northerliness', where 1 = north and -1 = south.
Cation exchange capacity	CEC_MOD	LRIS	The total amount of cations that can be held in the soil.

Generally, the variables are commonly used for describing factors such as climate and morphometry. However, the reason for some variables being in the dataset is less obvious. An example is establishment year, which in the case of Hundalee Forest, relates to the quality of redwood clones being planted and the standard of planting itself (S. Rapley, personal communication, September 12, 2022).

Table 3: Key statistics for variables used in the analysis

	Minimum	**Maximum**	**Mean**	**Std. Dev.**	**CV**
Deficit (mm)	51	123	99	21	21.2%
MPI	0.04	0.34	0.13	0.06	46.2%
TWI	3.22	8.04	5.06	1.28	25.3%
WEI	0.86	1.11	1.02	0.05	4.9%
Rainfall (mm)	968	992	977	5	0.5%
Elevation (m)	62	279	141	62	44%
Obs300 (m^3/ha/year)	2.2	31.6	13.9	7.2	51.8%
Cos_aspect	-1.00	0.88	-0.43	0.57	-132.6%

Important statistics of the variables are depicted in Table 3. These statistics provide an indication as to how much variation there is in each of the variables in the dataset. Variability is measured by the coefficient of variation (CV) percentage. A larger percentage indicates more variation relative to the mean. There were significant differences in the variability of the data, with CV values ranging from 0.5% to -132.6%.

7. Results

A maximal additive model (m1) was progressively simplified, using model R-squared, AICc and P-values for model parameters as criteria. Four models (m1-m4) were fitted as follows:

m1 = Obs300 ~ MPI + TWI + WEI + Elevation + Rainfall + CEC_MOD + Cos_aspect

m2 = Obs300 ~ MPI + TWI + Rainfall + Cos_aspect

m3 = Obs300 ~ MPI + TWI + Cos_aspect

m4 = Obs300 ~ MPI + TWI * Cos_aspect

The results of the model fitting are shown in Table 4.

Table 4: Model fitting process

Model	Adjusted R-squared	AICc
m1	0.325	216.6
m2	0.369	209.4
m3	0.413	203
m4	0.514	199.1

The first model, m1, is the maximal additive model with all predictor variables that were too highly correlated removed. Because of their high P-values, WEI, Elevation, and CEC_MOD were removed to create m2. This model showed improvement over m1 in terms of having a higher adjusted R-squared value and lower AICc. Again, based on high P-values, rainfall was removed from the model to make m3, which further improved the model performance. This model achieved the target of 3 predictor variables, so interaction terms were added to try and improve the model's performance even more. An interaction between TWI and cos_aspect finalized the model with m4. This model had the highest adjusted R-squared and lowest AICc value and was the best-performing model created for predicting observed 300 Index.

The final model produced was:

*Observed 300 Index ~ Morphometric protection index + Topographic wetness index * Cosine of aspect*

This model produced an adjusted R-squared value of 0.514, meaning that 51.4% of the variation in the model has been accounted for. Compared to the original model produced by Watt et al. in 2021, the percentage of variation explained by the model has increased by 51.4% from the Watt model's R-squared of 6.7%. This change in adjusted R-squared represents a significant increase of 44.7% in the prediction accuracy of the model.

Table 5: Key statistics of the final model

Statistic	Value
Residual standard error	5.106 on 26 degrees of freedom
Adjusted R-squared	0.514
P-value	<0.001
AICc	199.1

The original additive model with all variables suitable to be used in the analysis based on collinearity produced an AIC value of 216.6. By adjusting what variables were used in the model and changing the model from a fully additive model to an additive model with an interaction term, the AICc was reduced to 199.1.

Table 6: Statistics of each variable in the final model (m4)

	Estimate	Standard Error	t value	Significance
(Intercept)	-13.461	7.852	-1.714	NS
MPI	-41.770	14.905	-2.802	**
TWI	7.421	1.690	4.391	***
Cos_aspect	-18.629	8.803	-2.116	*
TWI:Cos_aspect	4.977	1.934	2.574	*

*Significance level (*** = $P < 0.001$; ** = $P < 0.01$; * = $P < 0.05$; NS = $P \geq 0.05$).*

The statistical significance, measured by P-values, of the variables in the final model are displayed in Table 6. A P-value less than 0.05 implies statistical significance, so it can be said that each of the variables used in the final model were significant predictors of observed 300 Index at Hundalee Forest.

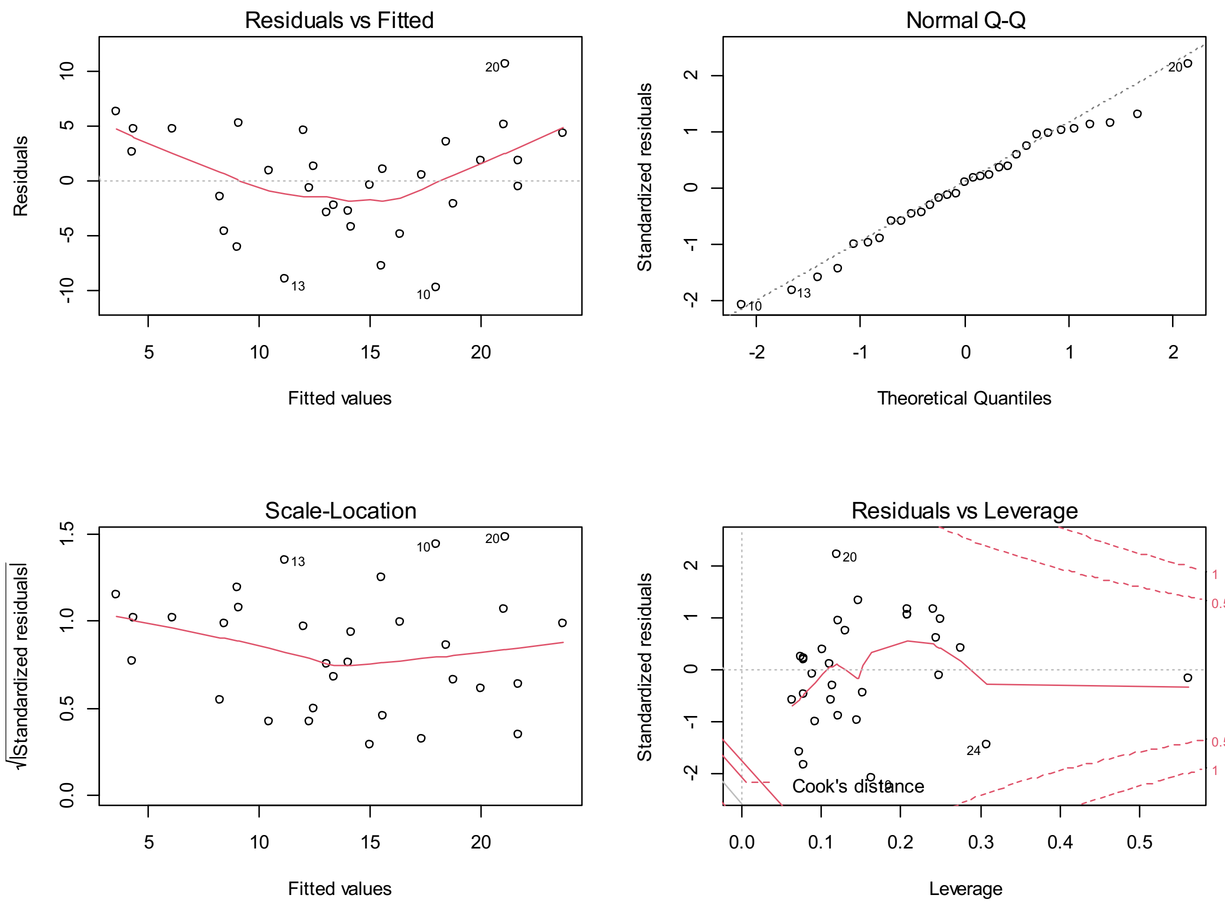

Figure 4: Residual plots for the final model

With linear regression modelling, four key assumptions must be met to ensure that the model is unbiased. These four assumptions are:

1. Linearity – there must be a linear relationship between the independent and dependent variable
2. Independence – the residuals are independent
3. Equal variance – the residuals have constant variance
4. Normality – the residuals are normally distributed

The residual plots in Figure 4 analyse the residuals of the model. The residuals vs fitted plot shows the linearity of the model. The red line is curved, which indicates that the model does not follow a linear pattern. The normal Q-Q plot is used to determine if residuals are normally distributed.

Generally, the points follow the dotted line which indicates that the residuals were normally distributed, despite some outliers that slightly deviate from the diagonal line. The scale-location plot shows the homoscedasticity, or how constant the variance is. A flat red line means that the assumption of equal variance is met which looks to be the case for this model, even though there is a section of the line that is not flat. The residuals vs leverage plot is used to identify any observations that have a strong influence on the coefficients of the model if they were excluded from the dataset. As none of the points fall outside of Cook's distance (the red dashed lines), there were no influential points in the regression model.

A way to get around the issue of assumptions not being met is with variable transformations. By transforming variables to their log, square root, and cube root forms, the variables can become closer to being normally distributed. However, when these transformations were applied in this analysis, it was found that the residuals of the model did not improve, and in some cases, the model's ability to predict observed 300 Index was reduced by the transformed variables.

The relationships between predicted 300 Index and observed 300 Index for the two different models are shown in Figures 5 and 6. The growth model developed by Watt et al. (2021) shows a very weak correlation between the two variables, with an R-squared value of 6.7%. In contrast, the new morphometric model for predicting 300 Index at Hundalee Forest has a much stronger correlation with the observed 300 Index values. This can be seen by the typically smaller deviation from the trendline and adjusted R-squared value of 51.4%.

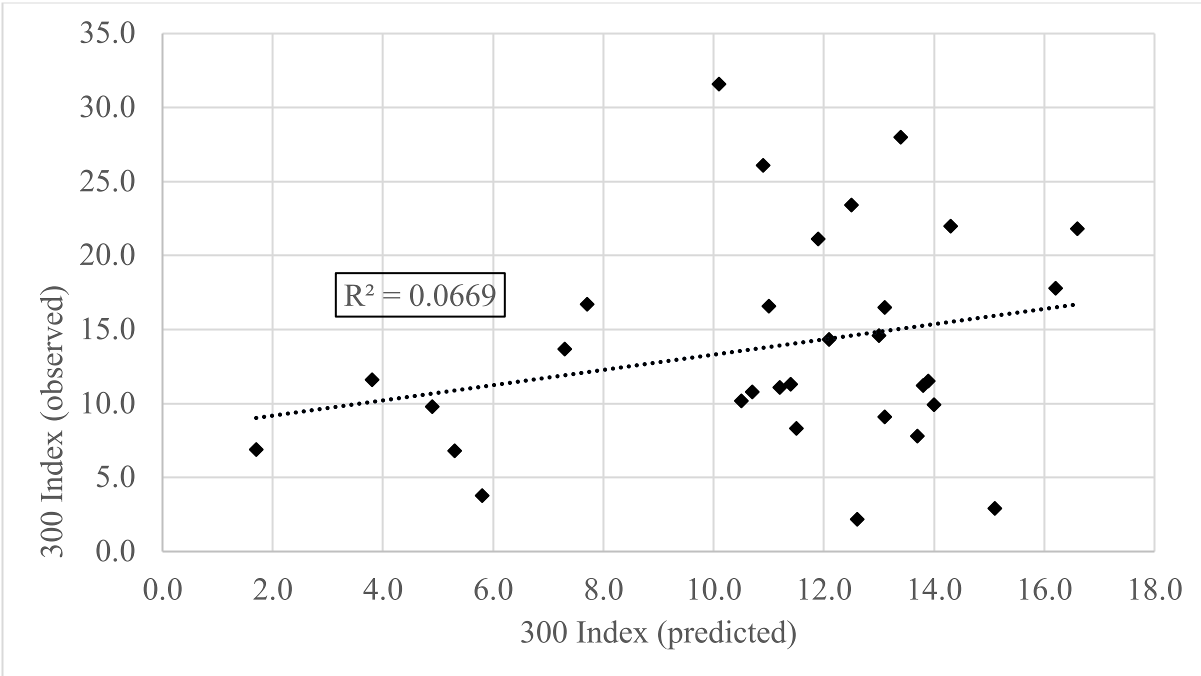

Figure 5: 300 Index predicted by the original Watt model vs observed 300 Index

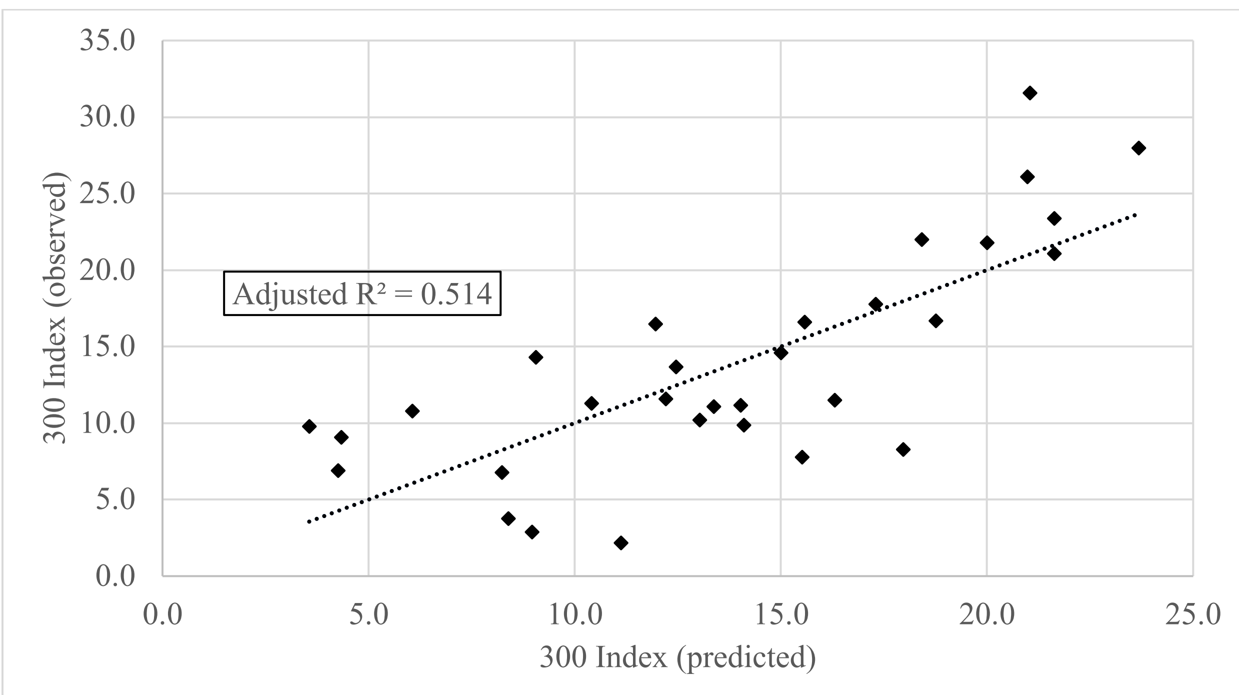

Figure 6: 300 Index predicted by the new morphometric model vs observed 300 Index

8. Discussion

The results of this study have accomplished the main objective of this dissertation by increasing the prediction accuracy of the coast redwood 300 Index model at Hundalee Forest. The new morphometric model developed improved the accuracy of the 300 Index prediction by 44.7%, from an initial performance of 6.7% compared to the new prediction accuracy of 51.4%. Another important finding from this study is that Salekin's 2019 study of how eucalyptus species' height and survival respond to the effects of morphometric variables on drier sites applies to coast redwood on a wetter site. This conclusion can be made as it was found that both morphometric protection index and topographic wetness index were statistically significant predictors of redwood productivity at Hundalee Forest.

While this study did produce superior results overall, some limitations might have prevented the results from being better. Firstly, one of the limitations of the study was the degree of correlation between five of the variables. It was found the climatic variables sourced from LRIS such as VPD, annual water deficit, minimum temperature and average temperature were highly correlated with elevation, with correlations above 95%. This is because these layers were typically created using elevation as a predictor variable (Leathwick et al., 2002). The consequence of having highly correlated predictor variables is that they were not suitable for use in the model due to lack of independence, so the effect of climate on redwood productivity at Hundalee Forest could not be explored fully. While these climate variables were used in the national redwood 300 Index model produced by Watt et al. (2021), they were not useful in this forest-level study of the drivers of redwood productivity. Climate variables also had limited use in this analysis as the layers vary over such a large scale that is not accounted for at the forest level. The CV values in Table 3 show that there is little variation among the climatic variables at Hundalee Forest, so these variables may have been limited in their ability to explain what drives productivity at Hundalee Forest. A further limitation related to layers derived from LRIS is that some layers were just not very useful in general. The main example of this limitation is with the cation exchange capacity layer, which suggested that Hundalee Forest had just three soil types and therefore three cation exchange capacity values. Soil types were grouped into polygons and mapped at approximately 1:50,000 scale, which is a very coarse resolution that does not provide much detail about soil properties at the site level. Since there was such little variation in cation exchange capacity described by the soil map, it was not found to be a very useful predictor of redwood productivity in this analysis.

Another limitation of this study was the small plot dataset available. Having only 31 plots available meant that the model would be restricted to having only around 3 predictor variables, as suggested by a general rule that a model should have one predictor variable for every 10 observations in the dataset. The implications of this are that a better model with more explanatory power could have been produced if there was an available dataset large enough to support more predictors of observed 300 Index. Also, the small dataset imposed constraints around having a fully interactive model. Having a full interaction model was not possible because there were not enough degrees of freedom, and a full interaction model did not work. Again, a larger dataset would have prevented this problem and could have helped with further increasing the ability of the model to accurately predict observed 300 Index. Despite having a small dataset which reduced the amount of statistical analysis able to be completed, this model has shown a way forward for further redwood productivity modelling. The results produced have shown that morphometric factors can be successful predictors of redwood productivity, and this should be acknowledged in the future.

It is uncertain how accurately the new model will predict observed 300 Index at other redwood sites around New Zealand. Site quality for forestry varies throughout New Zealand, and morphometry could be more or less important for explaining redwood productivity at other sites than Hundalee Forest. Plot data from other sites would need to be accessed to determine how a morphometric model performs at other locations, and is beyond the scope of this dissertation. Although it is unknown how the model will perform at sites other than Hundalee Forest, what can be said is that the model created for this dissertation has shown that morphometry is an important factor for redwood productivity, and that its effect on productivity should be considered as an approach when modelling redwood growth at other sites in the future.

A study conducted with fewer constraints around time and budget could produce a redwood productivity model that is further improved by gaining access to measured variables from sites around New Zealand. Such a study could ensure that a model applies to sites around the country rather than just Hundalee Forest, while also being able to fully explore the effect of different variables without being affected by high degrees of correlations from layers created as a function of elevation.

9. Recommendations

My recommendation for The NZRC is that they test the effectiveness of morphometric 300 Index modelling on their other New Zealand plantations located in the Central North Island. If it is found that the morphometric model is also more accurate than the existing redwood 300 Index model, The New Zealand Redwood Company will have access to a more effective growth model that can be used in their current operations and when looking to expand their estate. Further research into the effectiveness of a growth model that includes morphometric variables as predictors could also lead to a more accurate nationwide growth model being developed.

10. Conclusion

This study aimed to improve the ability of the coast redwood 300 Index model developed by Watt et al. (2021) to accurately predict redwood productivity at Hundalee Forest, a redwood plantation owned by The NZRC in North Canterbury. Results have also shown that the dryland eucalypt study by Salekin (2019) which explored the effects of morphometry on eucalyptus height growth and survival applied to coast redwoods as the productivity of redwood was influenced by morphometry. By completing a statistical analysis looking into the effect of climatic, site, and morphometric variables on redwood productivity, the final model for predicting redwood growth at Hundalee Forest was produced using the following variables:

*Observed 300 Index ~ Morphometric protection index + topographic wetness index * cosine of aspect*

This model improved the accuracy of the prediction, measured by adjusted R-squared, to 51.4%, an improvement of 44.7% from the original model's R-squared value of 6.7%.

This study had limitations, primarily a small dataset and high correlations among several of the climatic predictor variables. These issues could have prevented the model from being a more accurate predictor of redwood productivity, as well as imposing limitations around validating the effectiveness of the new model. It is uncertain how well the model will perform at other redwood plantations around New Zealand. Given that this study focused on Hundalee Forest, the model may not apply to all New Zealand redwood forests. Despite the limitations of the study, this research has improved the prediction accuracy of redwood 300 Index at Hundalee Forest and found that the site morphometry can be a statistically significant factor for predicting redwood productivity. Therefore, these factors should be accounted for when modelling coast redwood productivity in the future.

11. References

Armson, K. A. (1977). *Forest Soils: Properties and Processes*. University of Toronto Press. http://www.jstor.org/stable/10.3138/j.ctt15jjd7c

Bevans, R. (2022). *Akaike Information Criterion | When & How to Use It (Example).* Scribbr. https://www.scribbr.com/statistics/akaike-information-criterion/

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J. (2015): System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geosci. Model Dev., 8, 1991-2007, doi:10.5194/gmd-8-1991-2015.

Esri. (2020, February 20). *ArcMap* (10.8.1) [Software]. https://www.esri.com/en-us/arcgis/products/arcgis-desktop/resources

Frost, J. (n.d.). *Multicollinearity in Regression Analysis: Problems, Detection, and Solutions.* Statistics by Jim. https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/

Harris, A., & Baird, A. J. (2018, December 18). Microtopographic Drivers of Vegetation Patterning in Blanket Peatlands Recovering from Erosion. *Ecosystems*, 22(5), 1035–1054. https://doi.org/10.1007/s10021-018-0321-6

Hewitt, A.E., Balks, M.R., Lowe, D.J. (2021). Brown Soils. In: *The Soils of Aotearoa New Zealand*. World Soils Book Series. Springer, Cham. https://doi.org/10.1007/978-3-030-64763-6_4

Hocking, J. (2003, January 1). Alternative tree species on farms. NZGA: Research and Practice Series, 10, 91–101. https://doi.org/10.33584/rps.10.2003.2980

James, G., Witten, D., Hastie, T., and Tibshirani, R. (eds.). (2013). *An introduction to statistical learning: with applications in R*. New York: Springer.

Kimberley, M., West, G., Dean, M.G. & Knowles, L.R. (2005). The 300 index - A volume productivity index for radiata pine. *New Zealand Journal of Forestry*. 50. 13-18. http://www.nzjf.org.nz/free_issues/NZJF50_2_2005/8B99F9C0-92AB-49C3-94C8-EAA8759DC2F2.pdf

Leathwick, John & Morgan, Fraser & Wilson, Gareth & Rutledge, Daniel & Mcleod, Malcolm & Johnston, Kirsty. (2003). Land Environments of New Zealand: A Technical Guide.

Maclaren, J. P., Grace, J. C., Kimberley, M. O., Knowles, R. L., & West, G. G. (1995). Height growth of Pinus radiata as affected by stocking. *New Zealand Journal of Forestry Science*, *25*(1), 73-90. https://www.scionresearch.com/__data/assets/pdf_file/0018/17370/NZJFS2511995MACLAREN73-90.pdf

Maclaren, P. (2004). *Realistic alternatives to radiata pine in New Zealand - A critical review*. Report No. 90. Forest and Farm Plantation Management Cooperative. https://fgr.nz/documents/download/2598

Ministry for Primary Industries. (2022, May 5). *About New Zealand's forests.* https://www.mpi.govt.nz/forestry/new-zealand-forests-forest-industry/about-new-zealands-forests/

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rapley, S. (2018). Redwood In New Zealand. *New Zealand Journal of Forestry*, *63*(1), 29-33. https://www.nzjf.org.nz/free_issues/NZJF63_1_2018/FD170ED1-3709-4192-899B-BED5CBA1D91C.pdf

Rosemary, F., Vitharana, U., Indraratne, S., Weerasooriya, R., & Mishra, U. (2017). Exploring the spatial variability of soil properties in an Alfisol soil catena. *CATENA*, *150*, 53–61. https://doi.org/10.1016/j.catena.2016.10.017

Salekin, S., Bloomberg, M., Morgenroth, J., Meason, D. F., & Mason, E. G. (2021). Within-site drivers for soil nutrient variability in plantation forests: A case study from dry sub-humid New Zealand. *CATENA*, *200*, 105149. https://doi.org/10.1016/j.catena.2021.105149

Salekin, S., Mason, E. G., Morgenroth, J., Bloomberg, M., & Meason, D. F. (2019). Modelling the Effect of Microsite Influences on the Growth and Survival of Juvenile Eucalyptus globoidea (Blakely) and Eucalyptus bosistoana (F. Muell) in New Zealand. *Forests*, *10*(10), 857. https://doi.org/10.3390/f10100857

Satchell, D. (2018). *Trees for steep slopes.* New Zealand Farm Forestry Association/Forest Owners Association. https://www.nzffa.org.nz/system/assets/3025/Report_-_Trees_for_steep_slopes.pdf

Watt, M. S., Kimberley, M. O., Rapley, S., & Webster, R. (2021, November). Comparing volume productivity of redwood and radiata pine plantations in New Zealand. *Forest Ecology and Management*, *500*, 119628. https://doi.org/10.1016/j.foreco.2021.119628