

Chestnut High Forest Stands: Biometrical Study in the North of Portugal

L.F. Nunes, M.L. Monteiro and M.S. Patrício
Department of Forestry
Polytechnic Institute of Bragança, ESAB
Campus de Santa Apolónia
Apartado 172
5301-855 Bragança
Portugal

L.S. Martins
General Direction of Forest Resources
MAPF
1068-040 Lisbon
Portugal

T. Alpuim
Department of Statistics and Operations Research
University of Lisbon
1749-016 Lisbon
Portugal

Keywords: *Castanea sativa* Mill., high forest, regression analysis, volume equations, height-diameter equations

Abstract

Growing sweet chestnut (*Castanea sativa*) in high forest stands is a lucrative investment. Wood with medium and large dimensions can attain interesting profitability levels. Furthermore, forest producers are interested of new on this species and decision criteria for the best management are needed for landowners. So, accuracy is very important when evaluating volume in standing trees. Based on regression theory, predicting equations of wood volume are developed as a function of tree variables easily measured (fast and cheap), such as diameter at breast height (*DBH*) and tree total height (*h*). Variable selection in volume equations was based on a stepwise modified procedure (linear models) and using generalized-F tests (nonlinear models). The resulting volume equation, with *DBH* and *h* as regressors, was confronted with the equation recommended by the Portuguese Forest Services. In order to evaluate wood volume at the lowest costs, the height to diameter relationship was also analyzed. Candidate regression equations were evaluated based on fitting statistics and validation statistics, complemented with the use of generalized-F tests. The biological behavior of the equations was also considered.

INTRODUCTION

Castanea sativa wood quality is well recognized among the most important broadleaved species (Monteiro et al., 1990). The distribution area of chestnut is increasing nowadays in Portugal and, over the last decade, many new high forest stands were established supported by the European Union (M.L. Monteiro, pers. commun., 2001). The main objective for these young plantations is to produce high quality wood. Logs with medium and large dimensions can attain interesting profitability levels (125 €/m³ wood, referred to trees with *DBH* higher than 30 cm).

For chestnut high forest, there is a lack of decision criteria for adopting the best management. The accurate determination of wood volume is very important. Portuguese Forest Services usually use an official equation (Fortuna, 1969) based on data from forests located at the Estrela mountain (central Portugal). As a result of the generalized application of this official equation in Portuguese chestnut stands, there is evidence on the existence of revenue losses for some wood producers caused by the underestimation of volume and, consequently, wood monetary value (Monteiro et al., 1990).

Based on data obtained from three permanent research plots established in chestnut forested areas of northern Portugal, regression analysis was used to obtain volume equations as a function of variables easily measured in standing trees. An additional objective was modelling the height to diameter (*h-DBH*) relationship in order to minimize costs of volume estimation.

MATERIAL AND METHODS

Sampling Sites and Data Collection

In 1990, two permanent research plots were established in *Castanea sativa* high forest stands: one (0.05 ha) located in Bornes and the other (0.1 ha) located in Padrela. In 1992, two more permanent plots, both with an area of 0.05 ha were established in Marão (Fig. 1). Table 1 presents some characteristics of these plots. All trees from each plot were measured for *DBH* and *h*. For volume determination, the Bitterlich tele-relascope was used to obtain upper stem diameters and the corresponding heights. Tree volume was calculated by application of Smalian formula to logs. For the stump ($h \leq 0.15$ m) and the tip (diameter < 7 cm), the cylinder and the cone formulas, were applied respectively.

In 2000, a 2nd measurement was carried out re-evaluating *DBH* and *h* in all plots except in Padrela, where removal cutting already occurred.

Volume Equations

1. Variable Selection. Establishment of volume equations was based on data from the 1st measurement. From total data, 2/3 was used for fitting equations and 1/3 used for cross-validation. On both subsets, a balanced distribution of the trees by diameter classes was taken into account. For variable selection, linear and nonlinear equations were fitted in EXCEL using ordinary least squares (OLS).

The analysis of the simple correlation matrix for the volume predicting variables most found in forestry bibliography (fd^2h , d^2f , f , d^2h , d^2 , d , dh , h^2 , and h), was the basis of the methodology used in linear equations (Table 2). Variable *h* is defined as before, *d* is the *DBH* and variable *f* represents the tree form and refers to the ratio $d_{5.30}/d$, where $d_{5.30}$ is the diameter over bark at 5.30 m high (cm). This ratio was called Girard modified coefficient (*Cg*). For variable selection, a modified version of stepwise procedure was used where, at each step, a variable can enter in the model only if its multiple correlation coefficient with the variables already in the model is not higher than a value that was fixed in 0.9 (Nunes, 2001). Stepwise analysis as a useful tool for variable selection should be carefully used when collinearity exists among regressors, which may cause that the estimates of the regression coefficients are unstable and have large standard errors. (Afifi and Clark, 1996; Chatterjee et al., 2000; Maroco, 2003). The objective of our procedure was to avoid this type of problems. When there are two variables in the model (X_1 and X_2), all other variables with multiple correlation with (X_1 , X_2) greater than 0.9 were not allowed to enter.

Nonlinear equations were first fitted in linear form using logarithmic transformation. Then, all regression theory for linear models still applies. For variable selection we used *generalized-F tests*, (Ryan, 1997), fitting first the more general model (full model) and testing against the subset of reduced models (Nunes, 2001).

At the final, the best linear and nonlinear equations were compared based on fitting statistics and validation statistics (Fig. 2). Scatterplots of observed against predicted volume were also observed. Normal Q-Q plots showed no violation to normality of residuals.

2. Equations Comparison among Sampling Sites. Best final volume equations were re-estimated using total data with exclusion of inconsistent observations and compared among sampling sites using *generalized F-tests* (Sen and Srivastava, 1997).

Height-Diameter Equations

The data from the 1st and 2nd measurements was used for modeling the *h-DBH* relationship. The shape of the height curve changes with management practices, age and site quality (Loetsch et al., 1973). Because of that, equations were established by sampling sites, using in each case, 2/3 of available data for fitting equations and 1/3 for cross-validation.

The candidate equations were selected from the bibliography. Linear equations were fitted by OLS and by nonlinear least squares (NLS) using Gauss-Newton algorithm

in STATISTICA. Comparison of the equation's performance was mainly based on fitting and validation statistics mentioned before (Fig. 2). The significance of the regression coefficients was analyzed based on *T-tests*. Biological behavior of the models was also verified. Best final equations were compared but now, for each site, among sampling dates.

RESULTS

For each site, after evaluating performance, resulting *h-DBH* best equations fitted to total data, are presented:

Padrela: $h = 0.05368706 d^{3.33909309 - 0.45454766 \ln d}$ ($R^2=0.65$; $S=1.7$; $n=65$)

Bornes: $h = -5.02725666 + 8.14601227 \ln d$ ($R^2=0.71$; $S=1.5$; $n=145$)

Marão: $h = e^{4.02779067 - 30.89711940 (1/d)}$ ($R^2=0.48$; $S=5.2$; $n=104$)

h – tree total height (m)

d – DBH (cm)

R^2 – coefficient of determination

S – standard deviation of estimates (m)

n – number of observations

As the result of the applied methodology for variable selection and model comparison, best final volume equations follow (**a** – Bornes and Padrela, **b** – Marão):

1a: $V = -0.02498571 + 0.00081997 d^2$ ($R^2=0.93$; $S=0.132$; $CV=21.5$; $n=139$)

1b: $V = -0.27309556 + 0.00095811 d^2$ ($R^2=0.91$; $S=0.197$; $CV=24.1$; $n=53$)

2a: $V = 0.00009899(d^2 h)^{0.89732256}$ ($R^2=0.94$; $S=0.128$; $CV=20.8$; $n=139$)

2b: $V = 0.00009202(d^2 h)^{0.89260958}$ ($R^2=0.94$; $S=0.163$; $CV=20.0$; $n=53$)

3a: $V = 0.00025393(d^2 f)^{1.00798702} h^{0.41153730}$ ($R^2=0.98$; $S=0.080$; $CV=12.9$; $n=139$)

3b: $V = 0.00011551(d^2 f)^{1.01203721} h^{0.53162938}$ ($R^2=0.97$; $S=0.122$; $CV=14.9$; $n=53$)

V – volume (m³)

d – DBH (cm)

h – tree total height (m)

f – tree form given by Girard modified coefficient (*Cg*)

R^2 – coefficient of determination

S – standard deviation of estimates (m³)

CV – coefficient of variation (%)

n – number of observations

DISCUSSION

Height-Diameter Relationship

Attaining values of standard deviation of estimates (*S*), best final *h-DBH* equations give satisfactory results for Padrela and Bornes, estimating height with precisions of 1.7 and 1.5 metres respectively. Corresponding coefficients of variation (*CV*) are 7.1% and 7.6%. In opposite, equation for Marão only explains 48% of total height variation, presenting high values of *S* and *CV*. Site quality seems to be different concerning the two plots established in Marão and, if the selected equation is fitted separately for these plots, values for R^2 are still low but *S* values are similar to those for Bornes and Padrela.

Comparisons to analyse evolution of *h-DBH* relationship between measurements, was only possible in Bornes and Marão. The chestnut stand in Bornes maintains, since

initial plantation, a high level of growing stock. Competition increased and growth potential is becoming limited. We can see that 10 years after the 1st measurement, the slope of the height curve is almost the same. For Marão and considering plot 1, coefficients of the equation do not differ significantly between the 1st and the 2nd measurement. In opposite, significant difference exists in the slope for plot 2. This plot shows lower growth pattern than plot 1 and the plots from Bornes and Padrela.

Volume

For volume equations as a function of *DBH* (1a and 1b), the coefficient of variation (CV) was 26.7%, being slightly high. When variable *h* is added to the model, CV decreases by 4%-units. Moreover, when variable *f* enters in the model, CV decreases by 10.1%-units in relation to the equation with only *DBH* as regressor and by 6.5%-units in relation to the equation with regressors *DBH* and *h*. Inclusion of variable tree form in the volume prediction model, led to substantial improvement in precision. Results comparing each type of equation among sampling sites, showed that there was no significant difference in regression coefficients values between Bornes and Padrela but, on the contrary, significant difference exists either between Bornes and Marão and Padrela and Marão.

Based on a 892 trees sample from a high forest chestnut stand located in Estrela Mountain (central Portugal), volume equations were established (Fortuna, 1969), using either Huber and Smalian formulas for volume evaluation. These equations have been recommended by the Forest Services. Equation resulting from Huber formula application (official) is:

$$V = -0.00246 + 0.0001684 g + 0.0000303 gh \text{ (S=0.019 m}^3\text{; CV=8.93\%)}$$

where *g* is the basal area (cm²), *V* is the volume (m³) and *h* is the total height (m).

Resulting equations **2a** (Bornes and Padrela) and **2b** (Marão) were compared with the Forest Services official equation (FS) (Fig. 3). It can be seen that the FS equation clearly underestimates volume when applied in Bornes and Padrela, mainly for volume above 0.5 m³. In Marão, both equations (2b and FS) present similar performance for the entire interval of observed volumes.

CONCLUSIONS

Usually *DBH* and *h* explain almost total variability of variable volume. Nevertheless, in the particular case of this study, the tree form (*f*) is shown to be an important regressor for predicting chestnut wood volume, causing considerable reduction in standard deviation of estimates (S) values. However, tree form is a difficult and time wasting variable to measure. So, to evaluate as accurate as possible the volume of chestnut high forest stands from Bornes, Padrela and Marão, the following equations are proposed:

$$V = 0.00009899(d^2h)^{0.89732256} \text{ (Bornes and Padrela)}$$

$$V = 0.00009202(d^2h)^{0.89260958} \text{ (Marão)}$$

To estimate tree total height in Padrela and Bornes it is acceptable to use the equations, respectively:

$$h = 0.05368706 d^{3.33909309 - 0.45454766 \ln d}$$

$$h = -5.02725666 + 8.14601227 \ln d$$

In order to reduce time and costs in volume predictions, improvement in modelling the height-diameter relationship should be considered, mainly in Marão.

Literature Cited

- Afifi, A.A. and Clark, V. 1996. Computer-Aided Multivariate Analysis. 3rd Ed., Chapman & Hall, London.
- Alves, A.M. 1988. Técnicas de Produção Florestal. Instituto Nacional de Investigação Científica. 2^a Edição, Lisboa.
- Chatterjee, S., Hadi, A.S. and Price, B. 2000. Regression Analysis by Example. 3rd Ed., John Wiley, New York.
- Fortuna, E.R.M. 1969. Tabelas de volume para o castanheiro bravo. Secretaria de Estado da Agricultura. Direcção Geral dos Serviços Florestais e Aquícolas, Série Estudos e Informação 245, Lisboa.
- Loetsch, F., Zöhrer, F. and Haller, K.E. 1973. Forest Inventory, Vol. II. Verlagsgesellschaft, München.
- Maroco, J. 2003. Análise Estatística com utilização do SPSS. Edições Sílabo, Lisboa.
- Monteiro, M.L., Sales Luís, J.F. and Oliveira, A.C. 1990. Crescimento e produção de povoamentos de *Castanea sativa* Miller em alto fuste: Primeiros resultados. Actas do II Congresso Florestal Nacional, Faculdade de Economia, Porto. p.360-367.
- Nunes, L.F. 2001. Estudo biométrico com vista à elaboração de tabelas de volume para o castanheiro. Tese de Mestrado, FCUL, Lisboa.
- Ryan, T.P. 1997. Modern Regression Analysis. Theory, Methods and Applications. Springer Verlag, New York.
- Sen, A. and Srivastava, M. 1997. Regression Analysis. Theory, Methods and Applications. 2nd Ed., Springer Texts in Statistics, Springer Verlag, New York.

Tables

Table 1. Principal characteristics of research sites.

Site	Latitude (N) Longitude (W)	Altitude (m)	Exposure	Slope (°)	Soil type	Ecological zone ^a
Bornes	41° 29' 37" N 6° 55' 12" W	780	North	18	Schist	Montano SA - I.SA
Marão	41° 14' 24" N 7° 55' 38" W	940	South	7	Granite	Montano A.SA
Padrela	41° 30' 34" N 7° 36' 54" W	830	North	29	Granite	Montano SA

^a Alves (1988)

Table 2. Simple correlation matrix for regressors with interest in volume prediction.

	f	fd^2h	d^2f	d^2h	d^2	d	dh	h^2	h	V
f	1									
fd^2h	0.3695	1								
d^2f	0.3661	0.9805	1							
d^2h	0.2553	0.9862	0.9667	1						
d^2	0.2145	0.9589	0.9815	0.9759	1					
d	0.2527	0.9296	0.9644	0.9440	0.9812	1				
dh	0.3405	0.9705	0.9465	0.9783	0.9461	0.9466	1			
h^2	0.4467	0.8235	0.7421	0.8197	0.7187	0.7367	0.9022	1		
h	0.4987	0.7966	0.7261	0.7884	0.6961	0.7219	0.8828	0.9896	1	
V	0.3747	0.9716	0.9747	0.9537	0.9482	0.9242	0.9370	0.7656	0.7472	1

V – volume (m³)

d – DBH (cm)

h – tree total height (m)

f – tree form given by Girard modified coefficient (C_g)

C_g – Girard modified coefficient ($d_{5.30}/d$) where

$d_{5.30}$ – diameter over bark at 5.30 m high (cm)

(Correlation formula):

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Figures

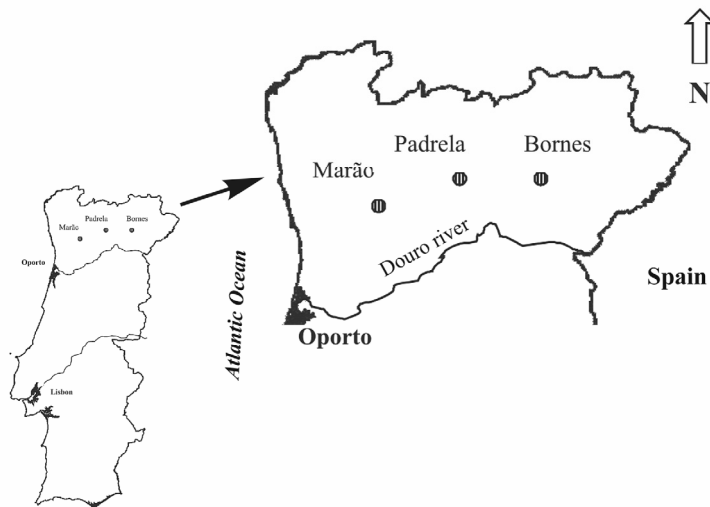


Fig. 1. Location of the research plots.

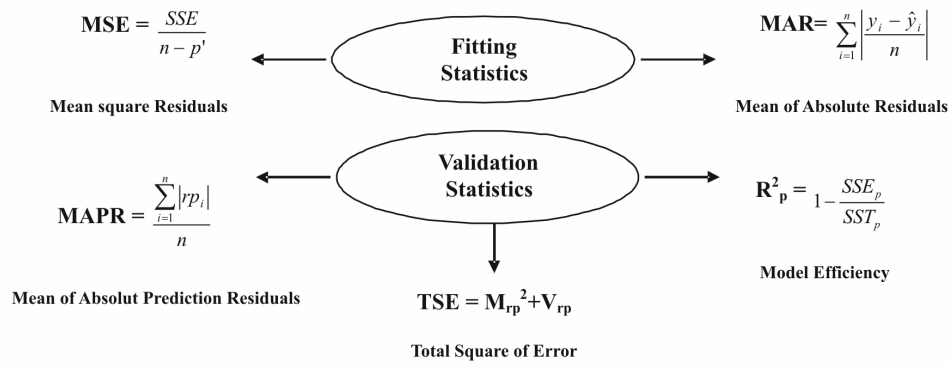
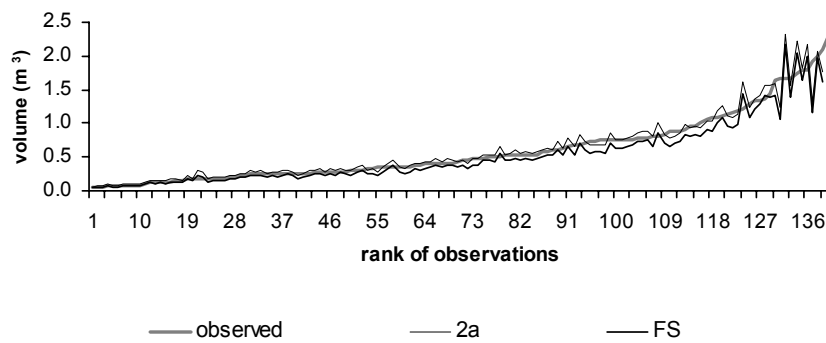


Fig. 2. Statistics used to compare equation's performance. SSE: residual sum of squares, rp_i prediction residual for observation, M_{rp}^2 : squared mean of prediction residuals, V_{rp} : variance of prediction residuals, SSE_p : residual sum of squares for validation dataset, SST_p : total sum of squares for validation dataset.

(Bornes and Padrela)



(Marão)

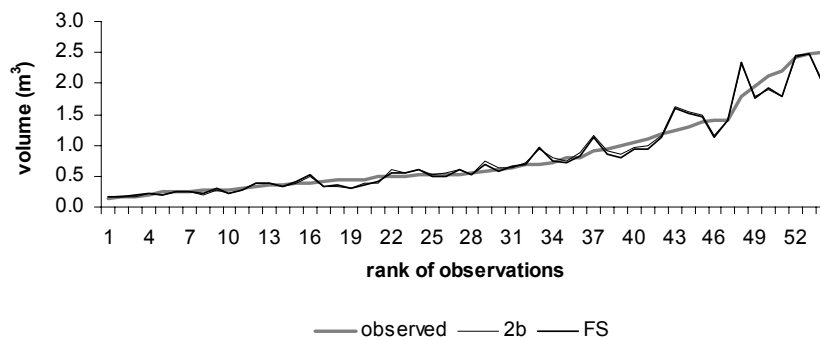


Fig. 3. Equations 2a and 2b compared with Forest Services official equation (FS).

