

**MODELAGEM DO CRESCIMENTO
FLORESTAL POR MEIO DE EQUAÇÕES
DIFERENCIAIS**

BRENO RODRIGUES MENDES

2005

BRENO RODRIGUES MENDES

DESCARTADO



ASSINATURA

Data 26 / 09 / 17

BIBLIOTECA UNIVERSITÁRIA
UFLA

**MODELAGEM DO CRESCIMENTO
FLORESTAL POR MEIO DE EQUAÇÕES DIFERENCIAIS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Mestrado em Engenharia Florestal, área de concentração em Floresta de Produção, para obtenção do título de "Mestre".

Orientador
Prof. Natalino Calegario



LAVRAS
MINAS GERAIS – BRASIL
2005

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Mendes, Breno Rodrigues

Modelagem do crescimento florestal por meio de equações diferenciais
/ Breno Rodrigues Mendes. -- Lavras : UFLA, 2005.

43 p. : il.

Orientador: Natalino Calegario.
Dissertação (Mestrado) – UFLA.
Bibliografia.

1. Modelo não-linear. 2. Crescimento florestal. 3. Equação diferencial.
4. Árvore individual. 5. Área basal. I. Universidade Federal de Lavras. II.
Título.

CDD-634.9285

BRENO RODRIGUES MENDES

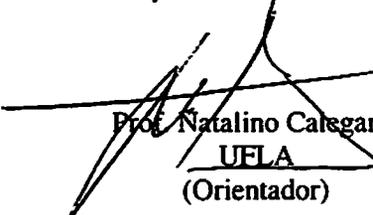
**MODELAGEM DO CRESCIMENTO
FLORESTAL POR MEIO DE EQUAÇÕES DIFERENCIAIS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Mestrado em Engenharia Florestal, área de concentração em Floresta de Produção, para obtenção do título de “Mestre”.

APROVADA em 22 de fevereiro de 2005

Prof. Jose Marcio Mello UFLA

Prof. Carlos Eduardo Silva Volpato UFLA



Prof. Natalino Categario
UFLA
(Orientador)

LAVRAS
MINAS GERAIS-BRASIL

A Deus,

por estar comigo em cada momento de minha vida;

OFEREÇO

*À minha querida mãe que esta ao lado de Deus, pelo seu infinito amor e por
tudo que sou.*

Ao meu pai, modelo de vida, amor e sabedoria.

Aos meus irmãos, pela grande admiração.

*Ao meu padrinho Luciano, a quem sigo os exemplos da vida pessoal e
profissional.*

*Aos meus amigos e familiares, pelo imenso prazer de viver este momento me
proporcionaram.*

DEDICO

AGRADECIMENTOS

À Universidade Federal de Lavras, à coordenação de Pós-Graduação e ao Departamento de Ciências Florestais pela oportunidade de realização do curso.

À CAPES, pela concessão da bolsa de estudos.

Ao Prof. Dr. Natalino Calegario, pelo incentivo, confiança, amizade, transferência de conhecimento, pelas sugestões na elaboração deste trabalho e harmonia na orientação desse trabalho. E por sua presença na definição da minha vida profissional.

Ao Prof. Dr. José Roberto Scolforo, pela amizade, pelo ensinamento durante a graduação e a pós-graduação e o apoio indispensável e irrestrito em todas as etapas do curso. Que sempre estarão guardados em minhas lembranças.

Aos professores do Departamento de Ciência Florestais, pelos ensinamentos e colaboração, em especial ao Prof. Luis Marcelo "Passarinho", pela amizade e atenção.

Aos funcionários do Departamento de Ciências Florestais, Teresinha, Rose, Patrícia e Gláucia, pelo auxílio, simpatia e dedicação.

Aos colegas de curso de graduação e pós-graduação pelo convívio durante estes períodos: Rodrigo, Cíntia, Gleyce, Gentil, "Claudão", Lucas, Gutenberg, Adauta, Elves, Lílian, Frederico, "Geninho", e Valdir.

Em especial, aos grandes amigos Sandro, Simone, Evandro, Edmilson, Maria Zélia, Marcela, "Totonho", Ivonise, Leonardo e Cláudio Thircher, pela ajuda, companheirismo, "cachaça" e colaboração em todos os nossos trabalhos.

À minha namorada "SOL", pelo companheirismo, incentivo e dedicação.

A todos aqueles que, de forma anônima, deram suas parcelas de contribuição, o que facilitou a concretização deste trabalho.

ÍNDICE DE TABELAS

TABELA 1 Coordenadas geográficas aproximadas e extensão da área da empresa Aracruz Celulose S.A.....	18
TABELA 2 Apresentação em síntese dos dados experimentais provenientes do inventário florestal contínuo, fornecido pela empresa Aracruz Celulose S.A ...	19
TABELA 3 Apresentação dos Modelos Completos obtidos a partir da decomposição dos parâmetros dos Modelos Originais	25
TABELA 4 Estatística comparativa entre os modelos: <i>Original I e Completo I</i> , seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual (EPR) de cada modelo avaliado.....	32
TABELA 5 Estatística comparativa entre os modelos: <i>Original II e Completo II</i> , seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual(EPR) de cada modelo avaliado.....	32
TABELA 6 Estatística comparativa entre os modelos: <i>Original III e Completo III</i> , seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual(EPR) de cada modelo avaliado	33

SUMÁRIO

ÍNDICE DE TABELAS	i
ÍNDICE DE FIGURAS	ii
RESUMO	iii
ABSTRACT	iv
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	3
2.1 Equações Diferenciais	3
2.2 Modelos Biométricos	6
2.3 Modelagem do Crescimento	7
2.4 Modelos de Regressão	12
3 MATERIAL E MÉTODOS	18
3.1 Dados	18
3.2 Desenvolvimento	19
3.2.1 Medidas de Acuracia	26
3.2.2 Análise de Resíduos	27
3.2.3 Análise gráfica dos resíduos	28
4 RESULTADOS E DISCUSSÃO	29
4.1 Análise individual dos dados	29
4.2 Estimativas dos parâmetros e medidas de acuracidade	31
4.3 Análise gráfica residual para os modelos	36
5 CONCLUSÃO	39
6 REFERÊNCIAS BIBLIOGRÁFICAS	40
ANEXOS	43

RESUMO

MENDES, Breno Rodrigues. **Modelagem do Crescimento Florestal por meio de Equações Diferenciais**. 2005. 44 p. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras¹.

Os modelos de crescimento são fundamentais para o planejamento das atividades florestais. Com o avanço tecnológico atual, existe um crescente interesse no desenvolvimento de modelos de crescimento por árvores individuais para expressar o crescimento em área basal. A área basal está diretamente correlacionada às práticas silviculturais, permitindo a otimização do sítio florestal. Assim, o principal objetivo deste trabalho foi o desenvolvimento de modelos não-lineares para árvores individuais, a partir de equações diferenciais, e a avaliação da qualidade de ajuste dos mesmos para expressar o crescimento em área basal. Os dados vieram da empresa Aracruz Celulose S/A, localizada na região costeira do Brasil, nos estados da Bahia e Espírito Santo e originaram-se de um inventário florestal contínuo de clones de *Eucalyptus* spp., onde suas respectivas medidas foram tomadas em um curto intervalo de tempo, proporcionando uma modelagem por árvores individuais, num total de 4683 observações individuais. O desenvolvimento dos modelos ocorreu a partir do trabalho de Ratkowsky (1989), que apresentou uma série de equações diferenciais de Bernoulli, das quais foram selecionadas possíveis equações que geram curvas côncavas para dar origem aos modelos de crescimento apresentados neste trabalho. A avaliação da precisão dos modelos foi verificada através da medida de acuracidade, erro padrão residual (EPR), e do gráfico de resíduo. Os resultados mostraram que o modelo completo com três parâmetros, desenvolvido a partir do modelo original com três parâmetros foi superior aos demais modelos desenvolvidos, devido à inclusão de novas variáveis, como: clone, altura total (HT), altura dominante (HD), diâmetro quadrático (DG), Área Basal (G), Índice de sítio (IS), Densidade (N) originando assim um novo modelo, denominado por *Modelo Completo III*. O ganho com precisão foi altamente significativo, quando comparado com os outros modelos. Conseqüentemente, este modelo gera informações com um alto grau de precisão e exatidão para o planejamento das empresas florestais.

¹Comitê de Orientação: Natalino Calegario (Orientador) – UFLA.

ABSTRACT

MENDES, Breno Rodrigues. **Modeling Forest Growth Based on differential Equations**. 2005. 44 p. Dissertação (Master in Forest Engineering) - Federal University of Lavras, Lavras I.

The growth models are important for planning of the forest activities. As consequent of the current technological advance, there is an increasing interest on the development of the growth models for individual trees, to express the growth in basal area. The basal area is directly correlated with silvicultural practices, allowing the optimization of the land uses. Thus, the main objective of this work was to develop nonlinear models for individual trees from differential equations and evaluate the fitting quality of them to express the growth in basal area. The stands used in this study are from the Aracruz Celulose S/A company, located in the costal region of Brazil, in the Bahia and Espírito Santo states. The database are from forest inventory of clonal Eucalyptus plantation, with measures taken in a short interval of time, providing individual trees modeling, with a total of 4683 observations. The model developments were based on exponential equations presented by Ratkowsky (1989) and converted in differential equations of Bernoulli class, which the solutions were used to growth models for this study. The evaluation model precise was done based on measurements of accuracy, Residual Standard Error (EPR), and graphics of residual. The results had shown that the complete model with 3 parameters, developed from the original model with 3 parameters, was superior and new variables were included on it to improve the precise, such as: clone, total height (HT), dominant height (HD), quadratic diameter (DG), basal area (G), site index (IS), Density (N). This model was considered a new model, called Complete Model III. The improvement of the precision was significantly when compared to the other models, facilitating the prediction and planning process by forest companies.

¹Guidance Committee: Natalino Calegario (Major Professor) – UFLA.

para expressar o crescimento. E apesar da existência de alguns poucos estudos na área florestal, os autores não se preocupam em desenvolver e mostrar a evolução da metodologia da modelagem do crescimento. Sendo que entre as diversas fases do desenvolvimento de modelos, a estimativa de parâmetros, a partir de dados experimentais, é de grande importância, pois a não disponibilidade de parâmetros confiáveis leva a resultados de simulação pouco realistas e até mesmo inúteis.

No setor florestal existe um crescente interesse no desenvolvimento de modelos de incremento em área basal, isto porque a área basal esta relacionada diretamente a práticas silviculturais, expressando o grau de densidade do povoamento permitindo a otimização da utilização do sítio florestal.

Hoje, as projeções feitas por especialistas do mercado *global*, indicam um considerável aumento no consumo de madeira para os próximos anos, principalmente no que se diz respeito à madeira destinada à serraria.

Diante desse fato, o principal objetivo deste estudo é o desenvolvimento de uma nova modelagem, que permita expressar o crescimento em área basal para povoamentos clonais de *Eucalyptus* spp, baseando-se em equações diferenciais, tornando possível, que o plano de suprimento das empresas florestais possa ser corretamente elaborado, produzindo assim, informações precisas e exatas para o planejamento florestal.

2 REFERENCIAL TEÓRICO

2.1 Equações Diferenciais

Atualmente o emprego de modelos matemáticos e estatísticos, na área biológica, tem aumentado consideravelmente em função do surgimento de novas técnicas, tanto experimentais, como de simulação. Tal fato se deu para facilitar o entendimento da complexidade dos processos biológicos de um modo geral. A dinâmica destes processos pode ser descrita matematicamente usando equações diferenciais.

As equações diferenciais são equações que envolvem derivadas de variáveis dependentes em relação a variáveis independentes. As técnicas de resolução de equações diferenciais são imprescindíveis para o trabalho com grandezas que variam, dependendo de outras que também estão mudando. Tornam-se, assim, instrumentos indispensáveis para quase todas as áreas do conhecimento, tanto na ciência pura como na aplicada.

Torres et.al. (2002), comentam que as equações diferenciais constituem uma das principais áreas da Matemática devido a sua imensa aplicabilidade em vários ramos da Ciência e da Engenharia, bem como pela sua riqueza teórica, onde se inserem várias das mais importantes idéias matemáticas, proporcionando contribuições importantes para a resolução de problemas que surgem no cotidiano de cada profissional.

Segundo Stamford (2004), chama-se de equação diferencial toda equação que relaciona uma função incógnita (função desconhecida) com uma ou mais das suas derivadas.

As derivadas definem quais são as variáveis dependentes e as variáveis independentes. Por exemplo, as notações $\frac{d^2y}{dx^2}$ e $\frac{dy}{dx}$ na equação $\frac{d^2y}{dx^2} = y \frac{dy}{dx}$ diz que nesta equação, y é a variável dependente e x é a variável independente.

Para esta outra equação $\frac{\partial^2 x(t,u)}{\partial t} = \frac{\partial^2 x(t,u)}{\partial u^2}$ as notações $\frac{\partial x}{\partial t}$ e $\frac{\partial^2 x}{\partial u^2}$ dizem que para esta equação, x é a variável exógena e u e t são as variáveis endógenas. Quando existir apenas uma variável independente como nestas equações diferenciais, denomina-se equação diferencial ordinária.

A ordem de uma equação diferencial é dada pela ordem da maior derivada da função desconhecida que aparece na equação. Como exemplo a equação $\frac{dT}{dt} = k(T - S)$ é uma equação de 1ª ordem e a equação $m \frac{d^2 x}{dt^2} = \frac{k}{x^2}$ é uma equação de 2ª ordem.

Luz et al (2001) ressalta a importância da aplicação de equações diferenciais ordinárias de primeira ordem, em vários segmentos da área biológica, e apresenta o seguinte exemplo em seu trabalho:

O crescimento de tumores, que experimentalmente foi observado que os microorganismos que se reproduzem de forma a ocorrer a “sua duplicação” (“mitose”), como as bactérias, tem sua taxa de crescimento proporcional ao volume de células divididas em um dado momento. Denotando por $V(t)$ o volume de células divididas no tempo (t).

Então

$$\frac{dV}{dt} = \lambda V$$

para alguma constante positiva λ . A solução é

$$V(t) = V_0 e^{\lambda(t-t_0)}$$

onde V_0 é o volume de células divididas no tempo inicial t_0 . Então o volume de células divididas cresce exponencialmente com o tempo, ou seja $V(t) \rightarrow \infty$ quando $t \rightarrow \infty$ quando, o que é impossível de ser mantido para sempre, temos,

então, um modelo de natureza razoável que tem melhor aplicabilidade em intervalos delimitados de tempo.

Por outro lado, o crescimento de tumores sólidos não é exponencial em relação ao tempo. Através de pesquisas verificou-se que uma boa aproximação de $V(t)$ que melhor se adequa aos dados obtidos da análise de vários tumores sólidos e dada pela equação

$$V(t) = V_0 \exp\left(\frac{\lambda}{\alpha}(1 - \exp(-\alpha t))\right) \quad (1)$$

onde $\exp(x) = e^x$, λ e α são constantes positivas. A equação (1) é conhecida como relação Gompertziana. A análise desta equação nos informa que o tumor cresce mais e mais lentamente com o passar do tempo e que o limite do volume de células divididas é aproximadamente:

$$V_0 e^{(\lambda/\alpha)}$$

Bronson (1993) apresenta a seguinte classificação para as equações diferenciais de 1ª ordem:

a) Forma diferencial: para uma equação diferencial de 1ª ordem diz-se na forma diferencial se está escrita como:

$$M(x, y)dx + N(x, y)dy = 0$$

b) Equação linear: chama-se equação linear de 1ª ordem a toda a equação diferencial que pode ser escrita na forma:

$$y' + p(x)y = q(x)$$

c) Equação Bernoulli: uma equação diferencial que pode ser escrita na forma:

$$y' + p(x)y = q(x)y^n$$

d) Equação homogeneizável: chama-se equação diferencial homogeneizável a toda a equação diferencial da forma:

$$f(tx, ty) = f(x, y)$$

e) Equação separável: chama-se equação diferencial separável ou de variáveis separáveis a toda a equação diferencial de 1ª ordem que pode ser escrita na forma:

$$g(y)dy + f(x)dx = 0$$

ou, ainda, na forma:

$$y' - \frac{f(x)}{g(y)}, \text{ com } g(y) \neq 0$$

Supondo-se f e g funções contínuas num dado intervalo.

f) Equação exata: uma equação diferencial de 1ª ordem do tipo:

$$M(x, y)dx + N(x, y)dy = 0$$

diz-se exata se $M(x, y)dx + N(x, y)dy$ é o diferencial total de alguma função $u(x, y)$, isto é, se existir alguma função $u(x, y)$, tal que:

$$du = M(x, y)dx + N(x, y)dy = 0$$

Muitos resultados já foram obtidos com a utilização de equações diferenciais ordinárias em vários segmentos da Ciência, pode-se perceber que na análise das aplicações, a teoria qualitativa das equações diferenciais ordinárias mostra-se um instrumento de grande importância, pois permite que se verifique previamente se a equação matemática utilizada para modelar o problema em questão realmente se adequa ao fenômeno descrito pelo modelo.

2.2 Modelos Biométricos

Para diversos estudos em todo mundo, está sendo utilizada uma tendência de modelagem do crescimento, que são os modelos biométricos, descritivos ou empíricos. Os modelos biométricos são essencialmente descritivos e identificados por equações matemáticas e estatísticas capazes de representar os dados experimentais de forma aceitável. Têm como propósito descrever o que foi observado experimentalmente ou coletado em um levantamento. Assim, este

tipo de modelo é resultado de um trabalho experimental e, geralmente, esse tipo de modelo é apresentado em gráficos nos quais são plotados os valores reais, os valores por ele estimados e os desvios (Pessoa et al., 1997).

Os modelos biométricos, baseados num grande número de observações feitas em parcelas permanentes, são caracterizados como a descrição das relações entre as variáveis envolvidas sem considerar, a princípio, os processos físicos e biológicos que governam o crescimento, embora modelos com comportamento biológico possam ser utilizados. Clutter et al. (1983) afirmam que a predição futura da produção de florestas envolve três diferentes abordagens: estimativa volumétrica corrente; estimativa da densidade corrente; e estimativa volumétrica futura. Em função disso, praticamente todos os florestais de renome da Alemanha do século XVIII e início do século XIX executaram estudos de produção baseados em observações do crescimento florestal, ou seja, dentro do conceito de modelos biométricos ou empíricos.

Segundo Burkhart (1999), o termo "empírico" normalmente é associado à imagem de uma metodologia expedita e de menor qualidade, porém, isto não representa a realidade. Os modelos empíricos não necessariamente são menos criativos, científicos, elegantes ou requerem menor capacidade intelectual. Possuem esta denominação, pois são funções ajustadas a dados observados em amostragem da população florestal de interesse, ou seja, **modelam** as "conseqüências" e não as "causas" dos processos fisiológicos.

2.3 Modelagem do Crescimento

Para se descrever a relação tamanho-idade em vegetais existe um grande número de modelos de crescimento disponíveis. Koops (1986) cita Bertalanffy, Gompertz, Richards ou Logística associado-os a modelos assintóticos de crescimento, com propriedades ligeiramente diferentes.

Calegario (2005) apresenta o modelo Logístico que é sigmoïdal e possui as características típicas para o ajuste do crescimento de qualquer variável biológica, ou seja, tem início na origem, apresenta acelerado ritmo inicial de crescimento, um ponto de inflexão onde tal ritmo decresce e, por fim, tende a um valor assintótico (Figura 1).

De acordo com Guimarães (1994), a realização da modelagem do crescimento em florestas, com condição necessária aos propósitos da prognose, deve ser baseada nos seguintes componentes:

- **base de dados:** os dados podem ser obtidos de parcelas permanentes, ou análise do tronco. Estas fontes devem captar as variações do crescimento ao longo do tempo;
- **variáveis envolvidas:** as variáveis selecionadas devem estar correlacionadas com o crescimento e serem de fácil obtenção. As variáveis mais utilizadas são: as idades atual e futura (I_1 e I_2), as áreas basais corrente e futura (B_1 e B_2), o índice de sítio (S), os números de árvores por hectare, atual e futuro (N_1 e N_2) e os volumes atual e futuro (V_1 e V_2);
- **formulação do modelo:** trata-se de uma explicação estatística para as variações contidas na base de dados. A complexidade do modelo está associada diretamente ao número de variáveis envolvidas e ao tipo de informação desejada. A formulação do modelo implica a elaboração de uma ou mais relações funcionais, que contêm as variáveis do povoamento e, ou, das árvores.

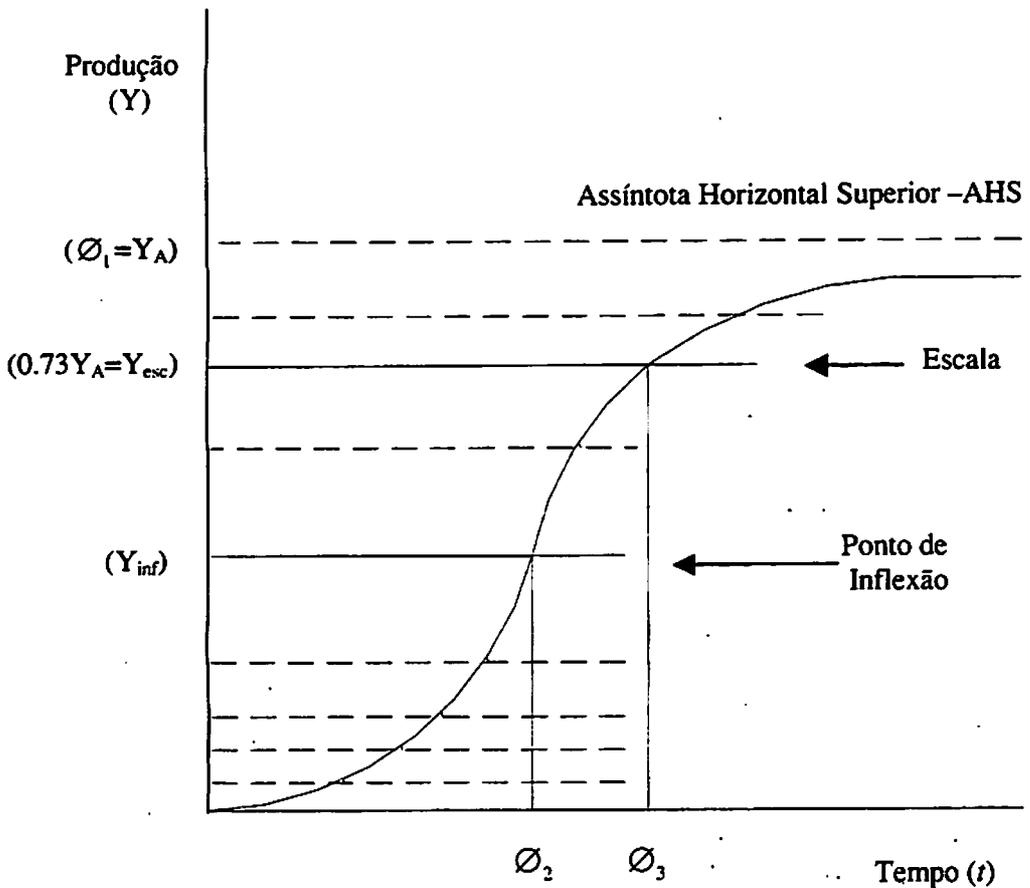


FIGURA 1 Modelo biológico (logístico) e o significado de seus coeficientes.

$$Y = \frac{\varnothing_1}{1 + \exp\left[\frac{(\varnothing_2 - t)}{\varnothing_3}\right]}$$

em que:

\varnothing_1 = Assíntota Horizontal Superior - AHS (mesma unidade de Y);

\varnothing_2 = Ponto de Inflexão (mesma unidade de t, ou seja, valor da idade quando a produção atingir $\varnothing_1/2$ e $Y_A/2$);

\emptyset_3 = Escala (mesma unidade de t , ou seja, diferença entre a idade onde a produção atingir $\approx 0.73 Y_A$ e a idade no ponto de inflexão).

Conforme Fitzhugh (1976), o termo curva de crescimento sugere uma curva sigmóide descrevendo uma seqüência de medidas de tamanho. Segundo ele, uma terminologia mais geral seriam curvas tamanho-idade.

A curva de crescimento de um organismo qualquer possui forma sigmoidal e apresenta três fases distintas: fase juvenil, que termina no ponto de máximo incremento corrente; fase de maturidade, que começa no ponto de máximo incremento corrente e termina onde o incremento corrente iguala-se ao máximo incremento médio; e fase de senescência, que começa a partir do máximo incremento médio (Husch et al., 1993).

Novamente, segundo Fitzhugh (1976), os objetivos primários para o ajuste de curvas de crescimento são descritivos, pois a informação contida em uma seqüência de pontos tamanho-idade está consolidada em relativamente poucos parâmetros e preditivos, pois os parâmetros de curvas de crescimento são utilizados separadamente ou em conjunto para prever taxas de crescimento, necessidades alimentares, respostas à seleção dentre outros fatores. Acrescenta, também, que os objetivos da análise determinam a escolha do método de ajustamento da curva de crescimento. Ele menciona que os algoritmos envolvendo interações são sensíveis à escolha dos valores iniciais e podem não convergirem para uma solução, além de ser possível encontrar estimativas dos parâmetros que sejam matematicamente corretas, mas absurdas biologicamente.

Os parâmetros de modelos de crescimento são geralmente estimados usando o procedimento iterativo do método dos mínimos quadrados generalizado (Marquardt, 1963). Para uma base de dados com poucos indivíduos requer-se o uso de algoritmos eficientes e computadores de alta velocidade. Estimativas iniciais precariamente determinadas podem aumentar o número de



iterações requeridas ou até impedir a convergência para uma solução razoável.

Os modelos de crescimento auxiliam pesquisadores e manejadores em diversas formas. Soares (1999) ressalta que na ciência florestal, os modelos de crescimento são utilizados para predição do crescimento dos povoamentos, fornecendo subsídios para a atualização de inventários passados, as prescrições silviculturais, a elaboração de planos de manejo e estudos de regulação florestal, dentre outros.

Segundo Scolforo (1998), o progresso das técnicas de computação, modelos de crescimento baseados em árvores individuais começam a ser mais implementados. A distinção básica entre este modelo e aquele por classe de diâmetro, é que este último fornece como resposta, informações por classe de diâmetro, a partir das informações de entrada (*inputs*), que são as estatísticas globais do povoamento.

Soares (1999) define os modelos de árvore individual como os modelos que simulam o crescimento de cada árvore e a probabilidade de mortalidade, fornecendo os resultados da produção em nível de povoamento e, ou, por classe de diâmetro.

Nos modelos para árvores individuais, o crescimento de cada árvore é projetado individualmente, a partir de equações apropriadas, e os resultados em relação ao povoamento são obtidos pela soma dos valores individuais (Bella, 1984).

Segundo Davis & Johnson (1987), os modelos para árvores individuais são os mais complexos. Estes modelos calculam os índices do grau de competição para cada árvore e utilizam na determinação do crescimento em diâmetro, altura e tamanho da copa.

Conforme a classificação apresentada por Munro (1974), citado em Scolforo (1998), os modelos para árvores individuais podem-se dividir em:

modelos dependentes da distância (espaçamento entre as árvores) e modelos independentes da distância.

Os modelos para árvores individuais de distância dependente utilizam índices de competição baseados na mensuração ou no mapeamento da distância de cada árvore em relação às árvores adjacentes, que estão competindo por luz, água e nutrientes (Davis & Johnson, 1987). O crescimento individual da árvore é simulado em função do seu tamanho, da qualidade do local e de uma medida de competição entre as árvores. O índice de competição varia de modelo para modelo, mas de modo geral é função do tamanho da árvore em relação ao tamanho e à distância das competidoras (Campos, 1997).

Nos modelos para árvores individuais de distância independente, o índice de competição é baseado na comparação das características de uma árvore (proporção da copa, DAP e altura total) em relação a todas as outras árvores em uma amostra, como, por exemplo, a razão entre o diâmetro da árvore e o diâmetro médio (Guimarães, 1994). Este tipo de modelo assume que todas as espécies e tamanhos de árvores estão uniformemente distribuídas no povoamento (Davis & Johnson, 1987).

2.4 Modelos de Regressão

Segundo Draper & Smith (1981), os modelos de regressão podem ser classificados em lineares, linearizáveis e não-lineares, em relação aos seus parâmetros.

Os modelos de regressão são lineares em relação aos parâmetros quando a primeira derivativa com relação aos seus parâmetros não é função dos seus parâmetros:

$$\frac{\partial f_i(x;\beta)}{\partial \beta_j}$$

para $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, p$.

Os modelos linearizáveis são aqueles que podem ser transformados em lineares através de anamorfose.

Os modelos de regressão são não-lineares em relação aos parâmetros, quando a primeira derivativa em relação aos parâmetros for uma função também dos parâmetros, ou seja:

$$\frac{\partial f_i(x; \beta)}{\partial \beta_j} = h(x, \beta)$$

pelo menos para algum i e j .

Medeiros (1999) comenta que Stevens (1951) estudou o seguinte modelo de regressão:

$$E(y_i) = \alpha + \beta * \rho^{x_i}, (i = 1, 2, \dots, n)$$

em que α , β , ρ são parâmetros com $0 < \rho < 1$ e x_i é o tempo. Ele ressalta o uso deste modelo em relações de crescimento, tanto animal como de plantas. Nesta equação, y_i tende assintoticamente para um certo limite, quando x_i tende para o infinito.

Quando o modelo é linear nos parâmetros, ele pode ser expresso como:

$$y = x\beta + \varepsilon$$

Na análise de regressão linear, os dois métodos de estimação mais empregados são o dos quadrados mínimos e o da máxima verossimilhança, sendo que conduzem aos mesmos estimadores, em presença de normalidade.

Quando os modelos de regressão são não-lineares, os métodos de estimação empregados também são, em geral, os de quadrados mínimos e de máxima verossimilhança. Entretanto, tais métodos aqui não conduzem a soluções explícitas para os estimadores.

Ao contrário dos modelos lineares, os estimadores são em geral soluções de sistemas de equações não lineares, necessitando-se de métodos numéricos

(iterativos) para sua resolução.

Machado (1980) usou médias de dados de crescimento ponderal e estimou os parâmetros da função logística, utilizando o método dos mínimos quadrados ordinário para chegar às equações normais e, em seguida, aplicando o método de Gauss-Newton, obteve a solução do sistema por meio de processo iterativo.

Draper & Smith (1981) apresentaram um procedimento para estimar os parâmetros em modelos não lineares, por meio do método dos mínimos quadrados ordinário, supondo o erro aleatório com média zero e variância σ^2 . O sistema de equações normais é então resolvido pelo método da linearização, "*Steepest Descent*" (gradiente) ou "*Marquardt's Compromise*".

O quadrado médio estimado dos parâmetros no modelo de regressão não-linear é diferente daquele apresentado para os modelos de regressão linear, pois não pode ser determinado por expressão matemática explícita. Os modelos não-lineares tem o dever de obter uma soma de quadrados mínima por algumas médias como num método iterativo, iniciando com um subconjunto de parâmetros iniciais (Ratkowsky, 1989). Há uma extensa literatura que discute e avalia estes métodos como Chambers (1973) ou Schmidt (1982). Entretanto, o conceito de iniciar a convergibilidade, que representa a facilidade de convergência dos parâmetros iniciais estimados, e sua relação com a amplitude na qual um modelo apresenta o comportamento "*close-to-linear*", é utilizado considerando somente o método de Marquardt, que tem a vantagem sobre os demais métodos como o método do gradiente (*steepest descent*) e de Gauss-Newton, pois agrega as melhores propriedades dos dois, convergência e rapidez.

A convergibilidade de um modelo para uma combinação de dados depende intensamente da forma do contorno da idêntica soma de quadrados do resíduo na superfície de resposta esboçada no espaço paramétrico (Ratkowsky, 1983).

Uma das limitações dos modelos não lineares é a escolha correta dos valores iniciais dos parâmetros para o processo de iteração, os quais irão gerar as estimativas pela convergência do algoritmo, baseado, na maioria das vezes, no método de Gauss-Newton (Calegario, 2004). O mesmo autor comenta que quando os valores iniciais dos parâmetros estão distantes das estimativas para a base de dados em questão, o processo de convergência do algoritmo falha e não ocorre à estimativa, ou o processo gera estimativas não confiáveis. Uma das formas de contornar o problema é o uso de expressões geradas com base na interpretação dos parâmetros do modelo e, ou, em expressões geradas pelo isolamento do parâmetro em função de outros parâmetros conhecidos e pontos específicos na base de dados.

Conforme Souza (1998), a determinação numérica dos parâmetros a serem estimados pode ser problemática no caso de modelos não-lineares. O sucesso na utilização de um procedimento para estimar os parâmetros vai depender da escolha apropriada da função resposta $f(x, \beta)$ e de bons valores iniciais. Draper & Smith (1981) apresentaram uma lista de alternativas para obtenção dos valores iniciais.

Gallant (1987) alerta com respeito à determinação de β , que o sucesso, em termos de convergência para β , não é garantido por qualquer método. A experiência indica que a falha no processo de convergência (quando não erramos as derivadas) depende da distância do valor inicial a resposta correta e do grau de parametrização da função resposta relativamente ao conjunto de dados utilizado. Estes problemas estão inter-relacionados no sentido de que funções respostas mais apropriadas conduzem a uma amplitude maior do raio de convergência (no espaço paramétrico). Quando a convergência não ocorre, deve-se tentar encontrar valores iniciais melhores (mais próximos de β) ou utilizar uma função resposta mais parcimoniosa (com menos parâmetros). Uma boa

maneira de verificar a precisão da solução numérica é provocando distúrbios nas condições iniciais e verificando se os novos valores iniciais conduzem à mesma solução.

Segundo Hoffmann & Vieira (1987), a validade dos estimadores de mínimos quadrados (do vetor de parâmetros), da estatística F na análise de variância, da distribuição t para construção de intervalos de confiança, dentre outros, está vinculada a certas pressuposições sobre o vetor de erros (ε). A primeira delas é que a média dos erros é nula, isto é:

$$E(\varepsilon) = 0$$

A seguir, tem-se que a variância do erro ε_i , $i = 1, 2, \dots, n$, é constante e igual a σ^2 , isto é:

$$E(\varepsilon \varepsilon) = I\sigma^2$$

Nesse caso, diz-se que os erros são homocedásticos. Em terceiro lugar, o erro de uma observação é não correlacionado com o erro de outra observação, isto é:

$$E(\varepsilon_i \varepsilon_j) = 0, i \neq j$$

A quarta pressuposição é que os erros têm distribuição normal.

A primeira pressuposição exclui, por exemplo, a existência de erros sistemáticos de medida da variável y. Se, a segunda pressuposição (erros homocedásticos) não é atendida, pode-se usar o método dos mínimos quadrados ponderados. Se a terceira não é obedecida, diz-se que há autocorrelação nos resíduos e neste caso pode-se usar modelos generalizados, ou então utilizar outros modelos. A quarta pressuposição é necessária para que, dentre outras estatísticas, seja possível usar o teste F da análise de variância e/ou construir intervalos de confiança para os parâmetros.

O método dos mínimos quadrados ordinário aplicado aos modelos não lineares, segue o seguinte procedimento; para o conjunto de variáveis $(x; y)$ e o modelo: $y = f(x; \beta) + \varepsilon$. O objetivo do método é encontrar um vetor β estimado, para o qual a norma quadrática de ε seja mínima. Considerando que $f(x; \beta)$ é diferenciável com respeito a β , pode-se escrever:

$$Z = [z_{ij}] = \left[\frac{\partial f_i(x_i; \beta)}{\partial \beta_j} \right]; i = 1, \dots, n, j = 1, \dots, p$$

Sabe-se que uma condição para a existência do mínimo é:

$$\frac{(\partial \varepsilon \varepsilon)}{\partial \beta_j} = 0; j = 1, \dots, p$$

β_{estimado}

isto é:

$$Z' [Y - f(x; \beta)] = 0$$

que é conhecido como Sistema de Equações Normais (SEN).

3 MATERIAL E MÉTODOS

3.1 Dados

Os dados vieram da empresa Aracruz Celulose S.A., localizada na região costal do Brasil, nos estados do Espírito Santo e Bahia. As áreas da empresa apresentam os tipos climáticos, segundo a classificação de Köppen, de Aw (Clima Tropical Úmido, com estação chuvosa e seca no inverno) e Am (Clima Tropical Úmido, sem estação seca pronunciada), no estado do Espírito Santo. No estado da Bahia a classificação, segundo Köppen é Af (Clima Quente em que o mês mais frio atinge temperatura superior a 18°C), Am (Clima Tropical Chuvoso, de monção) e Aw (Clima Quente com temperatura do mês mais frio superior a 18°C).

As coordenadas geográficas, aproximadas, para cada região administrativa e suas respectivas áreas são apresentadas na Tabela 1.

TABELA 1 Coordenadas geográficas aproximadas e extensão da área da empresa Aracruz Celulose S.A.

	Aracruz	São Mateus	Posto da Mata
Latitude (S)	19°35' e 20°15'	18°05' e 18°45'	17°15' e 18°05'
Longitude (W)	40°00' e 40°20'	39°30' e 40°10'	39°05' e 40°00'
Área (ha)	33.114,50	44.162,00	47.689,20

Fonte: EMBRAPA (2000)

Os dados do inventário florestal utilizados para este estudo são provenientes de parcelas permanentes de tamanho variado, de formato circular, dispostas em diferentes extratos de plantios de clones de *Eucalyptus* spp, onde suas respectivas medidas de diâmetro(tomada a 1,30m de altura) e altura total foram tomadas em um curto intervalo de tempo, propiciando dados de entrada concernentes a uma modelagem por árvores individuais e apresenta um total de

4.683 observações individuais distribuídos entre diferentes clones e idades como pode ser observado na Tabela 2.

TABELA 2 Apresentação em síntese dos dados experimentais provenientes do inventário florestal contínuo, fornecido pela empresa Aracruz Celulose S.A.

Clone	Nº da parcela	Idade(anos)	Nº de árvores/parcela	Nº de Medições	Total de observações
1	1	3,08-4,67	33	14	431
1	2	3,08-4,67	31	15	448
1	11	2,92-4,50	41	13	605
2	8	2,92-4,50	24	13	866
3	4	3,33-4,92	10	12	113
3	6	2,92-4,50	15	13	171
3	7	3,25-4,83	33	15	420
3	12	2,92-4,50	34	15	440
4	5	2,92-4,50	18	14	224
5	10	2,92-4,50	54	14	742
6	9	3,25-4,83	17	14	223

3.2 Desenvolvimento

O crescente interesse em se desenvolver modelos de crescimento gera diferentes idéias, para diferentes pesquisadores. A idéia para este estudo foi baseada na ausência de um modelo teórico que relacione a variável dependente área seccional (g) com a variável independente idade (t). O ponto inicial de desenvolvimento abrange uma investigação do diagrama de dispersão Figura 2.

O comportamento da relação entre estas duas variáveis pode ser visualizado na Figura 2, em que a taxa relativa de crescimento denominada incremento corrente anual relativo (ICA_R) da área seccional é expressa em anos. O diagrama de dispersão fornece uma boa idéia da forma da relação exibida pela relação entre as variáveis.

Pela linha em destaque do diagrama de dispersão é possível visualizar a

19

(1989), que pode ser escrita como:

$$g = t^{\beta_0} \quad (1)$$

onde: g = área seccional da árvore, em m^2 ; t = idade, em anos; β_0 = parâmetro

Na equação (1), se relacionou sua função incógnita com sua derivativa, gerando a seguinte expressão que descreve a taxa de crescimento relativa da área seccional em função do tempo, esta equação diferencial é classificada como equação diferencial de Bernoulli:

$$\frac{\partial g}{g \partial t} = t^{\beta_0} \quad (2)$$

Aplicando-se a antiderivativa à equação (2)

$$\int \frac{\partial g}{g} = \int t^{\beta_0} \partial t$$

a seguinte solução para equação (2), é obtida com o auxílio do software MAPLE V® :

$$\ln(g) = \frac{t^{(\beta_0+1)}}{\beta_0 + 1} \quad (3)$$

e na busca de uma equação que expressasse uma curva de crescimento exponencial, conforme a linha de tendência do diagrama de dispersão, foi utilizada a seguinte equação:

tendência de uma curva, que permite a se decidir por um dado modelo.

DIAGRAMA DE DISPERSÃO

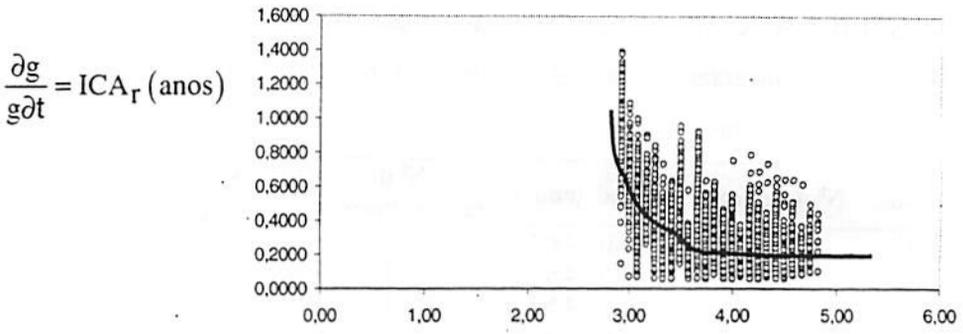


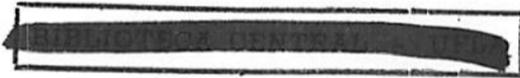
FIGURA 2 Diagrama de dispersão, que expressa a relação entre as duas variáveis de interesse.

Baseado no trabalho de Ratkowsky (1989) onde ele apresenta uma série de equações que expressam a relação entre uma variável dependente (y) e uma independente (x) em curvas côncavas e convexas, foram selecionadas possíveis equações que geram curvas côncavas, procurando atender a necessidade de se desenvolver um modelo exponencial conforme a tendência da linha em destaque da Figura 2.

Para se obter uma operacionalidade no desenvolvimento dos modelos, se optou por selecionar equações com 1, 2 e 3 parâmetros do trabalho de Ratkowsky(1989), totalizando três equações, sobre as quais se utilizou o conceito de equação diferencial, que é toda equação que relaciona uma função incógnita (função desconhecida) com uma ou mais das suas derivadas.

Uma importante ferramenta utilizada neste trabalho foi o software MAPLE V®, que permitiu obter as soluções destas equações diferenciais.

O desenvolvimento da modelagem de um novo modelo foi elaborado em



em que: g = área seccional da árvore, em m²; exp =exponencial, t = idade, em anos; β₀ = parâmetro

Na segunda etapa o desenvolvimento se deu a partir da primeira equação com 2 parâmetro selecionada, a partir do trabalho de Ratkowsky (1989), que pode ser escrita como:

$$g = \beta_0 * t^{\beta_1} \quad (4)$$

Na equação (4), relacionou-se a função incógnita com a derivativa, gerando a seguinte expressão, que descreve a taxa de crescimento relativa da área seccional em função do tempo, esta equação diferencial é classificada como equação diferencial de Bernoulli:

$$\frac{\partial g}{g \partial t} = \beta_0 * t^{\beta_1} \quad (5)$$

Aplicando-se a antiderivativa a equação (5)

$$\int \frac{\partial g}{g} = \int \beta_0 * t^{\beta_1} \partial t$$

a seguinte solução para equação (5) é obtida com o auxílio do software MAPLE V® :

$$\ln(g) = \frac{\beta_0 * t^{(\beta_1+1)}}{(\beta_1 + 1)} \quad (6)$$

$$g = \exp\left(\frac{\beta_0 * t^{(\beta_1+1)}}{(\beta_1 + 1)}\right) \text{ Modelo Original II}$$

em que g = área seccional da árvore, em m^2 ; \exp = exponencial, t = idade, em anos; β_0 e β_1 = parâmetros

Na terceira etapa o desenvolvimento se deu a partir da primeira equação com 2 parâmetros selecionados, a partir do trabalho de Ratkowsky (1989), que pode ser escrita como:

$$g = \beta_0 + \beta_1 * (t)^{\beta_2} \quad (7)$$

Na equação (7), se relacionou sua função incógnita com sua derivativa, gerando a seguinte expressão que descreve a taxa de crescimento relativa da área seccional em função do tempo, esta equação diferencial é classificada como equação diferencial de Bernoulli:

$$\frac{\partial g}{g \partial t} = \beta_0 + \beta_1 * (t)^{\beta_2} \quad (8)$$

Aplicando-se a antiderivativa à equação (8)

$$\int \frac{\partial g}{g} = \int \beta_0 + \beta_1 * (t)^{\beta_2} * \partial t$$

a seguinte solução para equação (8), é obtida com o auxílio do software MAPLE V[®] :

$$\ln g = \beta_0 * t + \left(\frac{\beta_1 * t^{(\beta_2+1)}}{\beta_2 + 1}\right) \quad (9)$$

e na busca de uma equação que expressasse uma curva de crescimento exponencial, conforme a linha de tendência do diagrama de dispersão, tornou-se necessário o rearranjo da equação (9), onde, por simples operação matemática retirou-se o antilog da equação (9) resultando a seguinte função exponencial de crescimento. Como na primeira etapa, esta função de crescimento foi denominado por modelo original 3:

$$g = \exp \left(\beta_0 * t + \left(\frac{\beta_1 * t^{(\beta_2)}}{\beta_2} \right) \right) \text{ Modelo Original III}$$

em que g = área seccional da árvore, em m^2 ; \exp = exponencial, t = idade, em anos; β_0 , β_1 e β_2 = parâmetros

Após a apresentação destas três etapas, três novos modelos teóricos foram definidos. Estes modelos foram denominados de modelos originais. Na busca de desenvolver novos modelos baseados em equações diferenciais, que possam resultar em estimativas mais precisas, estes modelos originais sofreram algumas modificações, como a inclusão de variáveis como Clone, altura total (HT), diâmetro (DAP), densidade (N), índice de sítio (IS), altura dominante (HD) e área basal (G).

Todas estas variáveis estão correlacionadas diretamente com o crescimento da área seccional (g) em relação ao tempo (t), desenvolvendo-se assim, três novos modelos, denominados modelos completos. Esta inclusão de variáveis foi feita de forma simples, através da decomposição dos parâmetros dos modelos originais, e a determinação de cada variável a ser incluída a cada parâmetro, foi feita de acordo com o seu grau de significância em relação ao parâmetro. O grau de significância foi verificado através do relatório de resultados fornecido pelo software estatístico S-plus®. A Tabela 3 apresenta os seguintes modelos completos desenvolvidos a partir dos modelos originais.

Tanto os modelos originais como os modelos completos foram ajustados, considerando a seguinte situação: os erros são independentes e identicamente distribuídos, com distribuição normal, média zero e variância constante σ^2 . Tendo-se utilizado novamente o software estatístico S-plus® para obtenção dos parâmetros.

Os valores iniciais requeridos para reduzir o número de iterações necessárias, para a solução do sistema de equações, foram obtidos com o uso da

seguinte alternativa: a plotagem da relação área seccional (g) versus idade (t) em planilhas do EXCEL. Esta é uma das alternativas que compõem uma lista para se obter valores iniciais para os parâmetros dos modelos não lineares, tendo sido apresentado por Draper & Smith (1981). Isto auxiliou a obtenção dos valores iniciais que foram determinados de forma iterativa, para se alcançar uma simulação mais realista possível.

TABELA 3 Apresentação dos Modelos Completos obtidos a partir da decomposição dos parâmetros dos Modelos Originais.

Modelo Original I

$$g = \exp \left(\frac{t^{(\beta_{00} + 1)}}{\beta_{00} + 1} \right)$$

Modelo Completo I

$$g = \exp \left(\frac{t \left(b_{00} + \sum_{i=1}^5 b_{0i} \text{Clone}(i) + b_{06} \text{DG} + b_{07} \text{HT} + b_{08} \text{HD} + b_{09} \text{G} + b_{10} \text{IS} + b_{111} \text{N} \right)}{\left(b_{00} + \sum_{i=1}^5 b_{0i} \text{Clone}(i) + b_{06} \text{DG} + b_{07} \text{HT} + b_{08} \text{HD} + b_{09} \text{G} + b_{10} \text{IS} + b_{111} \text{N} \right)} \right)$$

Modelo Original II

$$g = \exp \left(\frac{\beta_{00} * t^{(\beta_{10} + 1)}}{(\beta_{10} + 1)} \right)$$

Modelo Completo II

$$g = \exp \left(\frac{b_{00} + \sum_{i=1}^5 b_{0i} \text{Clone}(i) + b_{06} \text{DG} + b_{07} \text{HT} + b_{08} \text{HD} + b_{09} \text{G} + b_{10} \text{IS} + b_{10} * t \left(b_{10} + \sum_{i=1}^5 b_{1i} \text{Clone}(i) + b_{16} \text{HT} + b_{17} \text{G} + b_{18} \text{IS} + b_{19} \text{N} \right)}{b_{10} + \sum_{i=1}^5 b_{1i} \text{Clone}(i) + b_{16} \text{HT} + b_{17} \text{G} + b_{18} \text{IS} + b_{19} \text{N}} \right)$$

Continua...

TABELA 3 Continuação

Modelo Original III

$$g = \exp \left(b_{00} * t + \left(\frac{b_{10} * t^{(b_{20})}}{b_{20}} \right) \right)$$

Modelo Completo III

$$g = \exp \left(\frac{(b_{00} + \sum_{i=1}^5 b_{0i} \text{Clone}(i) + b_{06} \text{HT} + b_{07} \text{IS} + b_{08} \text{N} + b_{09} \text{DG}) * t + (b_{10} + \sum_{i=1}^5 b_{1i} \text{Clone}(i) + b_{16} \text{HT} + b_{17} \text{HD} + b_{18} \text{N}) * t}{b_{20} + \sum_{i=1}^5 b_{2i} \text{Clone}(i) + b_{26} \text{HT} + b_{27} \text{DG} + b_{28} \text{N}} \right)$$

em que: g = área seccional da árvore; t = idade; DAP = diâmetro a altura do peito; HT = altura total; HD = altura dominante; G = área basal; IS = índice de sítio, N = densidade e b₀, b₁ e b₂ = parâmetros .

A qualidade do ajusta dos modelos foi verificada por meio das seguintes análises:

3.2.1 Medidas de Acuracia

Como o desafio deste estudo era obter um modelo preciso e exato, foi determinado que se utilizasse erro padrão residual (S_{yx}) como critério de efeito comparativo, pois, este expressa o quanto, em termos médios, os valores observados variam em relação os valores estimados. A unidade do (S_{yx}) é a mesma da variável dependente (y). Significa que quanto mais próximo o valor de (S_{yx}) for de 0 (zero) , mais eficiente tende o modelo. (S_{yx}) é dado por:

$$S_{yx} = \sqrt{\text{QMRes}}$$

onde: QMRes = Quadrado médio do resíduo ou variável residual.

O quadrado médio do resíduo (QMRes) é:

$$QMRes = \frac{SQRes}{n - p}$$

Com isso, pode ser mostrado que o valor esperado do QMRes é a σ^2 , um estimador não tendencioso da variância, é dado por :

$$\sigma^2 = QMRes$$

3.2.2 Análise de Resíduos

O ajuste dos modelos ao conjunto de dados é muito útil para analisar, interpretar e fazer previsões sobre as questões de interesse. Entretanto, o desenvolvimento desses modelos exige uma série de conjunturas para o fenômeno, e uma boa modelagem não estaria completa sem uma adequada investigação da veracidade das mesmas. A existência e, conseqüente detecção da transgressão de algumas suposições, permite evitar o emprego de modelos pobres e de pouca utilidade e que acarretam baixa confiabilidade nos seus resultados.

Uma maneira de investigar este problema é estudando o comportamento do modelo no conjunto de dados observados, principalmente, as discrepâncias entre os valores observados e os valores ajustados, ou seja, pela análise de resíduos. Segundo Souza (1998), a idéia fundamental da análise de resíduos é a utilização dos resíduos da regressão no estabelecimento de técnicas visuais (gráficos) e formais (escalonares) para detectar os desvios, e o uso de procedimentos alternativos necessários para a remoção de anomalias uma vez detectadas. Tecnicamente, para cada observação i , temos associados o resíduo (ϵ_i) conhecido como erro propriamente dito que é igual a diferença entre o valor observado y_i e o ajustado \hat{y}_i , onde que para cada modelo ajustado o resíduo será dado por:

$$e = y - \hat{y}$$

Os resíduos, além de proporcionar medidas de qualidade para modelo e

seus estimadores, permitem o estudo de seu comportamento conjunto e, ou individual, baseando nas pressuposições dos erros (Bussab, 1988).

3.2.3 Análise gráfica dos resíduos

Embora não seja considerada uma medida de acuracidade, o gráfico de resíduo fornece informações fundamentais para seleção do melhor modelo de regressão ajustado.

Os resíduos ou erros são obtidos pela subtração dos valores reais (y) e os valores estimados, ou ajustados, (\hat{y}). O exame da distribuição dos resíduos é válido tanto para regressão linear quanto para regressão não-linear.

Os gráficos de resíduos possibilitam a detecção da normalidade, “*outliers*” e heterogeneidade da variância.

Os resíduos são analisados em função dos valores estimados para verificar qual dos modelos se adequou melhor aos dados.

4 RESULTADOS E DISCUSSÃO

4.1 Análise individual dos dados

O inventário florestal de clones de *Eucalyptus* spp para este estudo propiciou dados de entrada concernentes a uma modelagem por árvores individuais. Com os resultados deste inventário as seguintes análises foram realizadas, como mostra a Figura 3, a qual apresenta a taxa relativa de crescimento em área seccional expressa em anos. A Figura 4 apresenta o crescimento da área seccional, em relação ao tempo, expressa em m². Para efeito de ilustração, os gráficos destas figuras apresentam somente informações referentes a parcela 1 do clone 1 totalizando 33 árvores (Vide Tabela 2).

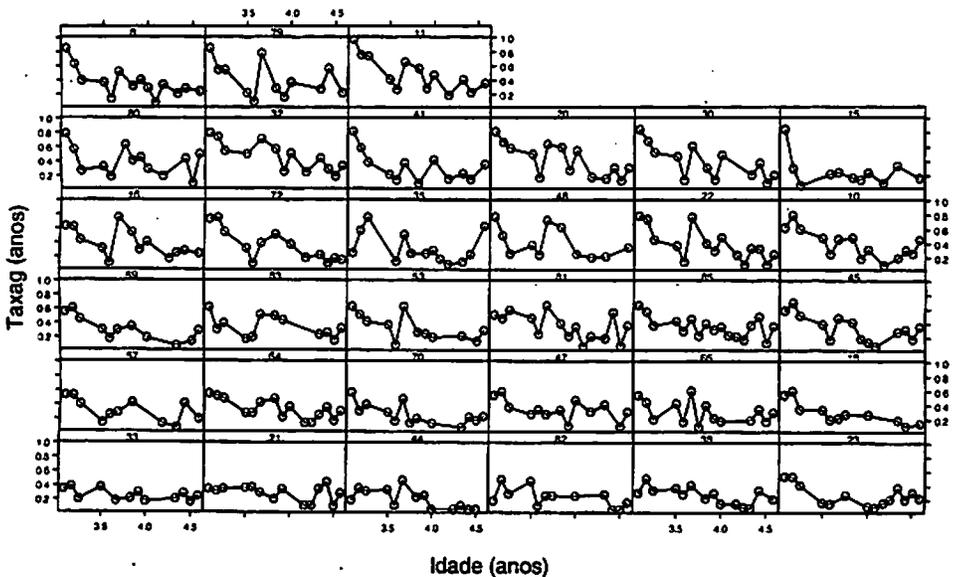


FIGURA 3 Taxa de crescimento relativo individual da área seccional para as árvores da parcela 1 do clone 1.

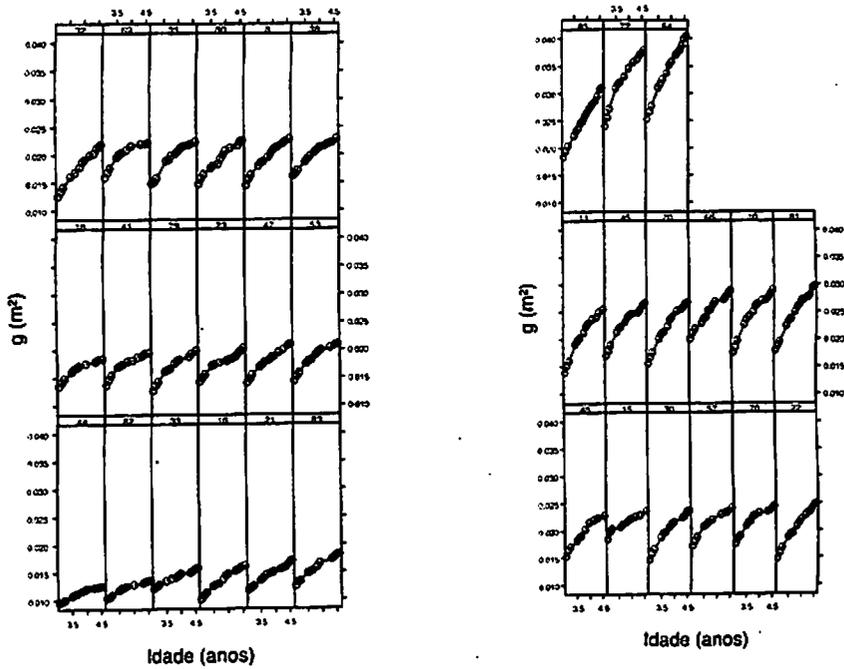


FIGURA 4 Crescimento individual em área basal para as árvores da parcela 1 do clone 1.

Analisando as Figuras 3 e 4, pode-se concluir que muitas das vezes as intervenções realizadas nos plantios de eucalipto, como o desbaste, são feitas de forma sistemática. Porém, com um simples diagnóstico realizado nestas figuras, é possível visualizar que a árvore número 44 é a árvore que estabilizou seu crescimento em área seccional abaixo de $0,015 \text{ m}^2$, bem antes de completar 5 anos de idade. Também pode se verificar que a mesma árvore não apresenta mais uma taxa de crescimento acentuada a partir dos 4 anos de idade, podendo ser considerada como uma árvore suprimida. Esta seria uma forma de determinar ao desbaste, quais seriam as árvores a ser eliminadas, favorecendo assim o

incremento em área seccional, para árvores como a de número 64, 72, 85, da Figura 4.

Também, pode-se verificar que as árvores de número 64, 72 e 85 possuem uma taxa de crescimento oscilante, tendendo à estabilização. Por exemplo, a retirada de árvores, como a de número 44, consideradas como competidoras, proporcionaria um novo alavanque de desenvolvimento para estas árvores.

Assim sendo, estima-se que com a retirada da árvore número 44, aos 3 anos de idade, ocasionaria, antecipadamente, um novo ritmo de crescimento às árvores que estão localizadas ao seu redor, devido a menor competição ocasionada por sua retirada. Simultaneamente, agregaria um valor econômico as árvores localizadas ao seu redor. Estas árvores serão destinadas à serraria, promovendo assim um menor ciclo de corte. Enquanto isso, a árvore número 44 seria destinada a fins energéticos.

4.2 Estimativas dos parâmetros e medidas de acuracidade

Os resultados da estimativa dos parâmetros dos modelos originais e modelos completos, com suas respectivas medidas (em síntese) de acuracidade fornecidas pelo software S-plus®, são apresentados de forma comparativa entre modelos originais e completos nas Tabelas 4, 5 e 6.

TABELA 4 Estatística comparativa entre os modelos: *Original I e Completo I*, seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual (EPR) de cada modelo avaliado.

<i>Modelo Original I</i>		EPR (m ²) = 0,00439129	EPR (%) = 24,85	
Parâmetro	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0	-1,191611	0,0001366597	-8.719.547	< 0,0001
<i>Modelo Completo I</i>		EPR (m ²) = 0,01720029	EPR (%) = 99,42	
Parâmetro	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
$\beta_{0\text{combinado}}$	-1,712978	0,008001524	-214.0814	< 0,0001

em que: β_0 = parâmetro estimado para o *Modelo Original I* e $\beta_{0\text{combinado}}$ = parâmetro estimado para o *Modelo Completo I*, combinado com as variáveis (Clone+Diâmetro quadrático + altura total + altura dominante + área basal + índice de sítio + densidade).

TABELA 5 Estatística comparativa entre os modelos: *Original II e Completo II*, seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual(EPR) de cada modelo avaliado.

<i>Modelo Original II</i>		EPR (m ²) = 0,00425539	EPR (%) = 24,59	
Parâmetro	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0	1,619701	0,05426565	29,8476	< 0,0001
β_1	-1,276822	0,00675446	-189.0339	< 0,0001
<i>Modelo Completo II</i>		EPR (m ²) = 0,002738841	EPR (%) = 15,83	
Parâmetro	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0	-2,801093	0,6827589	-4,10261	< 0,0001
β_0 .Clone1	-0,053449	0,0102010	-5,23959	< 0,0001
β_0 .Clone2	-0,048296	0,0051315	-9,41155	< 0,0001
β_0 .Clone3	-0,002973	0,0038368	-0,77491	0,4384

Continua...

TABELA 5 Continuação.

<i>Modelo Completo II</i>		EPR (m ²) = 0.002738841	EPR (%) = 15.83	
<i>Parâmetro</i>	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0 .Clone4	-0.005140	0,0025290	-2,03234	0,0422
β_0 .Clone5	0,046315	0,0095657	4,84181	< 0,0001
β_0 .DG	-0,014321	0,0015449	-9,26999	< 0,0001
β_0 .HT	0,058954	0,0011946	49,34924	< 0,0001
β_0 .G	0,011289	0,0040304	2,80085	0,0051
β_0 .IS	0,005920	0,0110088	0,53775	0,5908
β_0 .N	0,019822	0,0052510	3,77496	0,0002
β_1	-0,396432	0,1353794	-2,92830	0,0034
β_1 .Clone1	0,010502	0,0020046	5,23881	< 0,0001
β_1 .Clone2	0,010402	0,0010362	10,03838	< 0,0001
β_1 .Clone3	0,000413	0,0007968	0,51823	0,6043
β_1 .Clone4	0,000662	0,0005037	1,31386	0,1890
β_1 .Clone5	-0,009279	0,0018975	-4,89012	< 0,0001
β_1 .HT	-0,012839	0,0002364	-54,30398	< 0,0001
β_1 .G	-0,001573	0,0008469	-1,85784	0,0633
β_1 .IS	-0,000699	0,0022055	-0,31680	0,7514
β_1 .N	-0,004148	0,0010411	-3,98397	0,0001

em que: DG = diâmetro quadrático; HT = altura total; G = área basal; IS = índice de sítio, N = densidade e β_0 β_1 = parâmetros estimados.

TABELA 6 Estatística comparativa entre os modelos: *Original III e Completo III*, seguidos de seus respectivos: desvio padrão, do valor de t, probabilidade e erro padrão residual(EPR) de cada modelo avaliado.

<i>Modelo Original III</i>		EPR (m ²) = 0.00425539	EPR (%) = 24.50	
<i>Parâmetro</i>	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0	-0,298776	0,035289	-8,466443	< 0,0001
β_1	-1,276822	0,00675446	6,017650	< 0,0001
β_2	-0,769441	0,081007	-9,498078	< 0,0001

Continua...

TABELA 6 Continuação

<i>Modelo Completo III</i>		EPR (m ²) = 0.001974689	EPR (%) = 11,41	
<i>Parâmetro</i>	$\hat{\beta}$	Desvio Padrão	Valor de "t"	Pr > t
β_0	-4.215021	37.88313	-0,11126	0,9114
β_0 .Clone1	0,202262	0,06288	3,21662	0,0013
β_0 .Clone2	-0,066682	0,01987	-3,35669	0,0008
β_0 .Clone3	-0,020359	0,01437	-1,41635	0,1567
β_0 .Clone4	6,199987	45,71076	0,13564	0,8921
β_0 .Clone5	-0,856091	7,61382	-0,11244	0,9105
β_0 .HT	0,101991	0,00529	19,28961	< 0,0001
β_0 .IS	-0,017123	0,00321	-5,33225	< 0,0001
β_0 .N	0,098885	0,01669	5,92568	< 0,0001
β_0 .DG	-0,038631	0,00813	-4,75017	< 0,0001
β_1	4,343395	37,97768	0,11437	0,9090
β_1 .Clone1	-0,268697	0,07182	-3,74150	0,0002
β_1 .Clone2	0,066908	0,02434	2,74882	0,0060
β_1 .Clone3	0,036513	0,02169	1,68306	0,0924
β_1 .Clone4	-6,293245	45,75080	-0,13765	0,8905
β_1 .Clone5	0,862467	7,61654	0,11324	0,9098
β_1 .HT	-0,054419	0,00599	-9,09126	< 0,0001
β_1 .HD	0,018544	0,00294	6,31478	< 0,0001
β_1 .N	-0,118559	0,01575	-7,52939	< 0,0001
β_2	-4,407952	0,31674	-13,91645	< 0,0001
β_2 .Clone1	0,113054	0,01914	5,90554	< 0,0001
β_2 .Clone2	-0,026960	0,00646	-4,17548	< 0,0001
β_2 .Clone3	-0,015156	0,00844	-1,79656	0,0725
β_2 .Clone4	0,145626	0,03271	4,45138	< 0,0001
β_2 .Clone5	0,038074	0,00845	4,50503	< 0,0001
β_2 .HT	0,000168	0,00141	0,11875	0,9055
β_2 .DG	0,066016	0,00628	10,51137	< 0,0001
β_2 .N	0,046621	0,00329	14,15450	< 0,0001

em que: DG = diâmetro quadrático, HT = altura total; HD = altura dominante; G = área basal; IS = índice de sítio, N = densidade e $\beta_0 \beta_1 \beta_2$ = parâmetros estimados.

Entre os modelos originais, se destaca o *Modelo Original III*, pois o mesmo apresentou medidas de acuracidade significativas, como é observado na

Tabela 6, o erro padrão residual (EPR), expresso em termos absolutos (m^2) e em porcentagem (%).

Verifica-se nas Tabelas 4, 5 e 6 que a distribuição de t , para a construção de intervalos de confiança, confere os parâmetros de todos os modelos originais uma contribuição significativa para expressar a relação existente entre a variável área seccional (g) e a variável idade (t).

As análises feitas para os modelos completos mostram, surpreendentemente, que a inclusão de outras variáveis aos parâmetros do *Modelo Completo I*, não lhe proporcionou uma redução no valor do erro padrão residual como foi o caso do *Modelo Completo II* e *Modelo Completo III*, ao invés, gerou um elevado erro padrão residual, o que significa que a inclusão de novas variáveis aos parâmetros do *Modelo Completo I* não contribui significativamente para seu ajuste.

Porém, a inclusão de novas variáveis nos outros modelos indica que o *Modelo Completo III* mostra-se superior ao *Modelo Completo II*, pois, o mesmo apresentou menor valor de erro padrão residual.

A inclusão de novas variáveis nos parâmetros do *Modelo Completo III* mostra que os valores da distribuição de t para a construção de intervalos de confiança, não confere aos parâmetros do modelo, uma contribuição significativa para expressar a relação existente entre a variável área seccional (g) e a variável idade(t). Ao contrário no *Modelo Completo II*, a inclusão de novas variáveis proporcionou uma contribuição um pouco mais significativa para expressar a relação existente entre a variável área seccional (g) e a variável idade(t).

Todas estas variáveis que estão sendo incluídas de forma combinada a cada um dos parâmetros que compõem os modelos, estão de maneira geral umas relacionadas com as outras, ou seja, uma depende da outra, pois a retirada ou a inclusão de qualquer variável a qualquer parâmetro destes modelos, modificam

significativamente a estrutura de todo o ajuste. O que mostra que essa técnica de inclusão de variáveis aos modelos precisa ser melhor estudada, de forma detalhada, para cada situação de interesse, pois tem-se em mãos um novo modelo, que é complexo, porém muito promissor.

4.3 Análise gráfica residual para os modelos

Nas Figuras 5, 6 e 7, estão representados, de forma comparativa, os gráficos dos resíduos versus o valor da área seccional estimada, para todos os modelos.

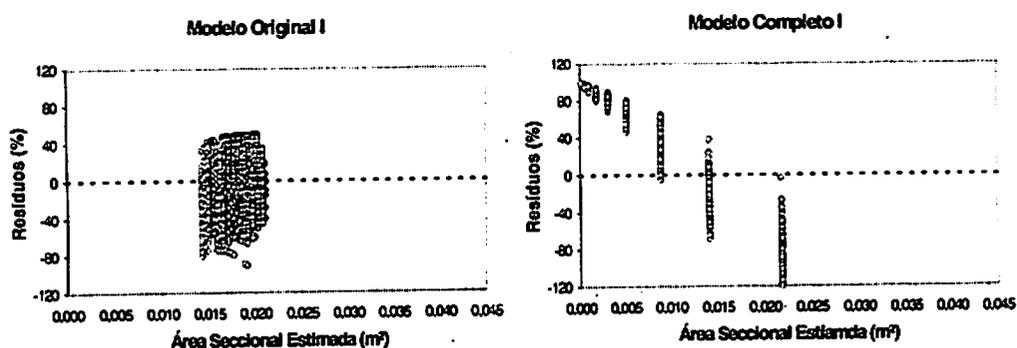


FIGURA 5 Distribuição dos resíduos percentuais em função da área seccional estimada de cada árvore, apresentado de forma comparativa entre os *Modelo Original I* e *Modelo Completo I*.

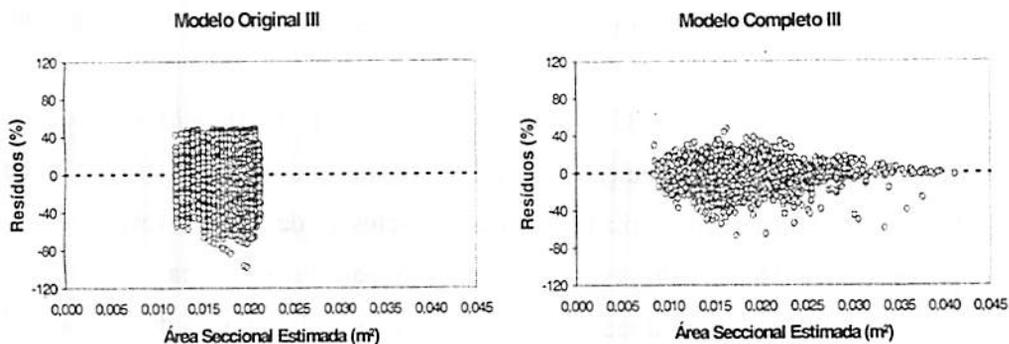


FIGURA 6 Distribuição dos resíduos percentuais em função da área seccional estimada de cada árvore, apresentado de forma comparativa entre os *Modelo Original II* e *Modelo Completo II*.

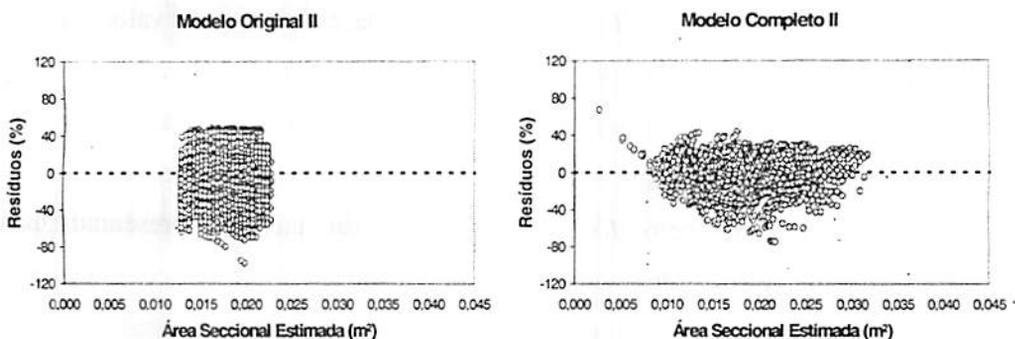


FIGURA 7 Distribuição dos resíduos percentuais em função da área seccional estimada de cada árvore, apresentado de forma comparativa entre os *Modelo Original III* e *Modelo Completo III*.

Analisando a distribuição dos resíduos percentuais nas Figuras 5, 6 e 7, verifica-se que os modelos originais apresentaram uma forma semelhante de

distribuição de resíduos, em que os resíduos estão distribuídos aleatoriamente em torno do eixo zero.

Para os Modelos Completos, verifica-se que o *Modelo Completo I* apresentou a pior distribuição de resíduos, indicando que a suposição de homocedasticidade foi violada, e que a inclusão de novas variáveis aos parâmetros do *Modelo Original I* não contribuiu para sua melhoria.

O grande ganho desta nova modelagem, esta na redução do erro, que é verificado através da análise gráfica dos resíduos percentuais aos *Modelos: Completo II e Completo III*, que exprime uma sensível melhoria na distribuição dos resíduos, em relação aos seus respectivos modelos originais, em que os resíduos foram distribuídos aleatoriamente em torno do eixo zero, de forma mais concentrada.

Outra análise de ganho também realizada mostra que os *Modelos: Completo II e Completo III* indicam precisão da estimativa do valor da área seccional, e que o *Modelo Completo III* se mostra superior ao *Modelo Completo II*. Neste, verifica-se a presença do menor valor observado da área seccional (0,008 m²) e do maior valor observado da área seccional (0,040 m²) em relação os valores estimados pelo *Modelo Completo III*. Em anexo, é apresentada uma cópia da planilha original dos dados do inventário florestal contínuo com destaque para os valores máximo e mínimo da área seccional.

5 CONCLUSÃO

De acordo com as análises desta nova técnica de modelagem, a partir de equações diferenciais, chegou-se as seguintes conclusões:

- Os modelos completos, com exceção do *Modelo Completo I*, apresentaram uma significativa redução no erro padrão residual. Verificado a análise gráfica de resíduos, é também concluído, que a presença de novas variáveis nos modelos originados de equações diferenciais, permitiu obter uma melhora na distribuição aleatória dos resíduos em torno do eixo zero, ocorrendo uma redução na heterogeneidade de variância existente nos modelos originais.
- Também, nesta mesma análise, o ganho com precisão é notado, principalmente com o *Modelo Completo III*, que quando comparado aos demais modelos mostra, que a presença dos valores observados é a mesma dos valores estimados no gráfico de resíduos.
- As variáveis mensuradas em um inventário florestal e aquelas relativas ao nível de parcela, como é o caso do diâmetro quadrático, da área basal, do índice de sítio, da altura dominante e da densidade, podem ser utilizadas no ajuste de modelos sem que haja um custo adicional, para obter tais informações.
- Esta nova modelagem afeta, diretamente, o plano de suprimento das empresas florestais, pois este pode ser corretamente elaborado, produzindo assim, informações com um alto grau de precisão e exatidão para o planejamento florestal, ocasionando uma redução de custo, às várias operações realizadas na cadeia produtiva de uma empresa florestal.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- BELLA, I. Growth, yield, and ENFOR. *Forestry Report*, Alberta, v.29, n.4, p. 1-8, 1984.
- BRONSON, R. *Schaum's outline of theory and problems of differential equations*, 2 ed., U.S.A., 1993, p.8-9.
- BURKHART, H.E. Development of empirical growth and yields models. In: **EMPIRICAL AND PROCESS BASED MODELS FOR FOREST TREE AND STAND GROWTH SIMULATION**, 1., 1997, Oeiras. *Anais...* Lisboa: Salamandra, 1999. p. 53-60.
- BUSSAB, W. O. **Métodos quantitativos: Análise de variância e de regressão**, 2 ed., São Paulo, 1988. 147p.
- CALEGARIO, N. **Melhoria da Qualidade de Ajuste de Modelos Biométricos Florestais pelo Emprego da Teoria dos Modelos Não Lineares Generalizados**. IPEF. (Submetido). 2004. 17p.
- CALEGARIO, N.; DANIELS, R.F. **Modeling dominant height growth based on a nonlinear mixed effect model: a clonal *Eucalyptus* plantation case study**. *Forest Ecology and Management*, v.204 (1):p 11-21. 2005.
- CAMPOS, L.J.D. **Prognose por classe de diâmetro a partir de modelos do tipo povoamento total**. Viçosa: UFV, 1997. 57 p. (Dissertação – Mestrado em Engenharia Florestal).
- CHAMBERS, J. **Fitting Nonlinear Models: Numerical Techniques**, *Biometrika* London, n.1.,p1-13.,apr.1973.
- CLUTTER, J.L.; FORTSON, J.C; PIENNAAR, L.V.; BRISTER, G.H.; BAILEY, R.L., 1983. **Timber Management: A Quantitative Approach**. New York: John Willey; Sons.333p
- DAVIS, L.S.; JOHNSON, K.N. **Forest Management**. New York: Mc Graw-Hill, 1987. 790p.
- DRAPER, N.R.; SMITH, H. **Applied regression analysis**. New York: Wiley, 1981.

FITZHUGH JR., H. A. Analysis of growth curves and strategies for altering their shape. *Journal of Animal Science*, v.42, n.6, p.1036-1050, 1976.

GALLANT, A. R. *Nonlinear statistical models*. New York: Wiley, 1987.

GUIMARÃES, D.P. **Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto**. Viçosa: UFV, 1994. 160p. Dissertação (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, 1994.

HOFFMAN, R., VIEIRA, S. **Análise de regressão: uma introdução à econometria**. 2 ed. Hucitec. São Paulo, 1987. 379 p.

HUSCH, B., MILLER, C.I., BEERS, T.W. **Forest mensuration**. 3.ed. Malabar: Krieger Publishing Company, 1993. 402p.

KOOPS, W. J. Multiphasic Growth Curve Analysis. *Growth* 50, p. 169-177, 1986.

LUZ, A.M.S.; CORREA, F.J.S.A. **Equações diferenciais ordinárias e aplicações**. Revista Virtual de Iniciação Acadêmica da UFPA. Vol 1, No 1, março de 2001; Disponível em: <<http://www.ufpa.br/revistaic>>. Acesso em: 20 nov. 2004

MACHADO, M. A. F. T. A função logística e sua aplicação a dados de crescimento. In: **Seminários de Pós-graduação em Experimentação e Estatística**. Piracicaba, ESALQ, USP, 1980. 19p.

MARQUARDT, D. W. An algorithm for least squares estimation of nonlinear parameters. *Journal Society for Industrial and Applied Mathematics*. Philadelphia. v.11(2), p.431-441, 1963.

MEDEIROS, H. A. **Qualidade do ajuste da função logística monofásica e difásica com estrutura de erros independentes e autorregressivos**. Lavras; UFLA, 1999, 73 p. (Dissertação –Mestrado em Agronomia).

PESSOA, M.C.P.Y.; LUCHIARI JÚNIOR, A.; FERNANDES, E.N.; LIMA, M.A. **Principais modelos matemáticos e simuladores utilizados para análise de impactos ambientais das atividades agrícolas**. Jaguariúna: EMBRAPA/CNPMA, 1997. 83p.

RATKOWSKY, D.A. **Handbook of nonlinear regression models**. New York and Basel: Marcel Dekker, 1989. 241p.

RATKOWSKY, D.A. **Nonlinear regression modeling, a unified practical approach**, New York: Marcel Dekker, 1983.

SCHMIDT, R. **Advances in Nonlinear Parameter Optimization**, Berlin: Springer-Verlag. 1982.420p.

SCOLFARO, J.R.S. **Modelagem do crescimento e da produção florestal para florestas plantadas e nativas**. Lavras: UFLA/FAEPE, 1998. 453p.

SOARES, C. P. B. **Um modelo para o gerenciamento da produção de madeira em plantios comerciais de eucalipto**; 1999. 82 p. Tese de Doutorado. Universidade federal de viçosa.

SOUZA, G. S. **Introdução aos modelos de regressão linear e não-linear**. Brasília, Embrapa, 1998. 505p.

STAMFORD, A. **Matemática para Economistas**. Pernambuco, 2004.

Disponível em:

http://www.stamford.pro.br/ARQUIVOS/20001_programa.doc/. Acesso em: 10 nov. 2004

TORRES, C.T; FILHO D.C. de M.; OLIVEIRA, A.S. **Análise de Fourier e Equações Diferenciais Parciais**. Pernambuco, 2002. Disponível em:

http://www.prpg.ufpb.br/~enic/2002/res_exatas.htm. Acesso em: 10 nov. 2004

VANCLAY, J.K. **Modelling forest growth and yield: applications to mixed tropical forests**. Wallingford: CABInternational, 1994. 312p.

Wraith, J.M.; Or, D. Nonlinear parameter estimation using spreadsheet software. **Journal of Natural Resources** 27: 13-19, 1998.

ANEXOS

Amostra dos dados provenientes do inventário florestal contínuo,
fornecido pela empresa Aracruz Celulose S.A..

1	A Clone	B parcela	C Narv	D idade	E g	F HD	G IS	H G	I N
4668	Clone 6	9	765	4,58	0,022	25,00	26,00	26,00	75,00
4669	Clone 6	9	765	4,67	0,022	25,00	26,00	26,00	75,00
4670	Clone 6	9	765	4,75	0,022	25,00	26,00	26,00	75,00
4671	Clone 6	9	769	3,25	0,020	25,00	26,00	26,00	75,00
4672	Clone 6	9	769	3,33	0,021	25,00	26,00	26,00	75,00
4673	Clone 6	9	769	3,42	0,022	25,00	26,00	26,00	75,00
4674	Clone 6	9	769	3,67	0,025	25,00	26,00	26,00	75,00
4675	Clone 6	9	769	3,75	0,026	25,00	26,00	26,00	75,00
4676	Clone 6	9	769	3,83	0,026	25,00	26,00	26,00	75,00
4677	Clone 6	9	769	4,00	0,027	25,00	26,00	26,00	75,00
4678	Clone 6	9	769	4,08	0,028	25,00	26,00	26,00	75,00
4679	Clone 6	9	769	4,17	0,029	25,00	26,00	26,00	75,00
4680	Clone 6	9	769	4,42	0,029	25,00	26,00	26,00	75,00
4681	Clone 6	9	769	4,50	0,030	25,00	26,00	26,00	75,00
4682	Clone 6	9	769	4,58	0,030	25,00	26,00	26,00	75,00
4683	Clone 6	9	769	4,67	0,030	25,00	26,00	26,00	75,00
4684	Clone 6	9	769	4,75	0,031	25,00	26,00	26,00	75,00
4685				Mayor =	0,040				
4686				Menor =	0,008				
4687									

em que: Clone = número do clone, Narv = número da árvore, idade = idade da árvore em anos, HD = altura dominante em metros; G = área basal em m²; IS = índice de sítio em metros e N = densidade.