Thesis Report

on

FACTORS DETERMINING DIAMETER GROWTH OF SHOREA ROBUSTA (SAL)

Agriculture and Forestry University

Forestry Campus, Hetauda

By

Binita Ghimire

B.Sc. Forestry

Fourth Year (Second Semester)

2023-08-17

Contents

List of Tables

List of Figures

1 Methodology

1.1 Study Area

1.2 Source of Data

1.3 Data analysis

2 Results

Tree diameter growth from different permanent sample plots.

2.1 Correlation for Feature Selection

Correlation method is simplest method, widely used for the variables having linear relationship, feature selection for modelling.

Table 1: Correlation of different features to basal area growth

Features	correlation coefficient
tree_density	0.50
min_temperature	-0.17
max_temperature	-0.16
crown_cover	0.16
elevation_nepal	0.15
aspect	0.15
ba_ha_old	0.14
lat	0.13
organic_layer_thickness	0.12
precipitation_accumulation	0.12
lon	-0.10
slope.map_nepal	0.09
climate_water_deficit	-0.08
wind_speed_at_10m	-0.07
palmer_drought_sensitive_index	0.06
organic_layer_type	0.04
soil_moisture	0.03
soil_depth	-0.01

Features	correlation coefficient
penetration_depth	0.00

2.2 Stepwise Regression

This method involves fitting a regression model with all potential predictor variables and then alliteratively adding or removing variables based on their significance until an optimal subset is obtained.

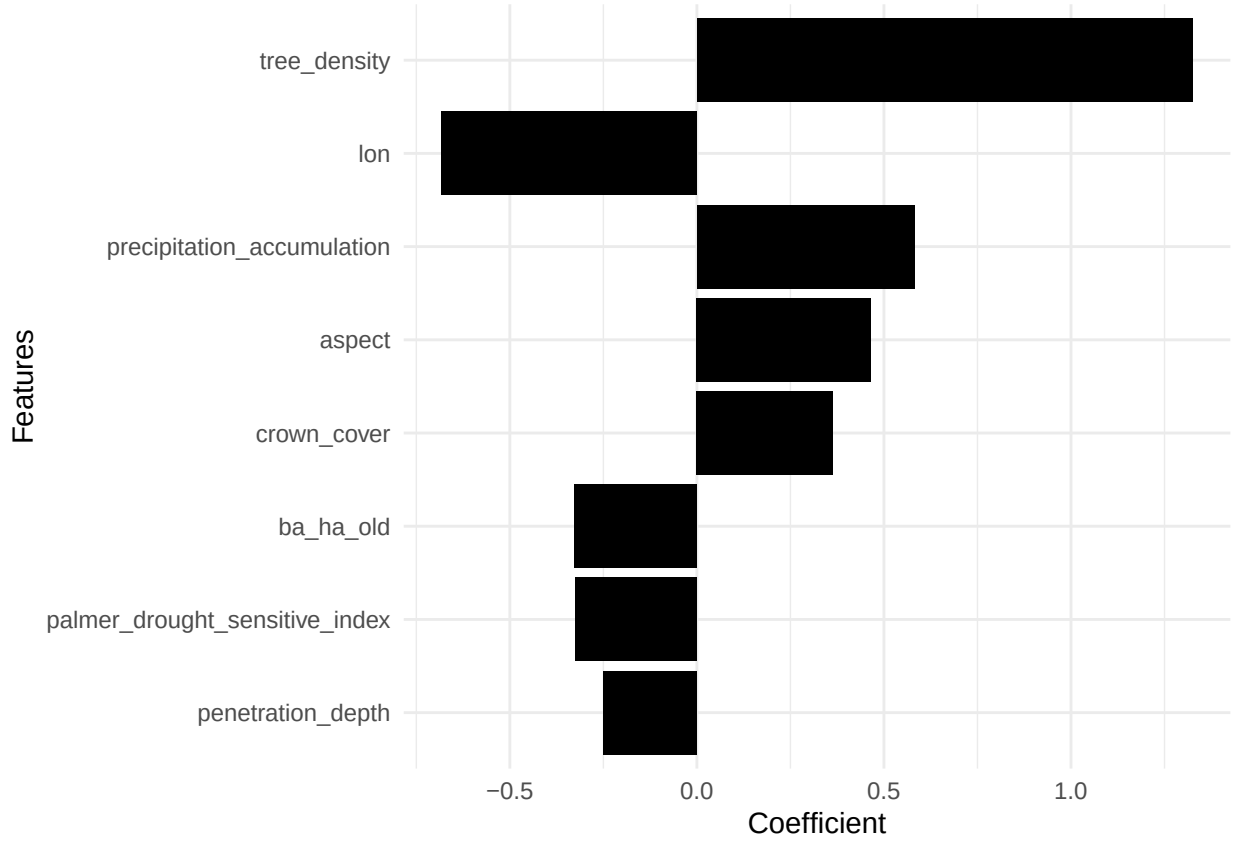

Figure 1: Feature and thier coefficients(in modulus) in order

3 Low variance filtering

The feature selection method "removing the factors having the lowest variability," is commonly known as "low variance filtering" or "constant feature removal." It is a straightforward technique used to eliminate features with very little or no variability in their values. In the context of feature selection, features with low variance do not contribute much information to the model because their values remain almost constant across all observations or samples. Such features are less likely to provide meaningful insights and might even add noise to the model, potentially leading to over fitting.

Table 2: Results from low variance filtering

Variables	Coefficient of variables
organic_layer_type	150.94
organic_layer_thickness	111.44
slope.map_nepal	97.81
tree_density	89.33
aspect	81.68
elevation_nepal	62.67
palmer_drought_sensitive_index	62.53
penetration_depth	46.44
ba_ha_old	39.48
soil_moisture	39.07
precipitation_accumulation	26.83
crown_cover	26.42
soil_depth	21.57
climate_water_deficit	21.28
min_temperature	9.51
max_temperature	6.66
wind_speed_at_10m	4.07
lon	2.54
lat	2.42

3.1 Multicollinearity Test

Table 3: Selected variables with thier VIF values

variable	VIF values
aspect	1.32
tree_density	1.43
organic_layer_type	1.54
organic_layer_thickness	1.66
crown_cover	1.70
wind_speed_at_10m	1.70
ba_ha_old	1.74

variable	VIF values
penetration_depth	1.79

3.2 RandomForest Regressor

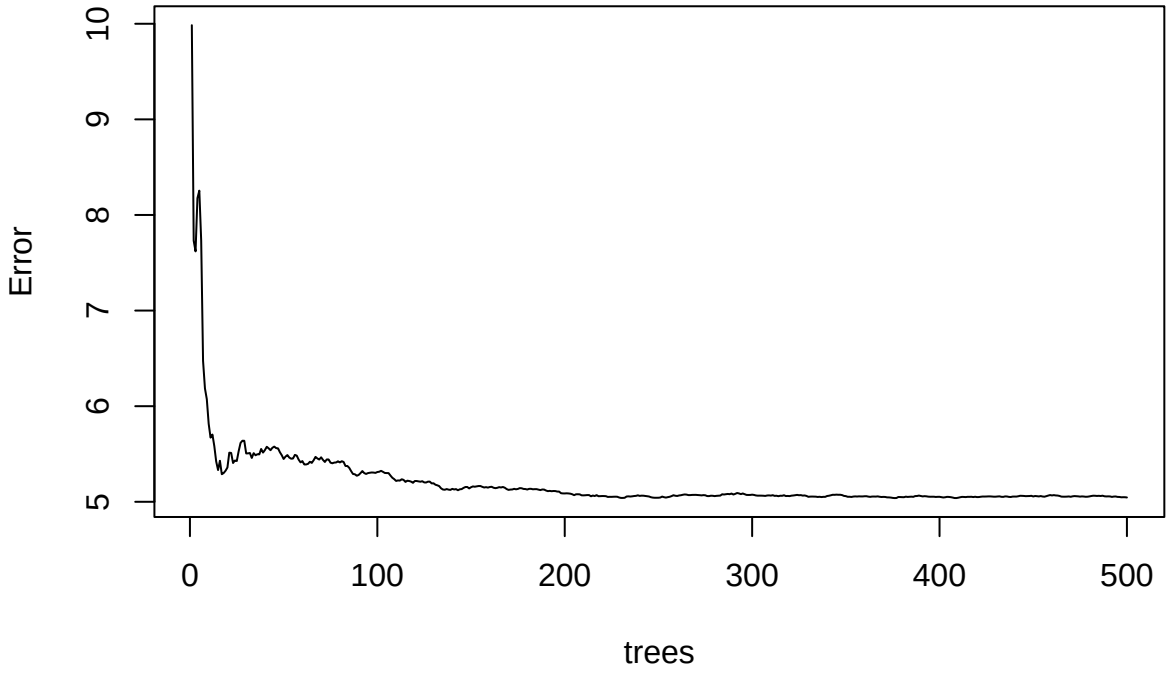

Table 4: Importance of feature in the model

Features	IncNode Purity Value
tree_density	264.713795
ba_ha_old	89.201379
penetration_depth	71.913346
lat	71.604930
min_temperature	57.178158
lon	55.333382
wind_speed_at_10m	53.999507
soil_moisture	53.205255
crown_cover	52.048419

Features	IncNode Purity Value
max_temperature	51.911490
elevation_nepal	50.208452
palmer_drought_sensitive_index	46.548952
slope.map_nepal	44.020634
climate_water_deficit	43.623167
precipitation_accumulation	42.737837
organic_layer_thickness	11.401235
soil_depth	10.999894
aspect	9.834277
organic_layer_type	9.390413