CHESTNUT HIGH FOREST STANDS BIOMETRICAL STUDY IN THE NORTH OF PORTUGAL

Luís Nunes[1], Maria do Loreto Monteiro[1], Maria do Sameiro Patrício[1], Teresa Alpuim[2] and Lucílio Martins[3]

[1] Agricultural College of Bragança Polytechnic, ESAB, Stª Apolónia, 5301-855 Bragança, Portugal.
[2] Department of Statistics and Operations Research, University of Lisbon, 1749-016 Lisbon, Portugal.
[3] Direction General of Forests, DGF, MADRP, 1068-040 Lisbon, Portugal.

1. INTRODUCTION

Timber production from chestnut high forest stands can be a lucrative investment for forest producers. To maximize landowners gains it is important to accurately estimate wood contents through volume determination.

Based on a tree dataset from permanent research plots installed in three pure regular high forest *Castanea sativa* stands located in the north of Portugal (Fig. 1), regression analysis was used to obtain volume equations as a function of standing tree easily measured variables and also for modeling the relationship of tree total height (*h*) to diameter at breast height (*d*).

A methodology was adopted to obtain models with the lowest possible bias in its volume estimates and, at the same time, with low correlation among regressors. Comparisons were made with equations used by the Portuguese Forest Services.

Fig. 1 – Location of permanent research plots

2. METHODS

SAMPLING SITES AND DATA COLLECTION

In 1990, Monteiro established two semi-permanent research plots in *Castanea sativa* high forest stands located in Bornes and Padrela. In 1992, two more permanent plots were established in Marão (Fig. 1). Dendrometrical parameters *d* and *h* were measured in all trees as well as upper stem diameters and the corresponding heights.

In 2000, a 2nd measurement on *d* and *h* was carried out in the research plots.

VOLUME EQUATIONS

Data from 1st measurement was used for volume equations establishment. Total dataset was split into two subsets, one with 2/3 of total dataset (estimation dataset) and other with 1/3 of total dataset (validation dataset). In both subsets, a balanced distribution of trees by diameter classes was taken into account, so as to maintain representativeness (Table 1).

Table 1 - Summary statistics for estimation and validation datasets

dataset	n.º of trees		*d* (cm)	*h* (m)	*f*	*V* (m³)
estimation	126	minimum	9.0	10.9	0.429	0.038
		maximum	52.4	33.7	0.999	2.506
		mean	28.6	21.1	0.788	0.711
validation	70	minimum	12.8	11.6	0.530	0.085
		maximum	56.0	30.5	0.956	2.428
		mean	26.9	20.7	0.801	0.583

Note: *f*: tree form; *V*: total volume; (a): *f* values refers to the ratio $d_{h/20}/d$, where $d_{h/20}$ is the diameter over bark at 5.30m high (cm). At this ratio we called Girard modified coefficient (*Cg*).

For model selection, linear and nonlinear models were worked apart. All the nonlinear models were fitted in linear form using the logarithmic transformation. So, all regression theory for linear models still applies.

For linear models we used a modified version of the stepwise procedure where, at each step, a variable can enter in the model only if its multiple correlation coefficient with the variables already included in the model is not higher than a certain value, which, in our study, was fixed in 0.9 (Nunes, 2001). The simple correlation matrix for the most common volume predictor variables found in forestry bibliography (fd^2h, d^2f, f, d^2h, d^2, d, dh, h^2, and h) was in the base of the described methodology.

Selection procedure for nonlinear models was based in *generalized F- tests* (Nunes, 2001).

Selection of the best final model was based in statistics that evaluate quality of estimation and predictive ability (Fig. 2). Best final models were re-estimated for total dataset and compared among sampling sites by *generalized F-tests*.

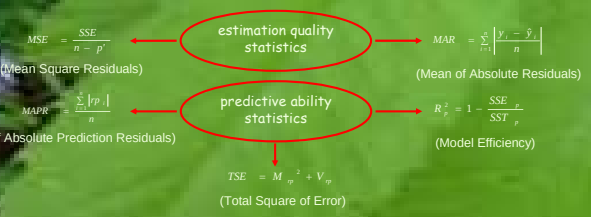

Fig. 2 – Statistics used to evaluate model's quality of estimation and predictive ability.

Note: SSE: residual sum of squares; *rpi* prediction residual for observation *i* of validation dataset; 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.

HYPSOMETRIC EQUATIONS

For modeling the relationship of the *h* and *d* data from de 1st and 2nd measurements were used. Equations were established by sampling sites and total datasets were also split (Table 2).

Table 2 – Summary statistics for height to diameter curves data

Site	n.º of trees estim./valid.		estimation *d* (cm)	estimation *h* (m)	validation *d* (cm)	validation *h* (m)
Bornes	97/48	minimum	11.9	14.9	10.3	13.4
		maximum	37.1	24.9	40.7	24.8
		mean	24.2	21.2	21.7	19.7
Marão	72/32	minimum	22.8	10.9	19.7	10.9
		maximum	52.4	33.0	54.5	33.7
		mean	34.1	22.1	33.2	21.1
Padrela	48/17	minimum	13.6	12.0	17.8	20.6
		maximum	52.0	26.7	56.0	25.9
		mean	31.9	23.1	35.0	23.6

From bibliography we selected a set of candidate models. Selection procedure was mainly based on the model's performance for quality estimation and predictive ability statistics mentioned before and adding the statistic R^2_{aj} (Adjusted R-Square) to the first group. Attention was paid to T-tests on the parameters and biological behavior of the models. Best final models were also compared but, for each site, among sampling dates.

3. RESULTS AND DISCUSSION

After selection procedure and comparing for different sampling dates, final height to diameter models for Bornes and Marão respectively follow:

Model: $h = b_0 + b_1 \ln d$ (Henricksen)

Date	Parameters	R^2	S (m)	CV %	n
1990	b_0 = -4.35761037	0.719	1.43	7.5	76
	b_1 = 7.76787290				
2000	b_0 = -1.87804031	0.640	1.49	7.0	69
	b_1 = 7.29916839				

Bornes

Model: $h = e^{(b0+b1(1/d))}$ (Michailov)

Plot	Date	Parameters	R^2	S (m)	CV %	n
1	1992 /2000	b_0 = 3.61286499	0.416	2.07	7.3	53
		b_1 = -9.71211856				
2	1992	b_0 = 3.31905843	0.493	1.81	13.0	26
		b_1 = -18.4687604				
	2000	b_0 = 3.30485143	0.523	1.51	9.4	25
		b_1 = -15.9927026				

Marão

Values for R^2 are generally low, especially in Marão. Other models will be tested in order to improve quality of estimation for height to diameter curve.

After selection procedure and comparing for different sampling sites we obtained the final volume models:

		R^2	S (m³)	CV%	n
$V = -0.02498571 + 0.00081997d^2$	(Bornes and Padrela)	0.934	0.132	21.5	139
$V = -0.27309556 + 0.00095811d^2$	(Marão)	0.910	0.197	24.1	53
$V = 0.00009899 (d^2h)^{0.89732256}$	(Bornes and Padrela)	0.939	0.128	20.8	139
$V = 0.00009202 (d^2h)^{0.89260958}$	(Marão)	0.938	0.163	20.0	53
$V = 0.00025393 (d^2f)^{1.00798702} h^{0.41153730}$	(Bornes and Padrela)	0.977	0.080	12.9	139
$V = 0.00011551 (d^2f)^{1.01203721} h^{0.53162938}$	(Marão)	0.966	0.122	14.9	53

R^2 - determination coefficient, S (m³) – standard deviation of model estimates, CV(%) – coefficient of variation, n – observations.

Compared to one-entry models, CV decreased by 6.5%-units when *h* is added. Moreover, CV decreased by 10.1%-units with inclusion of variable *f* in the volume prediction model, leading to substantial improvement in precision.

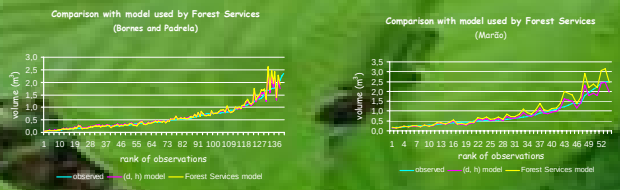

Fig. 3 – Comparison with Forest Services model .

For Bornes and Padrela the two models have similar volume estimation pattern for values less than 1m³ (Fig. 3). Up to 1 m³, both equations tend to produce estimates with increasing S values. Bias is more evident for values above 1.5 m³, mostly in Forest Services model. For Marão, Forest Services model clearly overestimates volume, namely from 0.5 m³ on.

4. REFERENCES

Nunes, L.F. 2001. Estudo biométrico com vista à elaboração de tabelas de volume para o castanheiro. Tese de Mestrado. FCUL-DEIO, Lisboa. 196 pp.