

# IMPORTÂNCIA DA DISTRIBUIÇÃO DE RESÍDUOS DE REGRESSÃO NA SELEÇÃO DE EQUAÇÕES DE VOLUME

Sergio Ahrens\*

## RESUMO

Características do ajustamento de cinco equações de volume foram avaliadas quanto ao seu desempenho na estimativa do volume comercial para lenha de bracatinga (**Mimosa scabrella** Benth.). Os critérios utilizados para a avaliação comparativa foram: a) exame da análise de variância para regressão, b) erro padrão da estimativa ( $s$  e  $s\%$ ), c) coeficiente de determinação ( $r^2$  e  $R^2$ ) e d) amplitude de distribuição gráfica dos resíduos volumétricos percentuais.

Dentre os modelos analisados, a equação do fator de forma constante  $V = 0,3879 d^2 h$  foi selecionada por apresentar  $s\% = 28,56$ ,  $r^2 = 0,9718$  e uma amplitude de distribuição consistente dos resíduos de  $\pm 30\%$  em relação ao volume cubado das 81 árvores-amostra. Estas características não recomendam o uso indiscriminado na equação. Na análise comparativa, entretanto, aquela foi considerada a função menos inadequada aos dados.

O estudo permitiu adicionalmente constatar que: a) um elevado coeficiente de determinação não é condição suficiente para indicar o uso de uma equação, b) a análise e a interpretação da distribuição de resíduos devem sempre ser consideradas na seleção de modelos matemáticos descritivos, e c) as técnicas de análise de regressão passo-a-passo e análise de regressão ponderada podem ser úteis em muitos problemas de regressão em pesquisa florestal.

## ABSTRACT

Five volume equations were evaluated in their prediction ability to estimate commercial volume for firewood of bracatinga (**Mimosa scabrella** Benth.). The criteria used for the comparative analysis were: a) examination of the analysis of variance for regression, b) standard error of the estimate ( $s$  and  $s\%$ ), c) coefficient of determination ( $r^2$  and  $R^2$ ), and d) percentage volume residuals distribution.

Among the models analysed, the constant form factor equation  $V = 0.3879 d^2 h$  was chosen because of its  $s\% = 28.56$ ,  $r^2 = 0.9718$  and a consistent distribution of the deviations ( $\pm 30\%$ ) in respect to the observed volume of 81 sample trees. These were found to be undesirable characteristics for a descriptive model. However, the selected function was, comparatively, the less inadequate for the data.

With this investigation, it was also possible verify that: a) a high coefficient of determination is not, and by no means, sole and sufficient condition to recommend an equation, b) the analysis and interpretation of the graphical distribution of residuals should always be considered in selecting descriptive functions, and c) stepwise regression analysis and weighted-least-squares analysis might be useful for several regression problems in forest research.

**PALAVRAS-CHAVE:** Análise de regressão; resíduos de regressão, equações de volume; **Mimosa scabrella**, bracatinga.

---

\* Eng.º Ftal., M.Sc., Pesquisador da Unidade Regional de Pesquisa Florestal Centro-Sul – URPFCS (PNPF-EMBRAPA/IBDF).

## 1. INTRODUÇÃO

O volume do tronco de uma árvore é função do seu diâmetro, altura e forma e um dos métodos mais difundidos e utilizados para a sua estimativa é o uso de equações de volume.

O número de modelos matemáticos para volumetria observados na literatura florestal é extremamente elevado. Segundo LOETSCH, ZOHREER & HALLER (1973) este fato é forte evidência de que não existe uma equação de volume aplicável para todas as espécies. Desta forma, investigadores têm procurado desenvolver equações de volume adequadas para cada situação específica. No entanto, a análise preliminar de modelos utilizados anteriormente com sucesso tem sido aceita como prática recomendável em estudos de volumetria, o que implica em uma avaliação comparativa do ajustamento das várias funções aos dados.

Este trabalho descreve em detalhes a metodologia utilizada na seleção de uma equação de volume comercial por árvore para bracinga (**Mimosa scabrella** Benth.). O estudo enfatiza os segmentos mais críticos do processo subjetivos de análise comparativa dos critérios utilizados para definir a adequabilidade dos modelos.

Embora o estudo tenha sido efetivado com equações de volume, a interpretação dos resultados da análise de regressão pode ser aplicada a outras investigações em que aquela técnica estatística seja utilizada. Os princípios teóricos que regem a disciplina são aceitos universalmente e aplicáveis à grande maioria das situações onde o objetivo seja a definição de um modelo matemático descritivo.

## 2. REVISÃO BIBLIOGRÁFICA

Uma análise crítica e compreensiva de um elevado número de equações de volume foi apresentada por SPURR (1952), que descreveu detalhadamente as vantagens e as deficiências de diversos modelos. Em adição ao trabalho de Spurr, outras funções de volume foram acrescentadas à literatura. Dentre as contribuições mais importantes citam-se aquelas de HUSCH (1965) e, mais recentemente, os trabalhos de HUSCH, MILLER & BEERS (1972) e LOETSCH, ZOHREER & HALLER (1973).

Os primeiros estudos sobre o desenvolvimento de equações de volume fizeram uso de métodos gráficos e subjetivos como a técnica de curvas harmonizadas e as tabelas de alinhamento (HUSCH, MILLER & BEERS 1972). Atualmente, entretanto, a literatura registra um domínio marcante de procedimentos algébricos na definição de funções de forma de tronco, quando volume é obtido por integração de uma função de forma (AHRENS 1980), e no ajustamento de equações de volume. Nestas últimas investigações, as técnicas de análise de regressão têm, invariavelmente, sido utilizadas. Em essência, a análise consiste em ajustar um modelo de regressão a um conjunto de dados a fim de descrever matematicamente a relação probabilística entre uma variável dependente  $Y$  (volume) e uma ou mais variáveis independentes  $X_j$ ,  $j = 1 \dots p$  (normalmente diâmetro à altura do peito,  $d$  ou DAP, e altura total das árvores,  $h$ , ou combinações destas).

O desenvolvimento de uma equação de volume é complexo, mas pode ser executado de maneira muito eficiente por meio do correto uso das técnicas de análise de regressão. As facilidades atualmente disponíveis para processamento eletrônico de dados simplificam tremendamente o processo de manipulação algébrica de informações em diversos centros de pesquisa (LOETSCH, ZOHREER & HALLER 1973). Segundo NETER & WASSERMAN (1974), entretanto, um dos mais difíceis e frequentes problemas em regressão tem sido a seleção de um modelo matemático que se ajuste adequadamente aos dados. Os conhecimentos de álgebra necessários são relativamen-

te simples, mas a interpretação dos resultados numéricos e gráficos da análise é um processo pragmático, depende muito da experiência prévia do investigador, requer um conhecimento dos fundamentos teóricos da matéria e envolve o julgamento subjetivo.

Antes de se desenvolver uma nova equação de volume para um conjunto específico de dados, recomenda-se submeter à análise, modelos matemáticos sugeridos em literatura, a fim de permitir uma avaliação comparativa do desempenho das diversas equações (SPURR 1952 e LOETSCH, ZOHRER & HALLER 1973). Para o exame das equações ajustadas a um mesmo conjunto de dados, diversos critérios são sugeridos (FURNIVAL 1961, DRAPER & SMITH 1966, NETER & WASSERMAN 1974) e dentre os mais utilizados citam-se:

- exame do quadro da análise de variância
- coeficientes de determinação ( $r^2$  e  $R^2$ )
- erro padrão da estimativa ( $s_{Y.X1, \dots, Xp}$ )
- teste F seqüencial
- teste F parcial
- exame da distribuição gráfica dos resíduos
- intervalos de confiança para os parâmetros do modelo ( $\beta$ 's)
- intervalos de confiança para as estimativas ( $\hat{V}_i$ )
- índice de Furnival.

Alguns destes critérios têm sido utilizados em maior ou menor intensidade em investigações conduzidas tanto no Brasil como em outros países. Trabalhos recentes de pesquisadores brasileiros indicam um crescente aprimoramento no uso destes critérios. Entre outros estudos citam-se os de SIQUEIRA (1977), COUTO (1977), PAULA NETO et al. (1977), SILVA et al. (1978), MACHADO (1979) e VEIGA (1981).

### 3. MATERIAL E MÉTODOS

O material experimental utilizado para este estudo foi obtido com 81 árvores de bracatinha (**Mimosa scabrella** Benth.) conforme descrito por AHRENS (1981). As variáveis observadas (medidas) em cada indivíduo foram:

V = volume comercial para lenha,  $m^3$  (cubagem, segundo o método de Smalian, utilizando-se secções de 0,50 m e considerando-se o tronco principal e ramificações até um diâmetro limite comercial com casca igual a 4,0 cm).

d = diâmetro à altura do peito, DAP, m.

h = altura total, m.

Ao conjunto de dados composto pelos valores observados para estas três variáveis em cada uma das 81 árvores-amostra, submeteu-se o ajuste de cinco funções de volume selecionadas para a análise, considerando-se a consistência com que diferentes autores obtiveram sucesso em trabalhos anteriores de pesquisa em volumetria. Adicionalmente, a escolha destas equações baseou-se também na análise comparativa descrita por SPURR (1952). As funções utilizadas são identificadas como segue:

- 1) Modelo de Naslund (Naslund 1940, citado por SPURR 1952)

$$V = b_1d^2 + b_2d^2h + b_3dh^2 + b_4h^2$$

- 2) Modelo compreensivo (MEYER 1944)

$$V = b_0 + b_1d + b_2d^2 + b_3dh + b_4d^2h + b_5h$$

- 3) Modelo australiano (STOATE 1945)  

$$V = b_0 + b_1d^2 + b_2d^2h + b_3h$$
- 4) Modelo do fator de forma constante (SPURR 1952)  

$$V = b_1d^2h$$
- 5) Modelo da variável combinada (SPURR 1952)  

$$V = b_0 + b_1d^2h$$

onde, V, d e h conforme já definidos, e  $b_0, b_1, \dots, b_5$  são estimativas dos parâmetros para cada equação.

Na análise de regressão linear, o ajuste ou solução das funções matemáticas foi obtido pelo método dos mínimos quadrados. Os pacotes de programação Statistical Analysis System – SAS, versão 1979 (SAS 1979) e Sistema de Análise Estatística – SAEST (PIMENTEL; PANIAGO & COSTA 1981) foram utilizados em todo o processamento necessário dos dados, análise estatística e elaboração de gráficos.

O ajuste de cada equação aos dados foi analisado comparativamente, observando-se os seguintes critérios:

- a) exame do quadro da análise de variância (MQ Resíduo, teste F para o modelo completo e teste F sequencial)
- b) erro padrão da estimativa (s e s%)
- c) coeficiente de determinação ( $r^2$ ) ou coeficiente de múltipla determinação ( $R^2$ )
- d) amplitude de distribuição gráfica dos resíduos volumétricos percentuais ( $e_i\%$ ) em função do diâmetro à altura do peito das árvores-amostra.

Estas características foram avaliadas de tal modo que fosse possível atender um único objetivo: a equação de regressão selecionada dentre os modelos incluídos na análise deveria permitir a obtenção de estimativas precisas do volume de árvores com qualquer d e h (respeitados os limites da amostragem), e apresentar desvios mínimos, previsíveis e independentes da dimensão das árvores.

#### 4. RESULTADOS E DISCUSSÃO

Após o ajuste das funções ao conjunto de dados, obteve-se a definição matemática dos modelos como segue:

- 1) Modelo de Naslund  

$$V = 3,6915d^2 + 0,2705d^2h - 0,0005dh^2 - 0,0001h^2$$
- 2) Modelo compreensivo  

$$V = -0,1655 - 0,1414d + 9,3493d^2 - 0,1651dh + 0,3651d^2h + 0,0240h$$
- 3) Modelo australiano  

$$V = 0,0330 + 3,8935d^2 + 0,2536d^2h - 0,0046h$$
- 4) Modelo do fator de forma constante  

$$V = 0,3879d^2h$$

- 5) Modelo da variável combinada  
 $V = 0,0040 + 0,3849d^2h$

Na Tabela 1 são apresentadas as características do ajuste de cada equação. A fim de indicar a equação de volume mais adequada aos dados, uma análise e interpretação dos critérios de seleção fizeram-se necessárias.

#### 4.1 **Análise dos critérios de seleção dos modelos**

##### 4.1.1 **Exame da análise de variância**

Os testes de hipótese para os modelos completos (Tabela 1) indicaram que todas as cinco funções de volume foram altamente significativas (teste F significativo ao nível  $\alpha = 1\%$ ) ou seja: existe uma associação (relação probabilística) entre volume e o conjunto específico de variáveis independentes para cada modelo. A existência dessa relação, no entanto, não permite afirmar que estimativas úteis podem, necessariamente, ser obtidas com qualquer uma das equações (NETER & WASSERMAN 1974).

Examinando a análise de variância para cada modelo, constatou-se que algumas variáveis independentes não apresentaram contribuição significativa quando acrescidas sequencialmente às equações (teste F sequencial). Este fato ocorreu com dh e  $d^2h$  no modelo compreensivo e com h no modelo Australiano. As estimativas de interseção ( $b_0$ ) para os modelos Compreensivo e Australiano e para a equação da Variável Combinada foram igualmente não significativas, o que elimina a necessidade deste termo naqueles modelos.

Desta forma, apesar da significância identificada para todos os modelos completos, somente algumas variáveis em cada equação tiveram contribuição efetivamente positiva para explicar alguma parcela da variabilidade em volume. A constatação deste fato, ausente da discussão dos resultados em diversos trabalhos de volumetria registrados em literatura, sugere que a Análise de Regressão Passo a Passo talvez seja, atualmente, uma melhor alternativa para a seleção de equações de volume.

##### 4.1.2 **Erro padrão da estimativa**

O erro padrão da estimativa ou erro padrão residual ( $s$ ,  $s_{Y \cdot X}$  ou EPR) é uma medida absoluta da dispersão dos resíduos de regressão. A sua determinação permite definir intervalos de confiança para os parâmetros de um modelo ( $\beta$ 's) e para as estimativas dos valores ajustados da variável dependente.

O exame desta estatística (DRAPER & SMITH 1966) indica que quanto menor o seu valor numérico, mais precisas serão as inferências elaboradas a partir de um determinado modelo. O erro padrão da estimativa pode ser também expresso como uma porcentagem da média aritmética da variável dependente ( $s\%$ ) e, neste caso, pode ser interpretado de forma análoga ao coeficiente de variação.

A variação observada para o erro padrão da estimativa ( $s$ ) dos modelos incluídos neste estudo foi considerada aceitável e não se justificou identificar a superioridade de algum modelo. Entre os valores extremos de  $s\%$  constatou-se uma diferença de 5,33%. Entretanto, devido à elevada magnitude dos valores calculados para este critério nas cinco equações, julgou-se conveniente analisar as demais características das equações para melhor avaliar o seu ajuste.

2 **TABELA 1** Características dos modelos de volumetria analisados

Modelo	Características dos modelos						
	SQ <sub>Res.</sub>	F	r <sup>2</sup>	R <sup>2</sup>	s	s%	e <sub>i</sub> (%) vs d <sub>i</sub>
1. Naslund	0,1707	1137,08**	—	0,9776	0,0468	25,74	± 60
2. Compreensivo	0,1359	532,64**	—	0,9726	0,0426	23,44	± 100
3. Australiano	0,1767	694,76**	—	0,9644	0,0479	26,36	± 80
4. Fator de forma constante	0,2133	2756,31**	0,9718	—	0,0519	28,56	± 30
5. Variável combinada	0,2145	1747,72**	0,9568	—	0,0521	28,67	± 30

SQ<sub>Res.</sub> = Somatório do quadrado dos resíduos de regressão

F = Valor de F calculado para o teste de hipótese da nulidade com o modelo (\*\* significativo para α = 1%)

r<sup>2</sup> = Coeficiente de determinação (R<sup>2</sup> = Coeficiente de múltipla determinação)

s = Erro padrão da estimativa (s =  $\sqrt{MQ_{Resíduo}}$ )

s% = Erro padrão da estimativa expresso em % (s% =  $\frac{\sqrt{MQ_{Resíduo}}}{\bar{V}} \times 100$ )

e<sub>i</sub>% vs d<sub>i</sub> = Amplitude de distribuição dos resíduos volumétricos (%) em função dos diâmetros

e<sub>i</sub>% =  $\frac{V_i - \hat{V}_i}{V_i} \times 100$  onde V<sub>i</sub> = volume observado (cubado) e  $\hat{V}_i$  = volume estimado

i = 1, ..., 81

### 4.1.3 Coeficiente de determinação ( $r^2$ ) e coeficiente de múltipla determinação ( $R^2$ )

Os valores encontrados para estas medidas descritivas de desempenho e ajuste foram igualmente elevados para todas as equações (Tabela 1). Pequenas discrepâncias não foram consideradas como uma limitação, pois a magnitude de  $r^2$  ou  $R^2$  não deve ser analisada, comparativamente, de forma rígida e radical.

Em teoria, quanto maior for o valor numérico de  $R^2$  (ou  $r^2$ , em regressão linear simples, como é o caso das equações 4 e 5), maior será a proporção da variabilidade observada na variável dependente (volume) que está sendo explicada com o modelo.

Segundo KOZAK (1976), entretanto, o coeficiente de determinação ( $r^2$  ou  $R^2$ ) é um critério descritivo muito útil em Análise de Regressão, mas tem as suas limitações e aplica-se somente aos dados (amostra) para os quais o modelo de regressão foi ajustado. Desta forma (BARRET 1974), ao interpretar um coeficiente de determinação, deve-se sempre considerar que a sua utilidade é parcial. O uso de intervalos de confiança para as estimativas pode ter maior significado descritivo e prático. Adicionalmente, a elaboração de gráficos conforme proposto por DRAPER & SMITH (1966), ANSCOMBE (1973) e NETER & WASSERMAN (1974), permite observar características complementares sobre o ajustamento de um modelo.

### 4.1.4 Amplitude de distribuição gráfica dos resíduos

O exame visual de gráficos é uma opção informal e subjetiva para analisar a distribuição de resíduos de regressão, embora, em muitas investigações, seja suficiente para uma adequada avaliação. Por este motivo, neste estudo, omitiu-se o uso de procedimentos analíticos (NETER & WASSERMANN 1974, BRENA et al. 1978) que permitiriam reportar resultados numéricos da análise de resíduos.

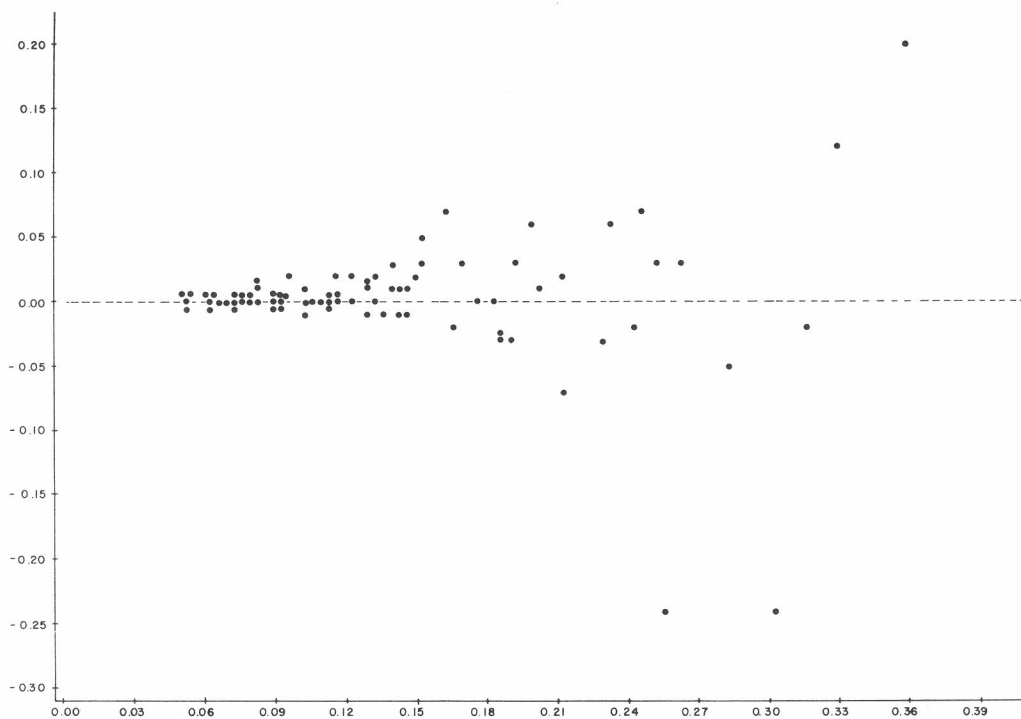
Evidências de heteroscedasticidade foram observadas para todos os modelos analisados. A Figura 1 ilustra o fato utilizando os resíduos volumétricos obtidos com a equação de Fator de Forma Constante: a variância dos resíduos, e assim também aquela da distribuição dos volumes observados, não é homogênea ao longo da amplitude de distribuição de DAP na amostra. Na realidade isto será praticamente sempre verdadeiro em estudos de volumetria: árvores pequenas têm menor variabilidade absoluta em volume que árvores com grandes dimensões.

Segundo NETER & WASSERMAN (1974), entretanto, a normalização da distribuição dos resíduos, que pode ser obtida por meio de uso de logaritmos nas variáveis originais, somente será desejável se o interesse for a determinação de intervalos de confiança para as estimativas. Caso a logaritmização não seja aplicada, as estimativas dos parâmetros do modelo serão de qualquer modo não-tendenciosas (unbiased), mas não serão aquelas de menor variância.

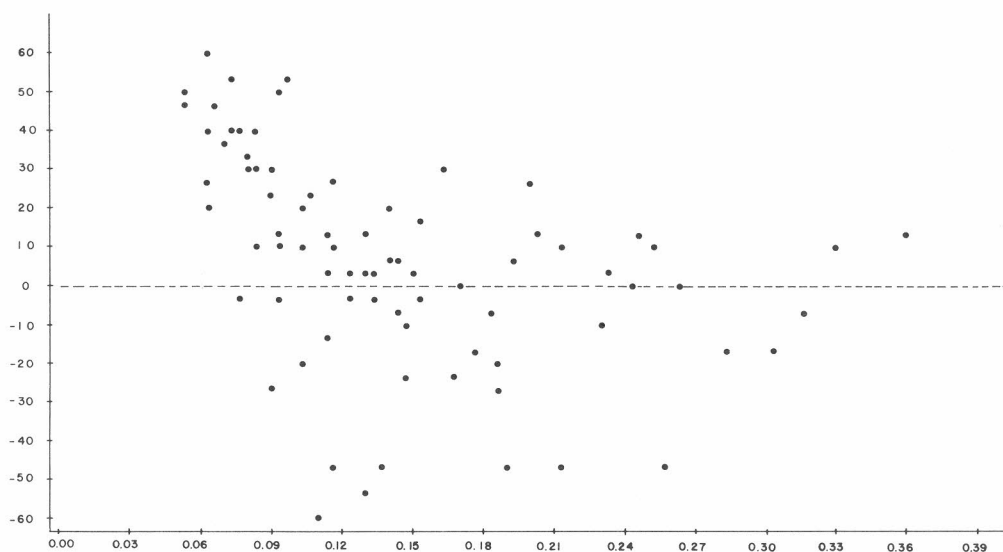
Neste estudo considerou-se suficiente avaliar a distribuição dos resíduos volumétricos percentuais (exame da variabilidade relativa), sem que tenha havido qualquer tentativa em se aplicar transformações ou análise de regressão ponderada aos dados. Esta decisão fundamentou-se na natureza preliminar e exploratória da investigação.

Os gráficos de distribuição dos resíduos percentuais (Figuras 2 a 6) complementam, desta forma, as informações necessárias para uma avaliação comparativa das equações ajustadas.

Os modelos de Naslund, Compreensivo e Australiano apresentaram amplitudes aproximadas de dispersão dos resíduos da ordem de 60, 80 e 100% respectivamente. Em adição à esta extrema variação dos resíduos, a sua distribuição sugere concentrações, o que indica que os modelos na Naslund (Figura 2) e Australiano (Figura 4) tiveram tendência a subestimar volumes para árvores com pequeno DAP, enquanto que, para árvores com as mesmas dimensões, superestima-

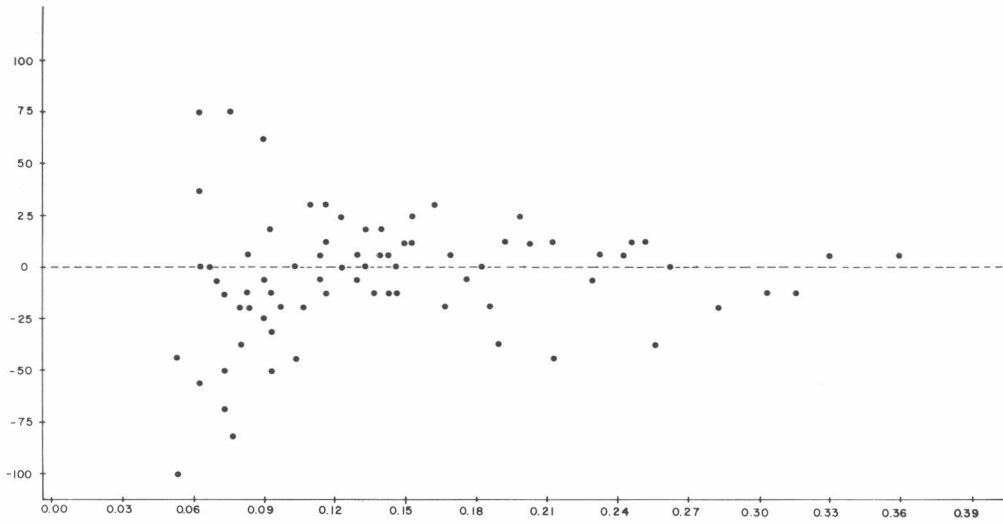


**FIG. 1** Distribuição dos resíduos volumétricos obtidos com a equação do fator de forma constante, equação 4, em função da amplitude observada de DAP.

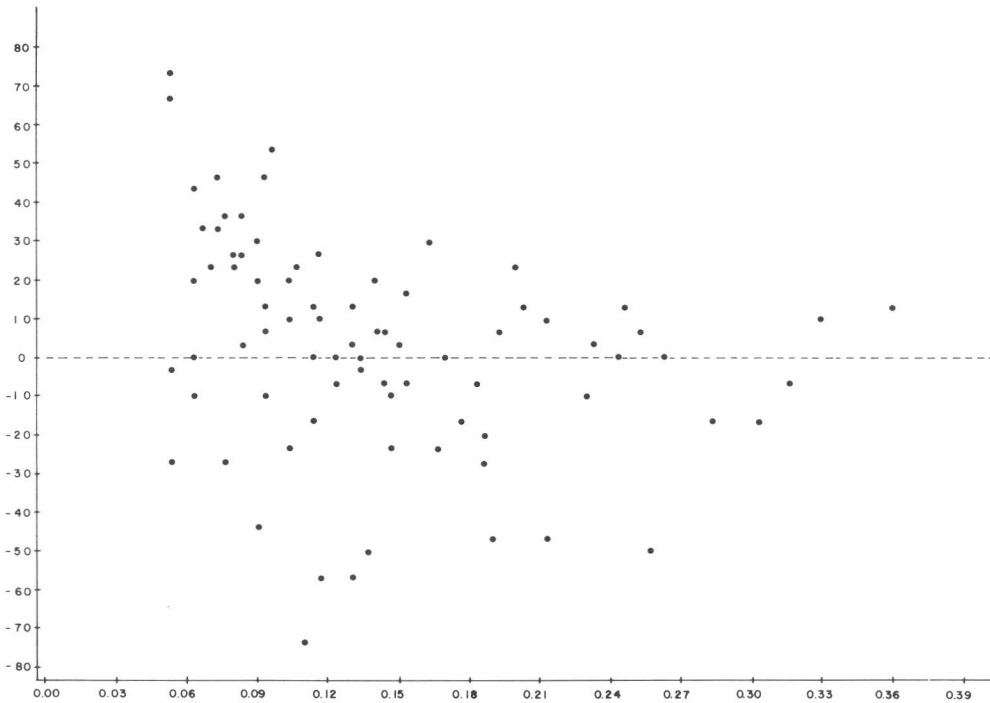


**FIG. 2** Distribuição dos resíduos volumétricos percentuais obtidos com o modelo de Naslund, equação 1, em função da amplitude observada de DAP.

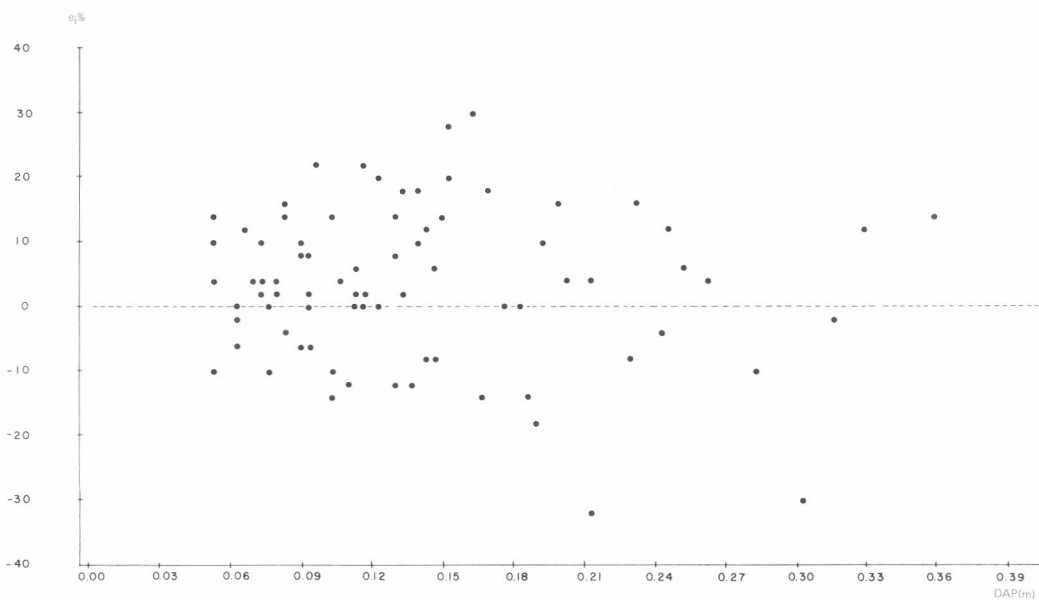




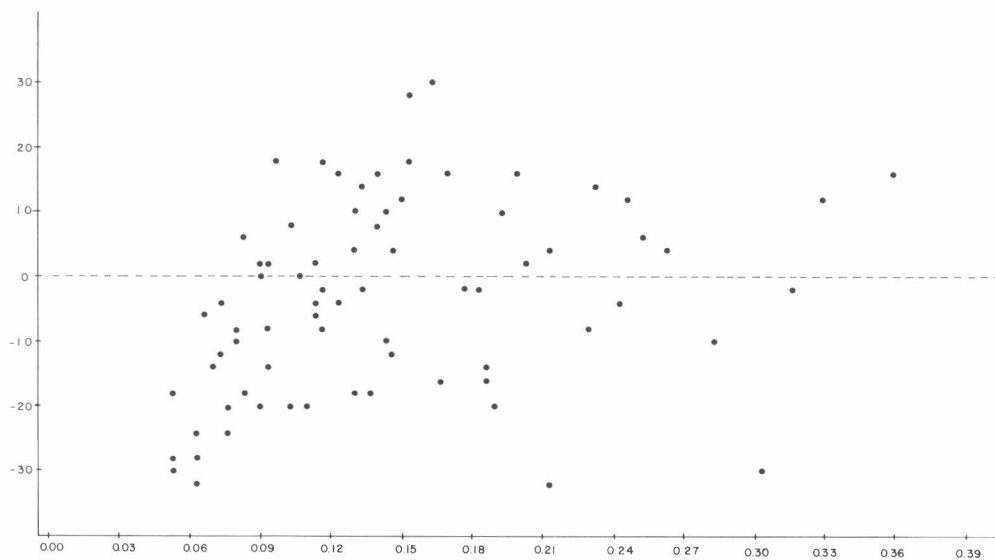
**FIG. 3** Distribuição dos resíduos volumétricos percentuais obtidos com o modelo Compreensivo, equação 2, em função da amplitude observada de DAP.



**FIG. 4** Distribuição dos resíduos volumétricos percentuais obtidos com o modelo Australiano, equação 3, em função da amplitude observada de DAP.



**FIG. 5** Distribuição dos resíduos volumétricos percentuais obtidos com a equação do fator de forma constante, equação 4, em função da amplitude observada de DAP.



**FIG. 6** Distribuição dos resíduos volumétricos percentuais obtidos com a equação variável combinada, equação 5, em função da amplitude observada de DAP.

tivas foram obtidas com o modelo Compreensivo (Figura 3). Uma redução considerável na magnitude da dispersão dos resíduos foi observada com as equações do Fator de Forma Constante (Figura 5) e da Variável Combinada (Figura 6). Ajustando-se estes dois modelos aos dados, foram definidas estimativas com desvios de até 30% em relação aos volumes cubados, sendo a equação do Fator de Forma Constante aquela que apresentou melhor consistência na distribuição dos resíduos ao longo da amplitude observada DAP.

#### 4.2 Seleção do melhor modelo

Todas as equações de volume foram altamente significativas e apresentaram elevado coeficiente de determinação, explicando, significativamente, desta forma, uma alta proporção da variabilidade observada do volume das árvores-amostra.

Aquelas características, no entanto, não amenizam os efeitos que os elevados valores para erro padrão da estimativa ( $s$  e  $s\%$ ) e para a manitude da dispersão dos resíduos podem ter sobre as estimativas volumétricas das árvores amostradas. Adicionalmente, a constatação de tendências nas estimativas obtidas com os modelos de Naslund, Compreensivo e Australiano será de grande influência no uso prático daquelas equações. Ressalte-se que os modelos de regressão foram exigidos a condensar e explicar a extrema variabilidade de volume que caracterizou a amostragem e isto, conforme evidenciado, nem sempre pode ser efetivado a contento (AHRENS 1981).

Desta forma, todas as funções analisadas foram inadequadas aos propósitos descritivos da investigação. Entretanto, em vista das características para cada modelo, e, sendo uma escolha necessária, pode-se afirmar que a equação do Fator de Forma constante ( $V = 0,3879d^2h$ ) foi a "menos inadequada" ao material experimental utilizado neste estudo.

#### 4.3 O uso de técnicas de seleção de variáveis

Analisando-se as equações de volume sugeridas em literatura, observa-se, na grande maioria das funções, que a expressão probabilística de volume ( $V$ ) como uma função do diâmetro à altura do peito ( $d$ ) e altura total de uma árvore ( $h$ ) implica em se utilizar o termo independente do modelo como sendo uma composição linear e aditiva de  $d$ ,  $h$ , e combinações ou transformações destas variáveis. Assim, a alternância, a natureza da combinação e a presença destas novas variáveis criadas a partir das variáveis observadas ( $d$ ,  $h$ ) são os elementos básicos de diferenciação entre os diversos modelos.

Por outro lado, com a disponibilidade de equipamentos para processamento eletrônico de dados, atualmente é possível aumentar a eficiência no desenvolvimento ou construção de modelos descritivos. DRAPER & SMITH (1966), NETER & WASSERMAN (1974) e KLEINBAUM & KUPPER (1979) descrevem alguns métodos extremamente úteis na seleção de variáveis. Dentre aqueles procedimentos, o uso de regressão passo a passo (stepwise regression analysis) permite avaliar a contribuição de cada variável quando incorporada ao modelo em desenvolvimento. Ou seja, pode-se "construir" um modelo matemático descritivo adequado aos propósitos de uma investigação específica. Desta maneira, quando aplicadas e interpretadas corretamente, as técnicas de seleção de variáveis e, em especial, a análise de regressão passo a passo, podem substituir, com muitos e importantes benefícios, o ajuste de funções pré-concebidas.

## 5. CONCLUSÕES E RECOMENDAÇÕES

Após a análise do ajustamento dos cinco modelos incluídos no estudo, a equação do Fator de Forma Constante ( $V = 0,3879d^2h$ ) foi selecionada como aquela de melhor desempenho. Esta não pode ser considerada, entretanto, uma boa equação para a estimativa de volume comercial de madeira de bracatinga. Apesar do elevado coeficiente de determinação ( $r^2 = 0,9718$ ), o erro padrão da estimativa expresso como uma porcentagem da média (coeficiente de variação) foi excessivamente alto ( $s\% = 28,56\%$ ). A elevada magnitude da dispersão dos resíduos percentuais de regressão ( $\pm 30\%$ ) não pode ser também desconsiderada, embora a sua consistência tenha sido sugerida.

Acredita-se que um melhor ajuste será obtido não necessariamente com um outro modelo de regressão, mas aplicando-se uma estratificação aos dados. O material experimental utilizado neste estudo apresentou uma extrema variabilidade de sítios, forma de implantação dos povoamentos, idade e dimensões (DAP e h) das árvores-amostra. Estes fatores podem ter influenciado a forma dos troncos, assim como contribuído para uma variabilidade no número de ramificações das árvores, o que por certo, reflete-se no volume comercial por árvore.

Adicionalmente, o maior valor comercial da espécie não está no seu volume individual mas no volume total por unidade de área. Assim, uma equação de volume por árvore terá utilidade somente para fins de pesquisa, quando estimativas precisas do volume são necessárias em estudos de espaçamento, fertilização e de competição entre espécies. Desta forma, futuros trabalhos de volumetria para bracatinga devem ser direcionados para duas principais alternativas:

- desenvolvimento de uma equação do volume por árvore, para uso em pesquisa e experimentação, envolvendo a amostragem com uma dispersão de idade entre quatro e dez anos e desenvolvidas a partir de mudas plantadas, e
- elaboração de um modelo para estimativa de volume comercial por unidade de área (metros estere/ha), para uso em empresas, descrevendo a relação entre volume comercial por hectare e, a área basal dos povoamentos ( $m^2/ha$ ) e sua altura média. Estudos desta natureza devem considerar as práticas de corte comercial em uso corrente.

Esta investigação permitiu também avaliar a utilidade de alguns critérios de uso corrente na seleção de modelos. Constatou-se que a significância estatística e um elevado coeficiente de determinação não são condições suficientes para indicar qualidades de ajustamento de um modelo de regressão. Embora a análise da distribuição gráfica de resíduos de regressão tenha caráter subjetivo, é um complemento útil no processo de seleção de equações e o seu uso deve ser intensificado em experimentação florestal.

Futuras investigações em seleção de equações para volumetria devem considerar também o desenvolvimento de equações por meio de Análise de Regressão Passo a Passo, bem como fazer uso das técnicas de Análise de Regressão Ponderada, quando conveniente.

## 6. REFERÊNCIAS

AHRENS, **A mathematical expression of stem form and volume for loblolly pine in southern Brazil**. Stillwater, Oklahoma State University, 1980. 59p. Tese Mestrado.

- \_\_\_\_\_. Um modelo matemático para volumetria comercial de bracatinga (**Mimosa scabrella** Benth.). In: SEMINÁRIO SOBRE ATUALIDADES E PERSPECTIVAS FLORESTAIS, 4., Bracatinga uma alternativa para reflorestamento, Curitiba, Jul. 1981. **Anais**. Curitiba, EMBRAPA/URPFCS, 1981. p.77-89 (EMBRAPA/URPFCS. Documentos, 5).
- ANSCOMBE, F. J. Graphs in statistical analysis. **The American Statistician**, Washington, DC, **27**(1):17-21, 1973.
- BRENA, D. A.; SILVA, J. N. M. & SCHNEIDER, P. R. Metodologia para verificação das condicionantes da análise de regressão. **Floresta**, Curitiba, **9**(2):25-45, 1978.
- COUTO, H. T. Z. do. Tabelas de volume para brotação de touças de **Eucalyptus saligna** Smith. **IPEF**, Piracicaba, (15):117-21, 1977.
- DRAPER, N. R. & SMITH, H. **Applied regression analysis**. New York, J. Wiley, 1966. 407p.
- FURNIVAL, G. M. An index for comparing equations used in constructing volume tables. **Forest Science**, Washington, DC, **7**(4):337-41, 1961.
- HONER, T. G. A new total cubic foot volume function. **The Forestry Chronicle**, Ottawa, **41**(4): 476-93, 1965.
- HUSCH, B. **Forest mensuration and statistics**. New York, Ronald Press, 1963. 474p.
- HUSCH, B.; MILLER, C. H. & BEERS, F. W. **Forest mensuration**. New York, J. Wiley, 1972. 410p.
- KLEINBAUM, D. G. & KUPPER, L. K. **Applied regression analysis and other multivariable methods**. North Scituate, Duxbury Press, 1978. 556p.
- KOZAK, A. **Notes on regression and correlation analysis as important tools in forest research**. s.n.t. Trabalho apresentado no XVI Congresso Mundial da IUFRO, Oslo, Noruega, Jun. 1976.
- LOETSCH, F.; ZOHRER, F. & HALLER, K. E. **Forest inventory**. Munique, BLV-Verlagsgesellschaft, 1973. v.2., 469p.
- MACHADO, S. do A. Tabela de volume para **Pinus taeda** na região de Telêmaco Borba, PR. **Floresta**, Curitiba, **10**(1):29-35, 1979.
- MEYER, W. H. A method of volume diameter ratios for board foot volume tables. **Journal of Forestry**, Washington, DC, **42**:185-94, 1944.
- NETER, J. & WASSERMAN, W. **Applied linear statistical models**; regression, analysis of variance and experimental designs. Homewood, Irwin, 1974. 842p.

- PAULA NETO, F. de; BRANDI, R. M.; RIBEIRO, J. C. & GUIMARÃES, D. P. Teste de aplicação de tabelas volumétricas para estimar a produção de plantações de **Eucalyptus paniculata** Sm., na região de Ipatinga, Minas Gerais. **Revista Árvore**, Viçosa, **1**(2):154-66, 1977.
- PIMENTEL, G. M.; PANIAGO, C. F. A. & COSTA, F. O. **Um sistema de análise estatística para microcomputadores**. Brasília, Centro de Computação da EMBRAPA, 1981. 29p. (Trabalho apresentado no XIV Congresso Nacional de Informática, Out. 1981).
- SAS INSTITUTE. **SAS user's guide**. Raleigh, 1979. 494p.
- SILVA, J. A. A. da; PAULA NETO, F.; BRANDI, R. M. & CONDÊ, A. R. Análise de modelos volumétricos para a construção das tabelas de volume comercial de **Eucalyptus** spp., segundo a espécie, região e os métodos de regeneração. **Revista Árvore**, Viçosa, **2**(1):86-99, 1978.
- SIQUEIRA, J. D. P. Tabelas de volume para povoamentos nativos de **Araucaria angustifolia** (Bert) O. Ktze, no Sul do Brasil. **Floresta**, Curitiba, **7**(1):7-12, 1977.
- SPURR, S. H. **Forest inventory**. New York, Ronald Press, 1952. 476p.
- STOATE, I. N. The use of a volume equation in pine stands. **Australian Forestry**, Camberra, **9**:48-52, 1945.
- VEIGA, R. A. de A. Comparação de equações de volume para **Eucalyptus saligna** Smith. I-equações aritméticas não formais. **Floresta**, **4**(3):81-97, 1973.

EQUACAO DE NASLUND - 1

21 23 FRIDAY, JULY 3, 1981 17

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE VTOT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR F	R-SQUARE	C.V.
MUDEL	3	7.46393002	2.48797667	1137.08	0.0001	0.977645	25.7410
ERROR	78	0.17066772	0.00218805		STD DEV		VTOT MEAN
UNCORRECTED TOTAL	81	7.63459774			0.04677657		0.18171989

SOURCE	DF	TYPE I SS	F VALUE	PR F	DF	TYPE IV SS	F VALUE	PR F
D2	1	7.29165126	3332.49	0.0001	1	0.04005642	18.31	0.0001
D2H	1	0.14716968	67.26	0.0001	1	0.12986725	59.35	0.0001
H2	1	0.02510908	11.48	0.0011	1	0.02510908	11.48	0.0011

PARAMETER	ESTIMATE	T FOR H0 PARAMETER 0	PR T	STD ERROR OF ESTIMATE
D2	3.69153959	4.28	0.0001	0.86277955
D2H	0.27054774	7.70	0.0001	0.03511741
H2	-0.00016463	-3.39	0.0011	0.00004860

MODELO COMPRENSIVO - 2

21 23 FRIDAY, JULY 3, 1981 . 22

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE VTOT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR F	R-SQUARE	C.V.
MODEL	5	4.82395703	0.96479141	532.64	0.0001	0.972610	23.4205
ERRJR	75	0.13584915	0.00181132		STD DEV		VTOT MEAN
CORRECTED TOTAL	80	4.95980618			0.04255963		0.18171989

SOURCE	DF	TYPE I SS	F VALUE	PR F	DF	TYPE IV SS	F VALUE	PR F
DAP	1	4.22216519	2330.99	0.0001	1	0.00002967	0.02	0.8985
D2	1	0.57592056	317.96	0.0001	1	0.01092097	6.03	0.0164
DH	1	0.00966793	5.34	0.0236	1	0.00934560	5.16	0.0260
D2H	1	0.00000022	0.00	0.9912	1	0.00656484	3.62	0.0608
HTOT	1	0.01620314	8.95	0.0038	1	0.01620314	8.95	0.0038

PARAMETER	ESTIMATE	T FOR HO PARAMETER 0	PR T	STD ERROR OF ESTIMATE
INTERCEPT	-0.16554345	-1.77	0.0805	0.09342950
DAP	-0.14444503	-0.13	0.8985	1.12851081
D2	9.34927602	2.46	0.0164	3.80754616
DH	-0.16508073	-2.27	0.0260	0.07267595
D2H	0.36512881	1.90	0.0608	0.19179261
HTOT	0.02400833	2.99	0.0033	0.00802715



MODELO AUSTRALIANO - 3

21 23 FRIDAY, JULY 3, 1981 27

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE VTOT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR F	R-SQUARE	C.V.
MODEL	3	4.78310294	1.59436765	694.76	0.0001	0.964373	26.3618
ERROR	77	0.17670324	0.00229485		STD DEV		VTOT MEAN
CORRECTED TOTAL	80	4.95980618			0.04790456		0.18171989

SOURCE	DF	TYPE I SS	F VALUE	PR F	DF	TYPE IV SS	F VALUE	PR F
D2	1	4.72686179	2059.77	0.0001	1	0.02696854	11.75	0.0010
D2H	1	0.04893222	21.32	0.0001	1	0.05612060	24.46	0.0001
HTOT	1	0.00730893	3.18	0.0783	1	0.00730893	3.18	0.0783

PARAMETER	ESTIMATE	T FOR HO PARAMETER 0	PR T	STD ERROR OF ESTIMATE
INTERCEPT	0.03305684	1.02	0.3102	0.03235730
D2	3.89354059	3.43	0.0010	1.13577653
D2H	0.25356459	4.95	0.0001	0.05127484
HTOT	-0.00459615	-1.78	0.0783	0.00257539



EQUACAO DA VARIÁVEL COMBINADA - 5

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE VTUT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR F	R-SQUARE	C.V.
MODEL	1	4.74531020	4.74531020	1747.72	0.0001	0.956753	28.6743
ERROR	79	0.21449598	0.00271514			STD DEV	VTUT MEAN
CORRECTED TOTAL	80	4.95980618			0.05210700		0.18171989

SOURCE	DF	TYPE I SS	F VALUE	PR F	DF	TYPE IV SS	F VALUE	PR F
D2H	1	4.74531020	1747.72	0.0001	1	4.74531020	1747.72	0.0001

PARAMETER	ESTIMATE	T FOR HO PARAMETER 0	PR T	STD ERROR OF ESTIMATE
INTERCEPT	0.00400145	0.56	0.5790	0.00718273
D2H	0.38486334	41.81	0.0001	0.00920599