IPEF Série Técnica, Piracicaba, 9(26): 1-36, mar.1993 ISSN 100 - 8137

A DETERMINAÇÃO DE EQUAÇÕES VOLUMÉTRICAS NA ENGENHARIA FLORESTAL

Frederico Pimentel Gomes
Carlos Henrique Garcia

INSTITUTO DE PESQUISAS E ESTUDOS FLORESTAIS
PRODUZINDO FLORESTAS COM CIÊNCIA

em convênio com

UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ E QUEIROZ"
Departamento de Ciências Florestais

Série Técnica IPEF ***(ISSN 100-8137)*** é uma publicação trimestral do **IPEF – Instituto de Pesquisas e Estudos Florestais**. Publlica contribuições originais, que se enquadram como anais de encontros ou monografias, com o objetivo de atualizar o conhecimento sobre temas florestais de grande interesse prático. (tiragem de 300 exemplares)

Endereço
IPEF/CTI – Central Técnica de Informações
Av. Pádua Dias, 11 - Caixa Postal 530
13400-970 – Piracicaba, SP – Brasil
FONE (0194) 334124
FAX (0194) 33 6081
TELEX 19 7881 IPEF BR

SUMÁRIO

A DETERMINAÇÃO DE EQUAÇÕES VOLUMÉTRICAS NA ENGENHARIA FLORESTAL

[*]Frederico Pimentel Gomes
[**]Carlos Henrique Garcia

1. INTRODUÇÃO

As equações para determinação de volume sólido de essências florestais são de uso geral e indispensável na Silvicultura. Como todas elas são empíricas, faz-se necessário ajustá-las com freqüência, para daptá-las a diferentes espécies, idades, espaçamentos e regiões. Estudar-se-á principalmente as equações clássicas seguintes (VEIGA, 1984):

1. Equação de variável combinada de Spurr:

$$V = a + b\ D^2H$$

2. Equações australianas, de Stoate:

$$V - a + b\ D^2H + c\ D^2 + f\ H.$$

3. Equação de Schumacher-Hall, na forma logarítmica:

$$\text{Log } V = a + b \text{ Log } D + c \text{ Log } H,$$

ou na forma original:

$$V = A\ D^b\ H^c$$

Nestas equações, **V** é o volume sólido (com ou sem casca), **D** é o DAP e **H** é a altura.

Embora só essas equações sejam discutidas especificamente, os métodos expostos podem ser aplicados a muitas outras, com as ligeiras modificações necessárias.

2. A ESTIMATIVA DOS PARÂMETROS

Equação de Spurr

É linear nos parâmetros. Se tomar $X = D^2H$, transformar-se-á em

$$V = a + bX$$

Os parâmetros (ou coeficientes) **a** e **b** podem ser estimados facilmente pelos métodos clássicos de regressão (PIMENTEL-GOMES, 1990, capítulo 12, DRAPER &

[*] Consultor do IPEF – Caixa Postal 530 – 13400-970, Piracicaba, SP
[**] IPEF – Caixa Postal 530 – 13400-970, Piracicaba, SP

SMITH, 1981) ou por programas apropriados de computador, tias como o REG do SAS, ou o REGRESEQ do SAEG.

Equação de Stoate

Com $X_1 = D^2H$, $X_2 = D^2$, $X_3 = H$, ela se transforma em:

$$V = a + bX_1 + cX_2 + fX_3,$$

Linear nos parâmetros, que se podem estimar pelos métodos indicados para a equação de Spurr.

Equação logarítmica de Schumacher-Hall

Com $U = \text{Log } V$, $X_1 = \text{Log } D$, $X_2 = \text{Log } H$, ela se transforma em

$$U = a + bX_1 + cX_2,$$

igualmente linear nos parâmetros, que se podem estimar como nos dois casos anteriores.

Equação de Schumacher-Hall original

Esta equação, dada pela expressão

$$V = A\ D^b\ H^c,$$

Tem os parâmetros **b** e **c** como expoentes e, pois, não é linear nos parâmetros. É usual passá-la à forma logarítmica

$$U = \text{Log } V = \text{Log } A + b \text{ Log } D + c \text{ Log } H,$$

$$= a + b \text{ Log } D + c\text{Log } H,$$

onde $a = \text{Log } A$. Nesta forma, ela é linear nos parâmetros, e eles poderão ser estimados pelo programa REG do SAS, ou pelos seus equivalentes do SANEST ou do SAEG. Obtidas as estimativas $\hat{a}, \hat{b}, \hat{c}$, pode-se voltar à forma original e escrever:

$$V = 10^{\hat{a}}\ D^{\hat{b}}\ H^{\hat{c}}$$

Mas a equação assim obtida subestima o volume V (THÖNI (1967)) e não corresponde exatamente à que se obteria a partir da equação original $V = A\ D^b\ H^c$. O melhor, pois, é tomar $10^{\hat{a}}, \hat{b}, \hat{c}$ como valores iniciais de A, b, c e através do programa NLIN do SAS, ou o REGRENL do SAEG obter novas estimativas A_1, b_1, c_1 para esses parâmetros.

Equação de Spurr – Exemplo

A equação em pauta é $V = a + bX$, com $X = D^2H$. Usar-se-á como exemplo os dados da Tabela 1. O programa REG do SAS dá a seguinte análise da variância. (Listagem nº 1).

C. variação	**G.L.**	**S.Q.**	**Q.M.**	**F**	**Probab.**
Regressão Linear	1	0,39955	0,39955	126,79	0,0001
Resíduo	48	0,15125	0,00315		
Total	49	0,55080			

A equação obtida é:

$$V = 0{,}0540 + 0{,}2026\ D^2H,$$

com $R^2 = 72{,}54\%$. O coeficiente de determinação ajustado $\left(R_a^2\right)$ é dado pela fórmula

$$R_a^2 = 1 - \frac{N-1}{N-P-1}(1 - R^2),$$

onde N é o número de árvores da amostra e p é o numero de coeficiente de variáveis na equação de regressão.

Obtem-se pois:

$$R_a^2 = 1 - \frac{50-1}{50-1-1}(1 - 0{,}7254) = 0{,}7197 = 71{,}97\%.$$

Quando é grande o núemero de árvores na amostra, é pequena a diferença entre R^2 e R_a^2: tal é o caso usual no estudo de equações da Silvimetria.

TABELA 1 – Dados dendrométricos de 50 árvores de ***Eucalyptus saligna*** Smith de 10 anos de idade.

Arv.	**D(m)**	**H(m)**	**V(m^3)**
1	0,172	12,0	0,116
2	0,190	14,0	0,130
3	0,270	17,0	0,249
4	0,150	11,0	0,047
5	0,122	9,0	0,041
6	0,187	16,0	0,131
7	0,245	14,0	0,203
8	0,237	18,0	0,270
9	0,160	9,5	0,071
10	0,250	14,0	0,202
11	0,235	14,0	0,214
12	0,200	12,0	0,129
13	0,177	13,0	0,144
14	0,283	16,0	0,305
15	0,275	18,0	0,357
16	0,205	15,0	0,160
17	0,200	12,0	0,254
18	0,204	12,0	0,146
19	0,215	15,0	0,211
20	0,270	14,0	0,199
21	0,160	12,0	0,088
22	0,215	14,0	0,148
23	0,205	16,0	0,173
24	0,203	14,0	0,212
25	0,215	14,0	0,156
26	0,260	18,0	0,438
27	0,310	16,0	0,255
28	0,185	12,0	0,091
29	0,215	16,0	0,177
30	0,220	14,0	0,239
31	0,150	7,2	0,043
32	0,250	16,0	0,261
33	0,315	18,0	0,397
34	0,210	18,0	0,209
35	0,273	18,0	0,328
36	0,170	16,0	0,123
37	0,202	17,0	0,201
38	0,254	19,0	0,352
39	0,180	14,0	0,133
40	0,190	14,0	0,156
41	0,223	14,0	0,200
42	0,174	12,0	0,135
43	0,364	18,0	0,572
44	0,245	16,0	0,226
45	0,260	14,0	0,301
46	0,190	14,0	0,162
47	0,247	16,0	0,234
48	0,270	18,0	0,365
49	0,130	8,6	0,344
50	0,224	16,0	0,220

Nesta análise de variância, o Resíduo corresponde na verdade a Desvios da Regressão, com QMResíduo = 0,0315 e $S = \sqrt{0{,}0315} = 0{,}0561\,m^3$. Como a média geral é m = 0,2104 m^3, o coeficiente de variação seria:

$$CV = \frac{0{,}0561 \text{ x } 100}{0{,}2104} = 26{,}7\%.$$

A análise poderia ser feita também pelo programa GLM do SAS, com resultados equivalentes, neste caso, em que há uma só variável independente X = D^2H.

A análise poderia ser feita, ainda, pelo programa REGREGN do SAEG. Este programa obtém estimativas dos parâmetros e dos respectivos erros padrões, mas não fornece uma análise de variância. Aplica o teste **t** a cada parâmetro estimado, com exceção do parâmetro **a,** neste caso. Obtêm-se, assim, os resultados seguintes: (Listagem nº 2).

$\hat{a} = 0{,}0540$;

$\hat{b} = 0{,}2026$; $S(b) = 0{,}0180$; $t = 11{,}26$;

$R^2 = 72{,}55\%$; $R^2_a = 71{,}98\%$.

Equação de Stoate – Exemplo

É a seguinte:

$$V = a + b\,D^2H + c\,D^2 + f\,H.$$

C. variação	G.L.	S.Q.	Q.M.	F	Probab.
Regressão Linear	3	0,41212	0,13737	45,57	0,0001
Resíduo	46	0,13868	0,00301		
Total	49	0,55080			

O valor de R^2 é:

$$R^2 = \frac{SQRegressão}{SQTotal} = \frac{0{,}41212}{0{,}55080} = 0{,}7482 = 74{,}82\%.$$

Embora com 3 variáveis ($X_1 = D^2H$, $X_2 = D^2$, $X_3 = H$), em vez de uma ($x = D^2H$), como a equação de Spurr, o valor de R^2 é pouco maior (74,82%, em vez de 72,54%).

A equação obtida é:

$$V = 0{,}2061 + 0{,}4945\,D^2H - 5{,}0305\,D^2 - 0{,}0085\,H$$
$$= 0{,}2061 + (0{,}4945\,H - 5{,}0305)\,D^2 - 0{,}0085\,H.$$

Esta equação é muito estranha, pois para H < 10,17 m ela é decrescente, em relação à variável D. Por exemplo:

$$V(D^2; H{=}8) = 0{,}1381 - 1{,}0745\ D^2.$$

Para $D^2 < 0{,}0172$ ($D < 0{,}131$) ela é também decrescente em relação a H. Por exemplo:

$$V(D^2 = 0{,}0150; H) = 0{,}1306 - 0{,}00108\ H$$

Conclui-se, pois, ser perigoso o uso dessa equação, sem restrições rigorosas.

O programa REG do SAS testa os coeficientes separadamente, pelo teste **t**. Por ai se verifica que o coeficiente de D^2H (b) é significativo ao nível de 1% (P = 0,0013), o de D^2 (c) é significativo ao nível de 5% (P – 0,0472), e o de H (f) não é significativo (P = 0,1651). Poder-se-ia, pois, recalcular a equação de regressão com exclusão de seu último termo, isto é, com a forma $V = a + b\ D^2H + c\ D^2$.

Por outro lado, obtem-se $S = \sqrt{QMResíduo} = \sqrt{0{,}00301} = 0{,}0549$, com coeficiente de variação CV = 26,10%.

A análise pelo GLM do SAS dá mais detalhes, pois, além da análise da variância apresentada, igual à do programa REG, indica a contribuição de cada variável na SQRegressão. Isto é feito de duas maneiras: pela análise de tipo I e pela de tipo II. (Listagem nº 4).

A análise de tipo I dá os resultados seguintes.

C. Variação	**G.L.**	**Q.M.**	**F**	**Probab.**
D^2H	1	0,399548	132,53	0,0001
D^2	1	0,006578	2,18	0,1465
H	1	0,005997	1,99	0,1651

Estes valores de F são calculados em relação ao QMResíduo = 0,00301.

Na análise de tipo I, acima exposta, a Soma de Quadrados relativa à primeira variável mencionada (D^2H, no caso presente) não é ajustada. A Soma de Quadrados referente à segunda variável (D^2) é ajustada em relação à primeira (D^2H). Já a Soma de Quadrados relativa à terceira variável (H) é ajustada em relação às duas variáveis anteriores (D^2H e D^2). Em resumo, no tipo I a Soma de quadrados referente a cada variável é ajustada em relação a todas as variáveis anteriores. É evidente, pois, que a ordem em que se colocam as variáveis acarreta mudanças, que podem ser drásticas, nessa análise da variância de tipo I, que, em geral, se considera a mais indicada para estudos de regressão. É conveniente, pois, colocar as variáveis em ordem decrescente de importância. No caso discutido parece evidente que se deva começar por D^2H, mas a ordem das variáveis restantes poderia ser D^2, H ou H, D^2. Se, por exemplo, tivesse adotado a ordem H, D^2, D^2H, os resultados seriam os seguintes, para o tipo I de análise de variância.

C. Variação	**G.L.**	**Q.M.**	**F**	**Probab.**
H	1	0,237684	78,84	0,0001
D^2	1	0,139204	46,17	0,0001
D^2H	1	0,035235	11,69	0,0013

Já a análise do tipo III ajustada a Soma de Quadrados de qualquer das variáveis em relação a todas as demais. É evidente, pois, que no tipo III, a ordem de consideração das variáveis é indiferente: os resultados são os seguintes, em qualquer caso.

C. Variação	G.L.	Q.M.	F	Probab.
D^2H	1	0,035235	11,69	0,0013
D^2	1	0,012535	4,16	0,0472
H	1	0,005997	1,99	0,1651

Esta análise de tipo III confirma que a variável mais importante é D^2H, seguida por D^2 e ficando H em último lugar.

O programa REGRESEQ do SAEG faz, por método um pouco diferente, análise de regressão semelhante à do SAS, pelo tipo I. Compeende-se, pois, que também neste caso a ordem de entrada das variáveis é importante. Adotada a ordem D^2H, D^2, H tem-se primeiro a estimativa do parâmetro **b**, como se houvesse apenas esta variável, e lhe aplica o teste **t**. A estimativa é $\hat{b} = 0,2026$, e o valor de **t** é t = 11,26, com P = 0,0000. fornece também a Soma de Quadrados correspondente: $SQ(D^2H) = 0,399598$ e os coeficientes de determinação $R^2 = 72,55\%$, $R_a^2 = 71,98\%$. (Listagem nº 5).

A seguir, considera as variáveis D^2H e D^2, estima $\hat{b} = 0,3439$ (coeficiente de D^2H) e $\hat{c} = -2,8202$ (coeficiente de D^2) e aplica o teste **t** a estas estimativas, obtendo para $\hat{b}$, t = 3,56 e, para $\hat{c}$, t = -1,49. Fornece também a Soma de Quadrados relativa a D^2, ajustada em relação a D^2H (SQ = 0,006798), assim com os novos coeficientes de determinação: $R^2 = 73,78\%$, $R_a^2 = 72,67\%$.

Finalmente, acrescenta o programa REGRESEQ do SAEG a variável H, dá as novas estimativas de b, de c e de f, com respectivos testes **t** e, ainda, fornece o valor SQ = 0,005994 relativo à variável H, depois de ajustada em relação a D^2H e D^2. A equação finalmente obtida é:

$$V = 0,2050 + 0,4951\ D^2H - 5,039\ D^2 - 0,0085\ H,$$

sendo significativos ao nível de 5% apenas os coeficientes de D^2H e de D^2.

Notam-se pequenas discrepâncias entre os resultados obtidos pelo SAEG e os fornecidos pelo SAS. Tais discrepâncias devem ser conseqüência de problemas de aproximação numérica.

A Equação de Schumacher-Hall na Forma Logarítmica

A equação é:

$$\text{Log } V = a + b \text{ Log } D + c \text{ Log } H.$$

Com os dados da Tabela 1, o programa REG do SAS dá os resultados seguintes. (Listagem nº 6).

C. variação	G.L.	S.Q.	Q.M.	F	Probab.
Regressão	2	1,93956	0,96978	46,79	0,0001
Resíduo	47	0,97412	0,02073		
Total	49	2.91368			

A equação obtida é:

$$\text{Log V} = -0{,}5782 + 1{,}4808 \text{ Log D} + 0{,}7234 \text{ Log H},$$

com $R^2 = 66{,}57\%$ e $R_a^2 = 65{,}14\%$. O desvio padrão é $S = \sqrt{0{,}02073} = 0{,}144$.

A média de Log V é negativa: -0,736, e o mesmo ocorre, pois, com o CV = -19,5%.

A Equação Original de Schumacher-Hall

A equação é:

(1) $V = A\ D^b H^c$.

Aplicado o logarítmo decimal, obtém-se:

Log V = a + b Log D + c Log H,

onde a = Log A. Ajustada aos dados da Tabela 1, tem-se:

(2) Log V = -0,5782 + 1,4808 Log D + 0,7234 Log H,

de onde se conclui que a equação original seria:

(3) $V = 0{,}2641\ D^{1{,}4808}\ H^{0{,}7234}$

Mas o modo correto de ajustamento da equação original (1) não é este. O certo é aplicar o programa NLIN do SAS ou o programa REGREGN do SAEG ou programas análogos de outros aplicativos, uma vez que a equação (1) não é linear nos parâmetros. A equação(3), obtida através da transformação logarítmica apenas nos fornece os valores iniciais para o programa escolhido. No caso do programa NLIN do SAS, tomaram-se os valores iniciais:

A: 0,2 a 0,3 by 0,005
B: 1,30 a 1,50 by 0,10
C: 0,50 a 0,90 by 0,10

O método de Gauss-Newton do programa NLIN deu a equação

$$V = 0{,}1256\ D^{1{,}460}\ H^{0{,}9945}$$

e a análise da variância seguinte. (Listagem nº 7).

C. Variação	G.L.	S.Q.	Q.M.
Regressão	3	2,60392	0,86797
Resíduo	47	0,15945	0,00339
Total (não corrigido)	50	2,76337	

No entanto, num caso como esse, em que o parâmetro A se ajusta à média dos valores do volume (V), é preferível obter uma Soma de Quadrados da Regressão com subtração da correção $C = (\sum V)^2/50 = 2{,}21257$. Obtem-se então a seguinte análise da variância.

C. Variação	G.L.	S.Q.	Q.M.	F
Regressão (corrigida)	2	0,39135	0,19568	$57{,}72^{**}$
Resíduo	47	0,15945	0,00339	
Total (corrigido)	49	0,55080		

Neste caso tem-se:

$$R^2 = \frac{0{,}39135}{0{,}55080} = 71{,}05\%$$

A estimação dos parâmetros A, b, c pode ser feita também pelo programa REGREGN do SAEG. Usaram-se os valores iniciais:

(A = 0.1, 0.3, 0.264)
(B = 1.3, 1.6, 1.480)
(C = 0.5, 1.1, 0.723)

Por exemplo, para A, o intervalo proposto é [0.1; 0.3] com média 0.264.
As estimativas obtidas foram as seguintes: (Listagem nº 8).

$$\hat{A} = 0{,}12349\,;\ \hat{b} = 1{,}4572\,;\ \hat{c} = 1{,}0002$$

Resolveu-se tomar por base estas estimativas, como novos valores iniciais, e usar mais uma vez o programa REGREGN. As novas estimativas, mais refinadas, foram as seguintes: (Listagem nº 9).

$$\hat{A} = 0{,}12586\,;\ \hat{b} = 1{,}4598\,;\ \hat{c} = 0{,}99464$$

Elas são praticamente iguais às dadas pelo aplicativo SAS. Mas, infelizmente, o SAEG não dá a SQResíduo, e, assim, não permite o cálculo de uma análise de variânica. Mas dá o coeficiente de determinação $R^2 = 71{,}05\%$ e um gráfico dos resíduos, que é importante.

É importante salientar que a transformação logarítmica da variável depende (V) acarreta sempre subestimação para ela (THÖNI, 1967).

3. CRITÉRIOS PARA JULGAMENTO DAS EQUAÇÕES

Qual o critério para comparar equações de regressão, de modo a indicar qual a mais conveniente? Podem-se indicar os seguintes critérios em ordem de importância:

1. as propriedades matemáticas das funções;
2. o coeficiente de determinação R^2, às vezes substituído pelo coeficiente de determinação ajustado R^2;
3. o QMDesvios da Regressão, ou o desvio padrão (S) respectivo; e
4. a distribuição dos resíduos.

Discutir-se-á rapidamente esses critérios.

As Propriedades Matemáticas das Funções

Nos fenômenos biológicos, as equações correspondentes, embora empíricas, têm, geralmente, certas propriedades matemáticas conhecidas. Por exemplo, as equações de volume devem ser funções crescentes do DAP e da altura. Assim, não seria aceitável uma equação.

$$V = A + B\,X + C\,X^2,$$

com $X = D^2H$ e $C < 0$, que tivesse um máximo dentro do intervalo dos valores de X observados.

PIMENTEL-GOMES (1990, pp.229-235) dá um exemplo interessante em que, com os dados de um experimento de adubação fosfatada de milho, cujos níveis de P_2O_5 variaram de zero a 100 kg/ha. A análise estatística indicou a conveniência de uma equação de 3º grau.

$$Y = 4{,}712 + 0{,}276\,X - 0{,}00483\,X^2 + 0{,}0000256\,X^3,$$

onde Y é a produtividade do milho e X é o nível de P_2O_5. Mas, do ponto de vista agronômico, esta equação não convém, pois tem um ponto de máximo para X = 43,9 kg/ha de P_2O_5, o que seria razoável, mas tem também um ponto de mínimo para X = 81,9 kg/ha de P_2O_5. As médias de produtividade para os níveis de P_2O_5 eram as seguintes:

Nível de P_2O_5(kg/ha)	0	25	50	75	100
Produtividade Média	4,65	9,26	9,30	9,35	9,64

Estas médias demonstram que a produtividade praticamente se estabilizou para níveis de P_2O_5 de 25 kg/ha para cima. Tal comportamento, que não pode ser adequadamente representado por uma regressão polinomial, é que leva à excentricidade exibida pelo polinômio de 3º grau. Na verdade, a lei de Mitscherlich (PIMENTEL-GOMES, 1990) seria mais adequada para esse caso.

Analogamente, o estudo matemático de uma equação proposta para a estimativa do volume sólido pode identificar casos em que não seja conveniente.

Foi isso que aconteceu na equação de Stoate antes obtida (pág. 4). Com efeito, essa equação, que é

$$V = 0{,}2061 + 0{,}4945\ D^2H - 5{,}0305\ D^2 - 0{,}0085\ H$$

é decrescente em relação a D, quando se tem H < 10,17 m, e também decrescente em relação a H, para D < 0,131.

O Coeficiente de determinação (R^2)

Considere-se o experimento de adubação fosfatada de milho, acima referido (PIMENTEL-GOMES, 1990, pp.229-235), instalado em 4 blocos casualizados, com 5 níveis de P_2O_5 (zero, 25, 50, 75 e 100 kg/ha). A análise da variância dos dados de produtividade, em kg/parcela, é a seguinte:

C. de Variação	G.L.	S.Q.	Q.M.	F
Tratamentos	4	72,22	18.055	19,84**
Blocos	3	2,73	. . .	
Resíduo	12	10,92	0,910	
Total	19	85,87		

Os programas ANOVA e GLM, que realizam essa análise, dão também o valor de R^2, mas este R^2 nada tem a ver com a regressão que se pode obter, para estimar o efeito do nutriente. Na verdade, esse R^2, que o programa refere como R^2 do **modelo**, é definido como:

$$R^2 = \frac{\text{SQTotal - SQResíduo}}{\text{SQTotal}} = \frac{85{,}87 - 10{,}92}{85{,}87} = \frac{74{,}95}{85{,}87} = 0{,}873 = 87{,}3\%.$$

Se admitida regressão polinomial, pode-se separar os 4 G.L. de Tratamentos em 4 componentes: Linear (ou de 1º grau) Quadrático (ou de 2º grau), Cúbica (ou de 3º grau) e de 4º grau, com os resultados seguintes:

C. de Variação	G.L.	S.Q.	Q.M.	F
Regressão Linear	1	40,64	40,64	44,66**
Regressão Quadrática	1	21,28	21,28	23,38**
Regressão Cúbica	1	9,23	9,23	10,14*
Regressão de 4º grau	1	1,07	1,07	1,18
Tratamentos	4	72,22	. . .	
Resíduo	12	10,92	0,910	

Se considerar uma regressão de 3º grau, a equação será:

$$Y = 4{,}712 + 0{,}276\ X - 0{,}00483\ X^2 + 0{,}0000256\ X^3.$$

A Soma de Quadrados relativa a esta equação de regressão será:

$$\text{SQRegressão} = 40{,}64 + 21{,}28 + 9{,}23 = 71{,}15\ ,$$

e o valor de R^2 correspondente é:

$$R^2 = \frac{\text{SQRegressão}}{\text{SQTratamentos}} = \frac{71,15}{72,22} = 0,9852 = 98,52\%.$$

Mas se considerar somente uma equação de 2º grau, o resultado seria:

$$Y = 7,912 + 0,8448\ X - 0,0009856\ X^2.$$

Note-se que mudaram tanto o coeficiente de X como o de X^2, ao se eliminar o termo de 3º grau na equação de regressão.

A nova Soma de Quadrados da Regressão é:

$$\text{SQRegressão} = 40,64 + 21,28 = 61,92,$$

o que daria um novo

$$R^2 = \frac{61,92}{72,22} = 0,8574 = 85,74\%.$$

Ao passar da equação de 2º grau, $Y = a + bX + cX^2$, para a de 3º grau (com um parâmetro a mais),

$$Y = a + bX + cX^2 + dX^3,$$

o valor de R^2 aumenta necessariamente. Mas esses valores não são diretamente comparáveis, pois se referem a equações com número diferente de parâmetros: três na equação de 1º grau (a, b, c) e quatro na de 3º grau (a, b, c, d). Para torná-los comparáveis, é usual calcular o **coeficiente de determinação ajustado** R_a^2, dado pela fórmula (PIMENTEL-GOMES, 1990):

$$R_a^2 = 1 - \frac{n-1}{n-p-1}(1 - R^2),$$

onde **n** = 5 é o número de níveis da variável independente (X) e **p** é o número de coeficientes de variáveis na equação de regressão.

No caso da equação de 2º grau tem-se p = 2 e fica:

$$R_a^2 = 1 - \frac{5-1}{5-2-1}(1 - 0,8574) = 0,7148 = 71,48\%.$$

Para a equação de 3º grau tem-se p = 3 e fica:

$$R_a^2 = 1 - \frac{5-1}{5-3-1}(1 - 0,9852) = 0,9408 = 94,08\%.$$

Neste caso, a equação de 3º grau tem R^2 maior do que a de 2º, como acontecia com o R^2, mas em certas oportunidades a situação se inverte.

Uma dificuldade que ocorre com o coeficiente de determinação ajustado R_a^2 é que pode assumir valor negativo, o que é absurdo. Isto ocorre quando se tem $R^2 < p/(n-1)$. Por exemplo, no caso da equação de 2º grau acima referida, se tiver $R^2 = 0,40$, fica:

$$R_a^2 = 1 - \frac{5-1}{5-2-1}(1-0,40) = -0,20 = -20\%.$$

Nos casos de análises de variância em que se tenha:

$$\text{SQTotal} = \text{SQRegressão} + \text{SQResíduo},$$

O valor de R^2 é:

$$R^2 = \frac{\text{SQRegressão}}{\text{SQTotal}}$$

É exatamente isso que ocorre no caso da determinação de equações para estimativa de volume, em Silvicultura.

O Quadro Médio dos Desvios de Regressão

Considere-se os Quadrados Médios dos Desvios da Regressão para as equações seguintes, já estudadas:

Spurr $\quad V = a + b\,D^2H$

Stoate $\quad V = a + b\,D^2H + c\,D^2 + f\,H$

Schumacher-Hall $\quad V = A\,D^b\,H^c$

Os resultados são os seguintes:

Equação	Nº de G.L. do Resíduo	QMResíduo	Desvio Padrão
Spurr	48	0,00315	0,0561
Stoate	46	0,00301	0,0549
Schumacher-Hall	47	0,00339	0,0582

A comparação dos Quadrados Médios ou dos Desvios Padrões mostra que as três equações dão resultados praticamente equivalentes, embora a equação de Spurr tenha apenas dois parâmetros, e a de Stoate tenha quatro.

Para a equação de Schumacher-Hall na forma logarítmica

$$\text{Log } V = a + b \text{ Log } D + c \text{ Log } D,$$

temos QMResíduo = 0,02073, Desvio Padrão = 0,1440, mas estes resultados não são comparáveis com os das demais equações, pois se referem a Log V, e não à variável V. Para obter resultado comparável deve-se usar o Índice de Furnival.

O Índice de Furnival (1961)

O problema que se tem em vista é o de comparar o desvio padrão S relativo a uma equação de regressão em que a variável dependente é V, por exemplo:

$$V = a + b\ D^2H$$

com o desvio padrão S_1 referente a outra equação de regressão, em que a variável depende é uma função de V, por exemplo

$$\text{Log } V = a + b \text{ Log } D + c \text{ Log } H,$$

onde Log indica logaritmo decimal.

O problema teve uma solução aproximada deduzida por FURNIVAL (1961), que propôs, para isso, um índice que recebeu o seu nome. No caos particular desta equação logarítmica, em que a variável dependente é Log V, o índice de Furnival tem a seguinte expressão:

$$I = 2{,}3026\ [V]\ S_1,$$

onde [V] indica a média geométrica dos valores de variável V.

Mas o que é média geométrica? Suponha-se que a variável V tenha 2 valores, $V_1 = 8$ e $V_2 = 2$. Sua média geométrica [V] será:

$$[V] = (V_1 \text{ x } V_2)^{1/2} = (8 \text{ x } 2)^{1/2} = \sqrt{14} = 4.$$

No caso de 3 valores $V_1 = 8$, $V_2 = 2$, $V_3 = 5$, a média geométrica seria:

$$[V] = (8 \text{ x } 2 \text{ x } 5)^{1/3} = \sqrt[3]{80} = 4{,}309.$$

No caso de N = 50 valores de V, a média geométrica seria, pois:

$$[V] = (V_1 \text{ x } V_2 \text{ x } V_3 \text{ x} \ldots \text{x } V_{50})^{1/50},$$

e, no caso geral de N valores de V:

$$[V] = (V_1 \text{ x } V_2 \text{ x } V_3 \text{ x} \ldots \text{x } V_N)^{1/N},$$

O modo mais fácil de calcular esta expressão é a aplicação de logaritmos:

$$\text{Log}[V] = \frac{1}{N}(\text{Log } V_1 + \text{Log } V_2 + \text{Log } V_3 + ... + \text{Log } V_N)$$
$$= \frac{1}{N} \text{R Log V.}$$

No caso, dos valores da Tabela 1, tem-se:

$$\text{Log}[V] = \frac{1}{50}(-0{,}936 - 0{,}886 - 0{,}604 - ... - 0{,}658)$$
$$= \frac{1}{50}[-36{,}811]$$
$$= -0{,}73622,$$

logo

$$[V] = \text{antilog } (-0{,}73622) = 10^{-0{,}73622} = 0{,}184.$$

Já o valor do desvio padrão S_1, é obtido facilmente, pois tem-se, para essa equação logarítmica QMResíduo = 0,02073 logo $S_1 = \sqrt{0{,}2073} = 0{,}144$.

Conclui-se, pois, que o índice de Furnival relativo à equação logarítmica em discussão é:

$$I = 2{,}3026\ [V]\ S_1 = 2{,}3026 \times 0{,}184 \times 0{,}144 = 0{,}0610.$$

Este valor é que se pode comparar aos desvios padrões relativos à equação de Spurr ($S = 0{,}0561$), à de Stoate ($S = 0{,}0549$) ou à de Schumacher-Hall ($S = 0{,}0582$). E se conclui, finalmente, que a equação logarítmica perde para as outras, por lhe corresponder um índice de Furnival mais elevado.

4. A DISTRIBUIÇÃO DOS RESÍDUOS

É um outro critério importante de julgamento das equações. Mas o que são resíduos?

Suponha-se a equação de Spurr, por exemplo, para a qual se obtivem a expressão

$$V_c = 0{,}0540 + 0{,}2026\ D^2H$$

Onde V_c indica o valor de V calculado pela equação de regressão. Mas para cada uma das 50 árvores da amostra tem-se também o valor observado V_0. Denomina-se resíduo (ê) a diferença ê = $V_0 - V_c$. O valor absoluto médio desses resíduos já é avaliado pelo Desvio Padrão S, para o qual se tem:

$$S_2 = \text{QMDesvios da Regressão} = \text{QMResíduo}.$$

Mas o computador pode calcular os desvios para todas as árvores da amostra e indicá-los num gráfico em coordenadas cartesianas, com V no eixo das abscissas e os resíduos (ê) no eixo das ordenadas, ou vice-versa, como mostra a Fig. 1.

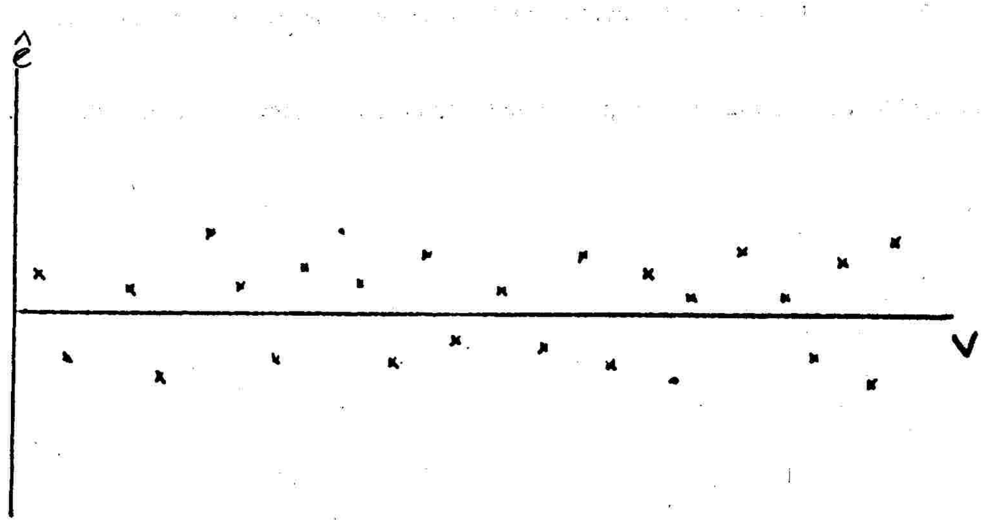

FIG. 1 – Gráfico dos resíduos em condições ideais.

Em condições ideais, os resíduos, positivos e negativos, devem distribuir-se aleatoriamente acima e abaixo do eixo das abscissas, quando é nele que se representam os valores V, ou à esquerda e à direita do eixo das ordenadas, no caso contrário. É o que acontece, aproximadamente na Fig. 1.

Já no caso da Fig. 2, vemos que os desvios positivos se acumulam em certas regiões, e os negativos em outras, o que é mau.

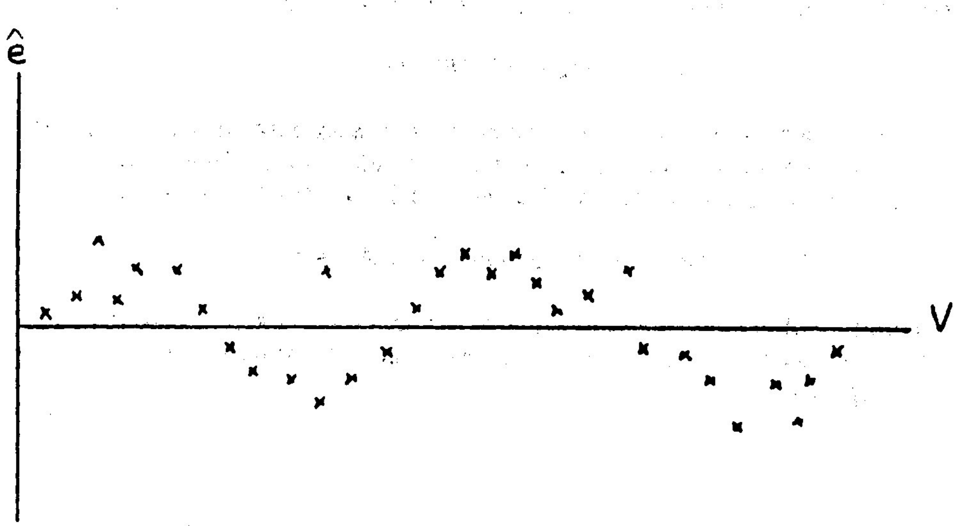

FIG. 2 – Gráfico dos resíduos, com os positivos acumulados em certas regiões e os negativos em outras, o que é mau.

Também é ruim o caso da Fig. 3, em que o valor absoluto dos desvios cresce com o valor da variável V.

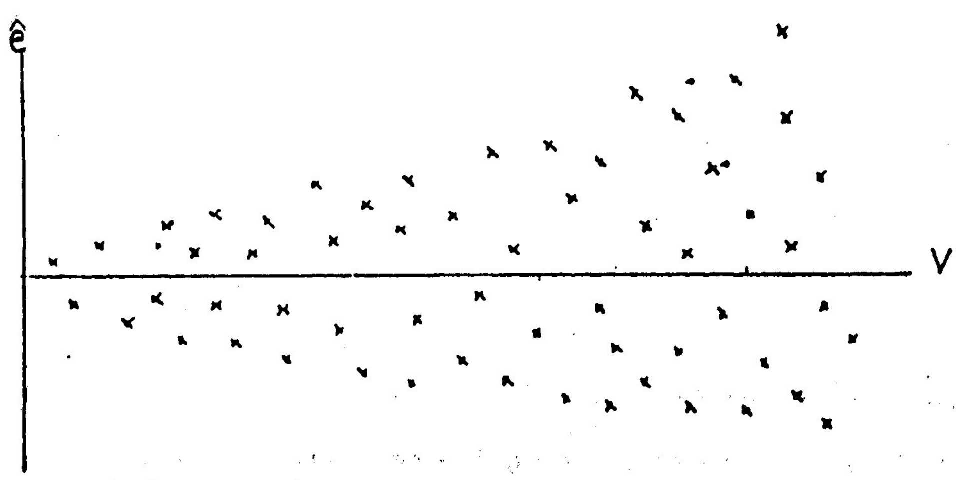

FIG. 3 – Gráfico dos resíduos, que crescem em valor absoluto com o valor de V

5. PARTICULARIDADES DO COEFICIENTE DE DETERMINAÇÃO

O coeficiente de determinação (R^2) é uma estatística geralmente mal compreendida. O assunto é vasto, mas discutir-se-á apenas os aspectos relativos à determinação de equações volumétricas.

Suponha-se que a equação Real de Regressão seja dada por Y = R(D, H). Sendo V_i o volume da árvore **i**, (i = 1, 2, ..., n), com D_i e H_i o DAP e a altura correspondentes, o modelo matemático será:

$$V_i = R(D_i, H_i) + e_i,$$

com $e_i N\cap(0, \sigma^2)$, isto é, com distribuição normal de média zero e variância σ^2. Deste modelo matemático decorre a igualdade:

(1) SQTotal = SQRegr. Real + SQResíduo.

Se a Regressão Real tiver **r** parâmetros, haverá uma estimativa imparcial de σ^2 correspondente a

$$S^2 = \frac{\text{SQResíduo}}{\text{N - r}}$$

Como não se conhece a Equação Real de Regressão Y = R(D, H), admite-se uma Equação de Regressão Adotada Z = A (D_i, H_i) e pode-se escrever:

$$\begin{aligned} V_i &= R(D_i, H_i) + e_i \\ &= A(D_i, H_i) + [R(D_i, H_i) - A(D_i, H_i)] + e_i \\ &= A(D_i, H_i) + U_i + e_i, \end{aligned}$$

onde U_i (i = 1, 2, ..., N) representa os Desvios de Regressão, isto é, a diferença entre o valor $Y_i = R(D_i, H_i)$ da Equação de regressão Real e o valor $Z_i = A(D_i, H_i)$ da Equação de Regressão Adotada. Tem-se, pois,

(2) SQTotal = SQRegr. Adotada + SQDesvios + SQResíduo.

No caso de ser a Equação de Regressão Adotada $Z = A(D_i, H_i)$ igual à Equação de Regressão Real $Y = R(D, H)$, os desvios $U_i = R(D_i, H_i) - A(D_i, H_i)$ serão nulos logo SQDesvios = 0 e se volta à igualdade (1).

Já no caso geral de ser $A(D, H) \neq R(D, H)$, que é o usual, tem-se SQDesvios > 0 e SQTotal = SQRegr. Adotada + SQResíduo Falso. Em tais condições, tem-se uma estimativa inflacionada $(SI)^2$ da variância

$$(SI)^2 = \frac{\text{SQRes.Falso}}{N - a}$$
$$= \frac{\text{SQRDesvios} + \text{SQResíduo}}{N - a},$$

onde **a** é o número de parâmetros da Equação de Regressão Adotada Z = A (D, H).

Discurtir-se-á separadamente os dois casos.

1º Caso: A(D, H) = R(D, H)

É válida então a igualdade (1):

(1) SQTotal = SQRegr. Adotada + SQResíduo,
= SQRegr. Real + SQResíduo.

Tem-se, pois:

$$R^2 = \frac{\text{SQRegr. Adotada}}{\text{SQTotal}}$$
$$= \frac{\text{SQTotal - SQResíduo}}{\text{SQTotal}} \qquad (3)$$
$$= 1 - \frac{\text{SQResíduo}}{\text{SQTotal}}$$

Como se deve ter sempre SQResíduo > 0, conclui-se que, mesmo no caso ideal de ser A(D, H) = R(D, H), tem-se $R^2 < 1$.

Por outro lado, sabe-se que, a partir de (3):

$$R^2 = \frac{\text{SQRegr. Adotada}}{\text{SQTotal}}$$

$$(4) \qquad = \frac{\text{SQRegr. Adotada}}{\text{SQRegr. Adotada} + \text{SQResíduo}},$$

onde SQRegr. Adotada = SQRegr. Real. Daí se conclui que só se pode ter $R^2 = 0$ se for SQRegr. Real = 0, o que é impossível.

A conclusão geral é, ois, de que, mesmo no caso de ser a Equação de Regressão Adotada igual à Equação de Regressão Real, tem-se sempre: $0 < R^2 < 1$.

Por outro lado, de (4) obtém-se:

$$R^2 = \frac{\text{SQRegr. Adotada}}{\text{SQTotal}} = \frac{\text{SQRegr. Real}}{\text{SQTotal}}$$

$$= \frac{\text{SQTotal - SQResíduo}}{\text{SQTotal}}$$

$$= 1 - \frac{\text{SQResíduo}}{\text{SQTotal}}$$

$$= 1 - \frac{(N - r)\, S^2}{\text{SQTotal}}$$

Conclui-se daí, que só se poderá ter valores de R^2 próximos de 1 (um) se for baixo o valor de S^2, isto é, do QMResíduo.

2º Caso: $A(D, H) \neq R(D, H)$

Nestas condições, tem-se SQDesvios > 0, e é válida a igualdade (2). Logo:
SQTotal = SQRegr. Adotada + SQDesvios + SQResíduo, e tem-se ainda:

$$R^2 = \frac{\text{SQRegr. Adot.}}{\text{SQTotal}}$$

$$= \frac{\text{SQTotal - SQDesvios - SQResíduo}}{\text{SQTotal}}$$

$$= 1 - \frac{\text{SQDesvios}}{\text{SQTotal}} - \frac{\text{SQResíduo}}{\text{SQTotal}}$$

Daí decorre que os valores baixos de R^2 podem ter as seguintes explicações:

A. desvios grande $U_i = R(D_i, H_i) - A(D_i, H_i)$, em valor absoluto;

B. valores altos de SQResíduo Falso = $(N - a)\, SI^2$, onde **a** é o número de parâmetros da Regressão Adotada, resultantes de heterogeneidade no povoamento; e

C. Valores baixos da SQRegr. Adotada, devidos, por exemplo, a intervalos de variação pequenos para D e para H.

CONCLUSÕES GERAIS

I. Mesmo no caso de ser a Equação de Regressão Adotada igual à Equação de Regressão Real, isto é, A(D, H) = R(D, H), e de se ter, pois, SQDesvios = 0, o valor de R^2 pode ser baixo ou alto, dependendo do valor da SQResíduo = (N – r) S^2.

Exemplo 1: SQRegr. Adotada = SQRegr. Real = 100, SQDesvios = 0, SQResíduo = 100.

$$R^2 = 1 - \frac{\text{SQResíduo}}{\text{SQTotal}} = 1 - \frac{100}{100 + 0 + 100} = 0{,}50.$$

Exemplo 2: SQRegr. Adotada = SQRegr. Real = 100, SQDesvios = 0, SQResíduo = 20.

$$R^2 = 1 - \frac{\text{SQResíduo}}{\text{SQTotal}} = 1 - \frac{20}{100 + 0 + 20} = 0{,}83.$$

II. Valores mais altos de R^2 não garantem maior adequação da Equação Adotada.

Exemplo 3: SQRegr. Adotada = 100, SQDesvios = 50, SQResíduo = 10.

$$R^2 = 1 - \frac{\text{SQDesvios} + \text{SQResíduo}}{\text{SQTotal}}$$

$$= 1 - \frac{50 + 10}{100 + 50 + 10} = 0{,}62.$$

Exemplo 4: SQRegr. Adotada = 100, SQDesvios = 10, SQResíduo = 90.

$$R^2 = 1 - \frac{\text{SQDesvios} + \text{SQResíduo}}{\text{SQTotal}} = 1 - \frac{10 + 90}{200} = 0{,}50.$$

Está claro que a Equação de Regressão Adotada se ajusta melhor à Equação de Regressão Real no Exemplo 4 do que no Exemplo 3. No entanto, o valor de R^2 dá idéia contrária, pois é maior no exemplo 3 do que no exemplo 4.

6. O MÉTODO DOS QUADRADOS MÍNIMOS PONDERADOS

Os métodos usuais de estimativa dos parâmetros de um modelo de regressão admitem que os erros de todas as observações tenham uma mesma variância. Na prática, porém, a variância desses erros costuma variar com o tamanho das árvores (FURNIVAL, 1961), de tal sorte que para árvores grandes essa variância é maior. Para contornar essa dificuldade, convém adotar o método dos quadrados mínimos ponderados, tendo como fator de ponderação a variável $W = D^2H$. Nestas condições a equação de Spurr $V = a + b\, D^2H$, por exemplo, é dividida por D^2H e fica:

$$U = \frac{1}{D^2H} \text{ x } V = a \text{ x } \frac{1}{D^2H} + b = a\,X + B,$$

Obtem-se pois a equação de regressão

(4) $U = a\,X + b$,

com $U = f(V) = \frac{1}{D^2H} \text{ x } V, \quad X = \frac{1}{D^2H}$.

Sendo S_1 o desvio padrão relativo à equação (4), ele não será comparável ao S das equações usuais de Spurr, de Stoate e de Schumacher-Hall. Para fazer a comparação, faz-se necessário calcular o índice de Furnival.

Sendo $U = f(V)$ a variável dependente, a expressão geral do índice de Furnival é:

$I = [f'(V)]^{-1}\, S_1$, onde $f'(V)$ é a derivada de $f(V)$ e os colchetes indicam média geométrica.

No caso da equação (4) tem-se

$$f'(V) = \frac{1}{D^2H}$$

e fica

$$I = [D^2H]\, S_1 .$$

No caso da equação de Stoate.

$$V = a + b\, D^2H + c\, D^2 + fH,$$

a divisão pelo fator de ponderação D^2H dá:

$$\begin{aligned}\frac{V}{D^2H} &= a\,\frac{1}{D^2H} + b + c\,\frac{1}{H} + f\,\frac{1}{D^2} \\ &= b + a\,X_1 + c\,X_2 + fX_3,\end{aligned}$$

com $X_1 = 1/D^2H$, $X_2 = 1/H$, $X_3 = 1/D^2$.

Por sua vez a equação de Schumacher-Hall fica:

$$\frac{V}{D^2H} = A\, D^m H^m,$$

com m = b-2, n = c-1

Em ambos os casos, o índice de Furnival tem a mesma expressão referente à equação de Spurr, isto é:

$$I = [D^2H]\ S_1.$$

Equação de Spurr – Exemplo

Considera-se a expressão

$$U = a\ X + b,$$

com $U = V/D^2H$, $X = 1/D^2H$ é o fator de ponderação para uso do método dos quadrados mínimos ponderados. O programa GLM do SAS aplicado aos dados da Tabela 2 dá a equação: (Listagem nº 10).

(5) $U = 0{,}2412 + 0{,}20095\ Z$

isto é:

$$\frac{V}{D^2H} = 0{,}2412 + 0{,}20095\,\frac{1}{D^2H},$$

ou ainda:

$$V = 0{,}2012 + 0{,}2412\ D^2H$$

A análise da variância relativa à equação (5) é a seguinte:

C. de variação	G.L.	S.Q.	Q.M.	F
Regressão	1	0,009104	0,009104	2,33
Resíduo	38	0,148766	0,003915	
Total	39	0,157870		

O desvio padrão é, pois, $S_1 = \sqrt{0{,}003915} = 0{,}0626.$

TABELA 2 – Dados dendrométricos de 40 árvores de **Eucalyptus saligna** Smith de 10 anos de idade.

Arv.	D(m)	H(m)	$V(m^3)$
1	0,190	14,0	0,130
2	0,270	17,0	0,249
3	0,150	11,0	0,047
4	0,187	16,0	0,131
5	0,245	14,0	0,203
6	0,237	18,0	0,270
7	0,250	14,0	0,202
8	0,235	14,0	0,214
9	0,200	12,0	0,129
10	0,177	13,0	0,144
11	0,283	16,0	0,305
12	0,275	18,0	0,357
13	0,205	15,0	0,160
14	0,200	12,0	0,254
15	0,204	12,0	0,146
16	0,215	15,0	0,211
17	0,270	14,0	0,199
18	0,160	12,0	0,088
19	0,215	14,0	0,148
20	0,205	16,0	0,173
21	0,203	14,0	0,212
22	0,215	14,0	0,156
23	0,260	18,0	0,438
24	0,310	16,0	0,255
25	0,185	12,0	0,091
26	0,215	16,0	0,177
27	0,220	14,0	0,239
28	0,250	16,0	0,261
29	0,315	18,0	0,397
30	0,210	18,0	0,209
31	0,273	18,0	0,328
32	0,170	16,0	0,123
33	0,202	17,0	0,201
34	0,254	19,0	0,352
35	0,180	14,0	0,133
36	0,190	14,0	0,156
37	0,223	14,0	0,200
38	0,174	12,0	0,135
39	0,364	18,0	0,572
40	0,245	16,0	0,226

Para a equação de Spurr sem ponderação obtém-se a expressão seguinte:

$$V = 0{,}0404 = 0{,}2123\ D^2H,$$

com $S = 0,0453$. Mas estes valores de S e S_1 não são comparáveis. Para que o sejam, calcula-se o índice de Furnival:

$$I = [D^2H]S_1 = 0,7303 \times 0,0626 = 0,0457$$

Verifica-se que este valor $I = 0,0457$ combina bem com o de $S = 0,0453$.

7. REFERÊNCIAS BIBLIOGRÁFICAS

DRAPER, N.R. & SMITH, H. **Applied regression analysis**, 2. ed. New York, John Wiley, 1981.

FURNIVAL, G.M. An index for comparing equations used in constructing volume tables. **Forest science**, Madison, 7(4): 337-41, 1961.

PIMENTEL-GOMES, F. **Curso de estatística experimental**, 13. ed. São Paulo, Nobel, 1990.

THÖNI, H. Transformations of variables uses in the analysis of experimental and observational data. **Technical report. ISU. Statistical Laboratory**, Ames (7), 1967.

VEIGA, R.A. de A. O uso de equações de volume em levantamentos florestais. In: SIMPÓSIO SOBRE INVENTÁRIO FLORESTAL, 2, Piracicaba, 1984. **Anais**. Piracicaba, ESALQ/IPEF, 1984. p.93-102.

ANEXOS

LISTAGEM N^{o} 1: ANÁLISE PELO PROGRAMA REG DO SAS

Model: MODEL1
Dependent Variable: V

Analysis of Variance

Source	DF	Sun of Squares	Mean Square	F Value	Prob>F
Model	1	0.39955	0.39955	126.794	0.0001
Error	48	0.15126	0,00315		
C Total	49	0.55080			
Root MSE		0.05614	R-square	0.7254	
Dep Mean		0.21036	Adj R-sq	0.7197	
C.V.		26.68527			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter = 0	Prob > :T:
INTERCEP	1	0.054002	0.01599490	3.376	0.0015
D2H	1	0.202614	0.01799370	11.260	0.0001

LISTAGEM Nº 2: ANÁLISE PELO REGRESEQ DO SAEG

MODELO = V FUNÇÃO D2H

ESTATISTICA SIMPLES

VARIAVEL	MÉDIA	DESVIO-PADRÃO
V	.2104	.1060
D2H	.7718	.4456

MATRIZ DE CORRELAÇÃO

	V	D2H
V	1.00000	.85175
D2H	.85175	1.00000

******VARIAVEL DEPENDENTE = V MODELO COMPLETO******

NOME	COEFICIENTE	DESVIO	T	BETA	PROBAB.
COSNTANTE	.539700E-01				
D2H	.202643E+00	.179922E-01	.112628E+02	.851752E+00	.0000

SOMA DE QUADRADOS DO MODELO = .3995977
SOMA DE QUADRADOS DE D2H = .3995977
COEF. DETERMINAÇÃO = .7254812
COEF. DETERMINAÇÃO AJUSTADO = .7197621

LISTAGEM Nº 3: ANÁLISE PELO REG DO SAS

Model: MODEL3
Dependent Variable: V

Analysis of Variance

Source	DF	Sun of Squares	Mean Square	F Value	Prob>F
Model	1	0.41212	0.13737	45.567	0.0001
Error	46	0.13868	0.00301		
C Total	49	0.55080			
Root MSE		0.05491	R-square	0.7482	
Dep Mean		0.21036	Adj R-sq	0.7318	
C.V.		26.10150			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter = 0	Prob > :T:
INTERCEP	1	0.206095	0.08888062	2.319	0.0249
D2H	1	0.494497	0.14464562	3.419	0.0013
D2	1	-5.030471	2.46708080	-2.039	0.0472
H	1	-0.008542	0.00605595	-1.410	0.1651

LISTAGEM N° 4: ANÁLISE PELO GLM DO SAS.

General Linear Models Procedure
Number of observations in data set = 50
General Linear Models Procedure

Dependent Variables: V

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.41212313	0.13737438	45.57	0.0001
Error	46	0.13868039	0.00301479		
Corrected Total	49	0.55080352			

R-Square	C.V.	Root MSE	V Mean
0.748222	26.10150	0.054907	0.21036000

General Linear Models Procedure

Dependent Variable: V

Source	DF	Type I SS	Mean Square	F Value	Pr > F
D2H	1	0.39954818	0.39954818	132.53	0.0001
D2	1	0.00657756	0.00657756	2.18	0.1465
H	1	0.00599739	0.00599739	1.99	0.1651

Source	DF	Type III SS	Mean Square	F Value	Pr > F
D2H	1	0.03523502	0.03523502	11.69	0.0013
D2	1	0.01253452	0.01253452	4.16	0.0472
H	1	0.00599739	0.00599739	1.99	0.1651

General Linear Models Procedure

Dependent Variable: V

Parameter	Estunate	T for HO: Parameter = 0	Pr > :T:	Std Error of Estimate
INTERCEPT	0.206094741	2.32	0.0249	0.08888062
D2H	0.494497341	3.42	0.0013	0.14464562
D2	-5.030471242	-2.04	0.0472	2.46708080
H	-0.008541514	-1.41	0.1651	0.00605595

LISTAGEM N° 5: ANÁLISE PELO REGRESEQ DO SAEG

REGRESEQ MODELO = V FUNÇÃO D2H H

ESTATISTICAS SIMPLES

VARIAVEL	MEDIA	DESVIO-PADRAO
V	.2104	.1060
D2H	.7718	.4456
D2	.0503	.0227
H	14.5060	2.7062

MATRIZ DE CORRELAÇÕES

	V	D2H	D2	H
V	1.00000	.85175	.81678	.65690
D2H	.85175	1.00000	.98294	.77758
D2	.81678	.98294	1.00000	.69003
H	.65690	.77758	.69003	1.00000

******VARIÁVEL DEPENDENTE = V MODELO COMPLETO******

PARAMETROS DA REGRESSAO

NOME	COEFICIENTE	DESVIO	T	BETA	PROBAB.
CONSTANTE	.539700E-01				
D2H	.202643E+00	.179922E-01	.112628E+02	.851752E+00	.0000

SOMA DE QUADRADOS DO MODELO = .3995977
SOMA DE QUADRADOS DE D2H = .3995977
COEF. DETERMINACAO = .7254812
COEF. DETERMINACAO AJUSTADO = .7197621

PARAMETROS DE REGRESSAO

NOME	COEFICIENTE	DESVIO	T	BETA	PROBAB.
CONSTANTE	.869234E-01				
D2H	.343893E+00	.966054E-01	.355977E+01	.144546E+01	.0004
D2	-.282022E+01	.189593E+01	-.148751E+01	-.604011E+00	.0718

SOMA DE QUADRADOS DO MODELO = .4063960
SOMA DE QUADRADOS DE D2H = .6798267E-02
COEF. DETERMINACAO = .7378237
COEF. DETERMINACAO AJUSTADO = .7266672

PARAMETROS DE REGRESSAO

NOME	COEFICIENTE	DESVIO	T	BETA	PROBAB.
CONSTANTE	.204968E+00				
D2H	.495086E+00	.143580E+00	.344816E+01	.208095E+01	.006
D2	-.503943E+01	.244797E+01	-.205862E+01	-.205862E+01	.0226
H	-.848055E-02	.600862E-02	-.141140E+01	-.216462E+00	.0824

SOMA DE QUADRADOS DO MODELO = .4123902
SOMA DE QUADRADOS DE D2H = .5994201E-02
COEF. DETERMINACAO = .7487063
COEF. DETERMINACAO AJUSTADO = .7323176

LISTAGEM Nº 6: ANÁLISE PELO REG DO SAS

Model: MODEL2
Dependent Variable: LV

Analysis of Variance

Source	DF	Sun of Squares	Mean Square	F Value	Prob>F
Model	1	1.93956	0.96968	46.790	0.0001
Error	47	0.97412	0.02073		
C Total	49	2.91368			
Root MSE		0.14397	R-square	0.6657	
Dep Mean		-0.73622	Adj R-sq	0.6514	
C.V.		-19.55467			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter = 0	Prob > :T:
INTERCEP	1	-0.578186	0.60774209	-0.951	0.3463
LD	1	1.480841	0.33436598	4.429	0.0001
LH	1	0.723442	0.36172034	2.000	0.0513

LISTAGEM Nº 7: ANÁLISE PELO NLIN DO SAS

Non-Linear Least Squares Interative Phase
Dependent Variable V Method: Gauss-Newton

Iter	A	B	C	Sum of Squares
0	0.300000	1.500000	0.700000	0.162170
1	0.270997	1.494329	0.732251	0.161767
2	0.247766	1.489424	0.761130	0.161477
3	0.228959	1.485193	0.786926	0.161232
4	0.213583	1.481550	0.809927	0.161007
5	0.188214	1.475285	0.850893	0.160935
6	0.171161	1.470769	0.882946	0.160668
7	0.147494	1.464309	0.932809	0.160540
8	0.126549	1.458769	0.988251	0.159760
9	0.125988	1.459713	0.994264	0.159450
10	0.125654	1.459765	0.995224	0.159450
11	0.125606	1.459777	0.995369	0.159450
12	0.125599	1.459778	0.995392	0.159450

NOTE: Convergence criteriom met.

Non-Linear Least Squares Summary Statistics Dependent Variable V

Source	DF	Sum of Squares	Mean Square
Regression	3	2.6039197832	0.8679732611
Residual	47	0.1594502168	0.0033925578
Uncorrected Total	50	2.7633700000	
(Corrected Total)	49	0.5508035200	

Parameter	Estimate	Asymptotic Std. Error	Asymptotic 95% Confidence Interval	
			Lower	Upper
A	0.125598613	0.15525249400	-.18672827444	0.4379254997
B	1.459778266	0.23763709107	0.98171539583	1.9378411364
C	0.995391729	0.36069347694	0.26977201136	1.7210114463

Asymptotic Correlation Matrix

Corr	A	B	C
A	1	0.7806546373	-0.978845088
B	0.7806546373	1	-0.638262231
C	-0.978845088	-0.63826231	1

LISTAGEM Nº 8: ANÁLISE PELO REGREGN DO SAEG – 1ª APROXIMAÇÃO

V FUNÇÃO A*D**B*H**C

AMPLITUDES (A = 0.1, 0.3, 0.264) (B = 1.3, 1.6, 1.48) (C = 0.5, 1.1, 0.723)

TOLERANCIA = .1000000E-02

ITERACOES PERMITIDAS = 60

ITERACOES EXECUTADAS = 22

NUMERO DE AVALIACOES = 59

NUMERO DE OBSERVACOES = 50

COEFICIENTE DE DETERMINACOA = .7105137E+00

PARAMETROS	ESTIAMTIVAS INICIAIS	LIMITES INFERIORES	LIMITES SUPERIORES	COEFICIENTES DE RESTRICAO	COEFICIENTES DA SOLUCAO	DESVIOS DOS COEFICIENTES
A	.264000E+00	.100000E+00	.300000E+00	.200000E-04	.123492E+00	.152451E+00
B	.148000E+01	.130000E+01	.160000E+01	.300000E-04	.145722E+01	.237291E+00
C	.723000E+00	.500000E+00	.110000E+01	.600000E-04	.100019E+01	.360258E+00

TABELA DE RESÍDUOS

NÚMERO DA OBSERVAÇÃO	V OBSERVADO	V ESTIMADO	RESÍDUO	GRÁFICO DOS RESÍDUOS STANDARTIZADOS -2.0 -1.0 0.0 1.0 2.0
1	.116000E+00	.114030E+00	.197044E-02	I*
2	.130000E+00	.153803E+00	-.238033E-01	* I
3	.249000E+00	.311666E+00	-.626663E-01	* I
4	.470000E-01	.856265E-01	-.386265E-01	* I
5	.410000E-01	.518421E-01	-.108421E-01	* I
6	.131000E+00	.171750E+00	-.407498E-01	* I
7	.203000E+00	.222772E+00	-.197725E-01	* I
8	.270000E+00	.272908E+00	-.290826E-02	*I
9	.710000E-01	.812403E-01	-.102403E-01	* I
10	.202000E+00	.229428E+00	-.274283E-01	* I
11	.214000E+00	.209647E+00	.435317E-02	I *
12	.129000E+00	.142059E+00	-.130587E-01	* I
13	.144000E+00	.128802E+00	.151985E-01	I *
14	.305000E+00	.314135E+00	-.913492E-02	* I
15	.357000E+00	.338946E+00	.180539E-01	I *
16	.160000E+00	.184087E+00	-.240871E-01	* I
17	.254000E+00	.142059E+00	.111941E+00	I *
18	.146000E+00	.146218E+00	-.217825E-03	*I
19	.211000E+00	.197317E+00	.136826E-01	I *
20	.199000E+00	.256657E+00	-.576569E-01	* I
21	.880000E-01	.102624E+00	-.146238E-01	* I
22	.148000E+00	.184161E+00	-.361605E-01	* I
23	.173000E+00	.196362E+00	-.233620E-01	* I
24	.212000E+00	.169375E+00	.426248E-01	I *
25	.156000E+00	.184161E+00	-.281605E-01	* I
26	.438000E+00	.312344E+00	.125656E+00	I X
27	.255000E+00	.358745E+00	-.103745E+00	* I
28	.910000E-01	.126803E+00	-.358028E-01	* I
29	.177000E+00	.210474E+00	-.334745E-01	* I
30	.239000E+00	.190435E+00	.485654E-01	I *
31	.430000E-01	.560420E-01	-.130420E-01	* I
32	.261000E+00	.262210E+00	-.121039E-02	*I
33	.397000E+00	.413118E+00	-.161184E-01	* I
34	.209000E+00	.228807E+00	-.198074E-01	* I
35	.328000E+00	.335360E+00	-.735989E-02	* I
36	.123000E+00	.149478E+00	-.264782E-01	* I
37	.201000E+00	.204203E+00	-.320268E-02	*I
38	.352000E+00	.318672E+00	.333284E-01	I *
39	.133000E+00	.142151E+00	-.915052E-02	* I
40	.156000E+00	.153803E+00	.219670E-02	I *
41	.200000E+00	.194231E+00	.576949E-02	I *
42	.135000E+00	.115967E+00	.190332E-01	I *
43	.572000E+00	.510006E+00	.619944E-01	I *
44	.226000E+00	.254604E+00	-.286035E-01	* I
45	.301000E+00	.242923E+00	.580771E-01	I *
46	.162000E+00	.153803E+00	.819671E-02	I *
47	.234000E+00	.257638E+00	-.236379E-01	* I
48	.365000E+00	.330003E+00	.349968E-01	I *
49	.344000E+00	.543413E-01	.289659E+00	I Y
50	.220000E+00	.223435E+00	-.343539E-02	*I

LISTAGEM Nº 9: ANÁLISE PELO REGREGN DO SAEG – 2ª APROXIMAÇÃO

REGREGN V FUNÇÃO A*D**B*H**C

AMPLITUDES (A = 0.1, 0.14, 0.12349) (B = 1.4, 5.1, 1.4572) (C = 0.9, 1.1, 1.000)

TOLERANCIA = .1000000E-02

ITERACOES PERMITIDAS = 60

ITERACOES EXECUTADAS = 8

NUMERO DE AVALIACOES = 18

NUMERO DE OBSERVACOES = 50

COEFICIENTE DE DETERMINACOA = .7105141E+00

PARAMETROS	ESTIAMTIVAS INICIAIS	LIMITES INFERIORES	LIMITES SUPERIORES	COEFICIENTES DE RESTRICAO	COEFICIENTES DA SOLUCAO	DESVIOS DOS COEFICIENTES
A	1.23490E+00	.100000E+00	.140000E+00	.400000E-04	.125856E+00	.155722E+00
B	.145720E+01	.140000E+01	.150000E+01	.100000E-04	.145978E+01	.238015E+00
C	.100000+01	.900000E+00	.110000E+01	.200000E-04	.994644E+00	.360689E+00

TABELA DE RESÍDUOS

NÚMERO DA OBSERVAÇÃO	V OBSERVADO	V ESTIMADO	RESÍDUO	GRÁFICO DOS RESÍDUOS STANDARTIZADOS -2.0 -1.0 0.0 1.0 2.0
1	.116000E+00	.114108E+00	.189246E-02	I*
2	.130000E+00	.153816E+00	-.238161E-01	* I
3	.249000E+00	.311636E+00	-.626365E-01	* I
4	.470000E-01	.856965E-01	-.386965E-01	* I
5	.410000E-01	.519148E-01	-.109148E-01	* I
6	.131000E+00	.171630E+00	-.406300E-01	* I
7	.203000E+00	.222936E+00	-.199357E-01	* I
8	.270000E+00	.272705E+00	-.270486E-02	*I
9	.710000E-01	.813862E-01	-.103862E-01	* I
10	.202000E+00	.229608E+00	-.276084E-01	* I
11	.214000E+00	.209778E+00	.422183E-02	I *
12	.129000E+00	.142211E+00	-.132107E-01	* I
13	.144000E+00	.128842E+00	.151581E-01	I *
14	.305000E+00	.314248E+00	-.924829E-02	* I
15	.357000E+00	.338822E+00	.181779E-01	I *
16	.160000E+00	.184068E+00	-.240677E-01	* I
17	.254000E+00	.142211E+00	.111789E+00	I *
18	.146000E+00	.146382E+00	-.381604E-03	*I
19	.211000E+00	.197321E+00	.136793E-01	I *
20	.199000E+00	.256909E+00	-.579088E-01	* I
21	.880000E-01	.102675E+00	-.146750E-01	* I
22	.148000E+00	.184234E+00	-.362340E-01	* I
23	.173000E+00	.196271E+00	-.232711E-01	* I
24	.212000E+00	.169418E+00	.425821E-01	I *
25	.156000E+00	.184234E+00	-.282340E-01	* I
26	.438000E+00	.312185E+00	.125815E+00	I X
27	.255000E+00	.358958E+00	-.103958E+00	* I
28	.910000E-01	.126913E+00	-.359131E-01	* I
29	.177000E+00	.210403E+00	-.334026E-01	* I
30	.239000E+00	.190522E+00	.484782E-01	I *
31	.430000E-01	.562197E-01	-.132197E-01	* I
32	.261000E+00	.262222E+00	-.122195E-02	*I
33	.397000E+00	.413111E+00	-.161107E-01	* I
34	.209000E+00	.228566E+00	-.195662E-01	* I
35	.328000E+00	.335231E+00	-.723103E-02	* I
36	.123000E+00	.149338E+00	-.263375E-01	* I
37	.201000E+00	.204032E+00	-.303185E-02	*I
38	.352000E+00	.318395E+00	.336050E-01	I *
39	.133000E+00	.142143E+00	-.914271E-02	* I
40	.156000E+00	.153816E+00	.218391E-02	I*
41	.200000E+00	.194326E+00	.567384E-02	I *
42	.135000E+00	.116050E+00	.189504E-01	I *
43	.572000E+00	.510185E+00	.618155E-01	I *
44	.226000E+00	.254602E+00	-.286015E-01	* I
45	.301000E+00	.243138E+00	.578621E-01	I *
46	.162000E+00	.153816E+00	.818393E-02	I *
47	.234000E+00	.257641E+00	-.236412E-01	* I
48	.365000E+00	.329867E+00	.351329E-01	I *
49	.344000E+00	.544401E-01	.289560E+00	I Y
50	.220000E+00	.223383E+00	-.338255E-02	*I

LISTAGEM Nº 10: ANÁLISE PELO GLM DO SAS.

General Linear Models Procedure
Number of observations in data set = 40
General Linear Models Procedure

Dependent Variables: U

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00910449	0.00910449	2.33	0.1355
Error	38	0.14876571	0.00391489		
Corrected Total	39	0.15787019			

R-Square	C.V.	Root MSE	U Mean
0.057671	22.98904	0.062569	0.27216907

General Linear Models Procedure

Dependent Variable: U

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Z	1	0.00910449	0.00910449	2.33	0.1355

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Z	1	0.00910449	0.00910449	2.33	0.1355

Parameter	Estunate	T for HO: Parameter = 0	Pr > :T:	Std Error of Estimate
INTERCEPT	0.2412355704	10.69	0.0001	0.2256827
Z	0.201206036	1.52	0.1355	0.01319398