Forest Research Centre, School of Agriculture, University of Lisbon
ForChange Technical Publications, TP1-2025

Modelling crown variables for maritime pine in Portugal: data and script examples

Muha Abdullah Al Pavel[1,2], Susana Barreiro[1], Margarida Tomé[1*]

Author affiliations:

[1]Forest Research Centre, Associated Laboratory TERRA, School of Agriculture, University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal
[2]Centre for Applied Ecology "Prof. Baeta Neves" (CEABN/InBIO), School of Agriculture, University of Lisbon, Tapada da Ajuda, 1349-017, Lisbon, Portugal

* Corresponding author: Margarida Tomé, magatome@isa.ulisboa.pt

1. Introduction

This technical publication aims to provide some descriptions/instructions for the R scripts that were used to develop a set of compatible models to predict tree crown variables (crown ratio, crown length and height to the base of the crown) for maritime pine (*Pinus pinaster* Aiton.) in Portugal described in Al Pavel et al (2025).

2. Data used

Data used in this study came from the database of permanent plots and trials available in Portugal for maritime pine. A total of 83230 tree measurements, from 9 different trials, was available in the data set after excluding plots in which trees had been pruned, dead trees and the trees located in the border zone of trial plots. Additionally, once the stand variables and competition indices were calculated, trees without the measurement of the height to the base of the crown were deleted. A detailed analysis of the tree data for each plot led to the exclusion of 47 additional points considered to be measurement errors. The final data set used to develop the crown ratio and crown length models included 13151 tree measurements covering a wide range of site indices (*S*), stand ages (*t*), and stand densities per ha (*N*) and is available in the GitHub repository of ForChange-CEF (*Pb.hbc.Rdata*):

https://github.com/ForChange-CEF/ModellingCrownVarMaritimePine_DataScriptExamples

The file contains the variables described in Table 1.

Table 1. Description of the variables in *Pb.hbc.Rdata* file used for modelling the crown variables.

Variable	Description
Cod_Par	Plot code
Id_Arv	Tree identifier
t	Stand age (years)
Cod_Par_Med	Code combining the measurement id and the plot code Cod_Par
Id_Trat	Treatment code (identifying the occurrence of pruning and thinning, and also the thinning weight)
Cod_Dom	Dominant tree code (1-dominant tree; 0-otherwise)
Cod_Est	Code for tree status (0-alive and healthy; 1-phytosanitary problems; 2-problems with height; 3-defect at dbh level; 4-dead; 5-forked; 6-stump of a previously thinned tree; 7-problems with height and dbh level; 8-reduced vitality; 9-other situations (see observations); 10-beat-up; 64-marked for thinning
d	Diameter-at-breast-height (cm)
h	Total tree height (m)
hbc	Height to the base of the crown (m)
Area	Plot area (m^2)
ddom	Quadratic mean diameter of the dominant trees (cm)
hdom	Dominant height (m)
S	Site index (m), base age 50 years
N	Stand density per ha (ha^{-1})
G	Stand basal area (m^2 ha^{-1})
G_di	Basal area of trees larger than tree i, $G>di$ (m^2ha^{-1})
cr	Crown ratio, (no units)
cl	crown length, (m)
h_d	Tree slenderness ratio, h/d
d_h	Inverse of tree slenderness ratio, d/h
invd	Inverse of diameter-at-breast-height, $1/d$ (cm)
invt	Inverse of age,$1/t$ (years)
invh	Inverse of total tree height, $1/h$ (m)
t2	Square of age, t^2 (years)
h2	Square of total tree height, h^2 (m^2)
N1000	$N/1000$
dg	Quadratic mean diameter, (cm)
distm	Mean distance between trees, $mdist=100/\sqrt{N}$ (m)
Fw	Relative spacing also known as Wilson's factor (Fw), $Fw=100/(hdom\sqrt{N})$
SDI	Stand density index, $SDI=N(dg/25)^{1.83926}$
invS	Inverse of dominant height, $1/hdom$
invN	Inverse of stand density per ha, $1/N$

invG	Inverse of stand basal areas, *1/G*
invhdom	Inverse of dominant height, *1/hdom*
invddom	Inverse of quadratic mean diameter of the dominant trees, *1*/*ddom*
invdg	Inverse of quadratic mean diameter, *1/dg*
S21000	$S^2/1000$
N10002	$N1000^2$
G2	G^2
hdom2	$hdom^2$
ddom21000	$ddom^2/1000$
dg2	dg^2
d_dg	Tree diameter relative to quadratic mean diameter, *d*/*dg*
d_ddom	Tree diameter relative to quadratic mean diameter of dominant trees, *d*/*ddom*
APB	Area potentially available for tree *i*, *APBi=gi/G*10000*
G_di_G	*G>di/G*
h_hdom	Total tree height relative to dominant height, *h*/*hdom*
Trial	Trial code

3. R-Scripts

3.1 SCRIPT1: Example for selecting the *f(X)* using the Logistic function

This script (*SCRIPT1_cr_SelectingTheFunction_f(X)_ExampleWithLogisticFunction(Table6).R*) is available in the GitHub repository of ForChange-CEF: https://github.com/ForChange-CEF/ModellingCrownVarMaritimePine_DataScriptExamples

It starts by loading the R data set "*Pb.hbc.Rdata*" that must be located in the same folder. Otherwise, the user must use the command *setwd()* in order to indicate the folder where the "*Pb.hbc.Rdata*" is located. In row 34 and following rows the user must test different variables in order to find the formulation of the *f(X)* function.

The model selected is fitted in rows 72-74, followed by the calculus of several statistics of model performance (as explained in the main text) and by the analysis of the regression assumptions.

The results are written in the *EXCEL* file whose name is given in row 176 (in the script provided as an example for the *cr* variable modelled with the logistic function the name is "*cr_exponential.xlsx*") and that includes two sheets, one with the parameter estimates and another with the evaluation statistics.

This script can be easily adapted to fit other functions or to model the variable *cl*.

3.2 SCRIPT2: Validating a selected Exponential function using different fitting methods

This script (*SCRIPT2_cr_ValidatingTheFitOfTheSelectedExponentialFunctionFitWithDifferentMethods-(Table7).R*) is available in the GitHub repository of ForChange-CEF: https://github.com/ForChange-CEF/ModellingCrownVarMaritimePine_DataScriptExamples

It starts by loading the R data set "*Pb.hbc.Rdata*" that must be located in the same folder. Otherwise, the user must use the commend *setwd()* in order to indicate the folder where the "*Pb.hbc.Rdata*" is located.

The user must run the script from the beginning up to row 62 (load the data and prepare the output files). Then the user must run the script from row 66 four times, changing the METHOD for each one of the four runs. The script validates the model fit with the four METHODs using the leave-one-out technique based on 28 plots that were randomly selected from each trial and saves the results of the validation statistics as well as the validation results at tree level – rows 195–401. Rows 403–428 make the plots of observed versus estimated values for the three crown variables, whereas rows 430–447 analyse the box-plots of residuals over classes of *S*, *N* and *t*.

The results are written in three *EXCEL* files. File names are assigned in rows 451–453. In the script provided as an example for the *cr* variable modelled with the exponential function the names of the three files are: "*cr_exponential_params.xlsx*" containing the parameter estimates; "*cr_exponential_evaluation.xlsx*" containing the statistics of model performance; "*cr_exponential_data_val.xlsx*" containing the output of the validation for each tree. All files contain four sheets, one for each one of the METHODS tested: "*OLS*", "*OLS&AR1*", "*MIX*" and "*MIX&AR1*".

This script can be easily adapted to fit other functions or to model the variable *cl*.

3.3 SCRIPT3: Simultaneous fitting of crown variables using seemingly unrelated regression (SUR)

This script (*SCRIPT3_SUR_cr_(Table8).R*) is available in the GitHub repository of ForChange-CEF: https://github.com/ForChange-CEF/ModellingCrownVarMaritimePine_DataScriptExamples

It starts by loading the R data set "*Pb.hbc.Rdata*" that must be located in the same folder. Otherwise the user must use the command *setwd()* in order to indicate the folder where the "*Pb.hbc.Rdata*" is located. Based on crown ratio (*cr*), the one that was selected in the paper

The script fits the system of equations to estimate the crown variables (crown ratio, crown length and height to the base of the crown) and saves the parameter estimates and respective confidence intervals in rows 53–89. Then it validates the system using the leave-one-out technique based on 28 plots that were randomly selected from each trial and saves the results of the validation statistics as well as the validation results at tree level in rows 90–165. Rows 167–192 make the plots of observed versus estimated values for the three crown variables and in rows 195–205 produce the box-plots of residuals by classes of *S*, *N* and *t*.

The results are written in three *EXCEL* files whose names are assigned in rows 209–21: "*cr_exponential_SUR_params.xlsx*" containing the parameter estimates and respective confidence intervals; "*cr_exponential_SUR_evaluation.xlsx*" containing the statistics of model performance and "*cr_exponential_SUR_data_val.xlsx*" containing the output of the validation for each tree.

4. Some examples of outputs

The outputs – *EXCEL* files – from the 3 scripts are also available in the GitHub repository of ForChange-CEF: https://github.com/ForChange-CEF/ModellingCrownVarMaritimePine_DataScriptExamples

5. References

Al Pavel MA, Barreiro S, Tomé M (2025) Compatible models to predict tree crown ratio, crown length and height to the base of the crown for maritime pine (*Pinus pinaster* Aiton.)**,** Ann For Scince**,** *(accepted with minor revisions)*