

# UMA PROPOSTA PARA A GERAÇÃO DE AMOSTRAS ALEATÓRIAS NOS PROBLEMAS DE SIMULAÇÃO EM MODELOS DE PLANEJAMENTO<sup>1</sup>

Luiz Alexandre Peternelli<sup>2</sup>, Gilson Fernandes da Silva<sup>3</sup> e Helio Garcia Leite<sup>4</sup>

**RESUMO** – Um modelo de predição do preço da celulose foi ajustado usando-se o tempo e o preço defasado como co-variáveis. A partir das estimativas dos parâmetros obtidas, foram propostas 48 possíveis tendências futuras para o preço da celulose. Posteriormente, três métodos de simulação foram usados para prever os valores futuros definidos pelas várias tendências:  $M_1 \Rightarrow P_{cel,f} = \mu$ ;  $M_2 \Rightarrow P_{cel,f} = \mu + \epsilon_p$  e  $M_3 \Rightarrow \mu_f + \epsilon_f$ , em que  $m$  é a parte sistemática do modelo, e  $e_f$  corresponde ao componente estocástico. Para as simulações foram usados o método de Monte Carlo e a distribuição triangular. Para comparar os valores simulados pelos três métodos com os conhecidos valores futuros nas várias tendências, foi usada a diferença relativa média entre os valores. No caso da ausência de tendência, os métodos  $M_1$  e  $M_2$  foram satisfatórios, apesar de o método  $M_2$  incluir distúrbios ao redor da média. No caso de haver tendência real, o método  $M_3$  teve a melhor “performance”, mesmo sendo influenciado pela acurácia na predição da tendência.

Palavras-chave: Simulação, análise de risco e planejamento florestal.

## ***A PROPOSAL FOR RANDOM SAMPLE GENERATION IN SIMULATION PROBLEMS OF PLANNING MODELS***

**ABSTRACT** – A cellulose price prediction model was adjusted using time and lagged price as covariates. From the model parameter estimates, 48 possible trends were proposed for future cellulose price. Following, three simulation methods were used to predict the future values defined by the various trends:  $M_1 \Rightarrow P_{cel,f} = \mu$ ;  $M_2 \Rightarrow P_{cel,f} = \mu + \epsilon_p$  and  $M_3 \Rightarrow \mu_f + \epsilon_f$ , where  $m$  is the systematic part and  $e_f$  is the stochastic component. The Monte Carlo method and a triangular distribution were used for the simulation. To compare the values simulated by the methods and the future values of the various trends, the Average Relative Difference was used. In case of no trend,  $M_1$  and  $M_2$  were satisfactory, although  $M_2$  included disturbances around the mean. In the case of a real trend,  $M_3$  had the best performance, though it was influenced by the accuracy in the predicted trend.

Keywords: Simulation, risk analysis and forest planning.

### **1. INTRODUÇÃO**

No que diz respeito ao planejamento florestal, grande parte das empresas baseia-se na experiência dos seus tomadores de decisão para selecionar a estratégia ótima de manejar a floresta. Contudo, isso nem sempre é possível, devido ao elevado número de variáveis que

afetam a tomada de decisão (VOLPI et al., 1999; SILVA et al., 2000). Técnicas de pesquisa operacional determinísticas, principalmente as de Programação Matemática, têm sido utilizadas para resolver esse tipo de problema (SCOLFORO, 1991; RODRIGUES et al., 1998; RODRIGUES et al., 1999; RODRIGUEZ e BORGES, 1999; VOLPI et al., 1999; SILVA et al., 2000).

<sup>1</sup> Recebido em 02.02.2004 e aceito para publicação em 05.04.2006.

<sup>2</sup> Departamento de Informática, Universidade Federal de Viçosa (UFV), 36570-000 Viçosa-MG, Brasil. E-mail: <peternelli@dpi.ufv.br>.

<sup>3</sup> Centro de Ciências Agrárias da Universidade Federal do Espírito Santo, Alto Universitário, Caixa Postal 16, 29500-000 Alegre-ES. E-mail: <gfsilva2000@yahoo.com>.

<sup>4</sup> Departamento de Engenharia Florestal da UFV. E-mail: <hglete@ufv.br>.



Embora a grande maioria dos modelos de planejamento utilizados seja determinística, grande parte das variáveis desses modelos é tipicamente aleatória, tornando-os limitados em relação ao tipo de resposta que possam vir a oferecer. Alguns autores têm manifestado preocupação na consideração de variáveis aleatórias em modelos de planejamento florestal, principalmente variáveis econômicas e operacionais, dentre os quais cabe citar Pickens e Dress (1988), Hof et al. (1992), Volpi et al. (1999) e Protil (2000). Nos trabalhos citados, fica clara a importância de se considerarem os impactos estocásticos na tomada de decisão, principalmente para análises de risco e geração de cenários. Protil (2000) considerou que a análise de variáveis estocásticas em processos de planejamento florestal, quando comparada com as análises puramente determinísticas, apresenta uma série de vantagens, dentre as quais a maior riqueza de informações e a maior segurança e confiabilidade na tomada de decisão.

De acordo com Volpi et al. (1999), várias das informações na área florestal são amostrais, obedecendo às equações do tipo  $Y_k = \mu + \varepsilon_k$ ,  $k = 1, 2, \dots, n$ , em que  $\mu$  é a parte sistemática que é avaliada de fato e  $\varepsilon_k$  é o componente estocástico, que entra como fonte de perturbação. Contudo, Souza (1999) argumentou que o desenvolvimento tecnológico pode determinar tendências tanto na redução quanto no crescimento dos custos e da produção, respectivamente. Quando isso ocorre, o parâmetro  $\mu$  da equação  $Y_k = \mu + \varepsilon_k$  passa a não representar adequadamente as variações das variáveis aleatórias consideradas ao longo do tempo, pois este termo pressupõe variações constantes ao longo do tempo, enquanto na prática tendências de crescimento ou decrescimento ou quaisquer outros tipos de tendência podem vir a acontecer.

O objetivo deste trabalho foi apresentar uma proposta para obtenção de amostras aleatórias a serem usadas nas análises de risco e geração de cenários em modelos de planejamento florestal, nos casos em que a variável aleatória tenha alguma tendência definida ao longo do tempo. A eficiência do método proposto, comparada com a de dois métodos usuais, foi avaliada via simulação estocástica e com o auxílio de dados reais de uma série histórica para o preço da celulose.

## 2. MATERIAL E MÉTODOS

### 2.1. O método proposto

O método proposto é extremamente simples e foi

idealizado com vistas a incorporar eventuais tendências na geração de amostras aleatórias a serem usadas nas análises de risco e geração de cenários em modelos de planejamento florestal, além de permitir o uso de informações empíricas ou subjetivas simples na determinação dos ruídos em torno dos valores médios.

#### 2.1.1. O modelo

O método se baseia no seguinte modelo estatístico:

$$Y_k = \mu_k + \varepsilon_k,$$

em que  $Y_k$  é o valor gerado para a observação no tempo  $k$ ,  $\mu_k$  é a parte sistemática do modelo, aquela que incorpora eventuais tendências e que será determinada usando-se qualquer técnica conhecida de regressão, séries temporais etc.; e  $\varepsilon_k$  corresponde ao erro aleatório, ou ruído, inerente à variável aleatória em apreço e determinado conforme explicado adiante neste trabalho.

Para a geração de cenários em determinado tempo  $k$ , basta obter o componente sistemático do modelo no tempo  $k$  e gerar os valores da variável de interesse, tomando-se diversos valores de  $\varepsilon_k$ . Alternativamente, basta gerar diversos valores de  $Y_k$  diretamente, após fixar o tempo  $k$ .

#### 2.1.2. Obtenção dos erros aleatórios ou ruídos

A estimação dos valores  $\varepsilon_k$  é também essencial nas análises de risco e geração de cenários em modelos de planejamento florestal. Esses valores devem ser obtidos com base em alguma distribuição de probabilidades.

Em processos de simulação estocástica, o sucesso em obter boas previsões depende grandemente da distribuição de probabilidade escolhida (LAW e KELTON, 1991). Conforme comentado por Wagner (1986), do ponto de vista prático quatro abordagens podem ser utilizadas para obter distribuições de probabilidade: usar introspecção, empregar dados históricos, achar aproximações convenientes e enunciar axiomas descritivos.

Quando poucas informações ou apenas probabilidades subjetivas estão disponíveis, sugere-se a utilização de distribuições mais simples e que resultem em uma boa aproximação. A distribuição triangular é uma que se destaca neste caso, devido à sua simplicidade, capacidade de expressar distribuições assimétricas e dependência de poucos parâmetros na

sua definição (ANDERSON, 1977, citado por RODRIGUEZ, 1987; LAW e KELTON, 1991).

Convém definir as características de uma variável aleatória que tenha distribuição triangular. Se uma variável aleatória  $X$  tem distribuição triangular com parâmetros  $a$ ,  $c$  e  $b$ , ou seja, valor mínimo, valor máximo e valor modal, respectivamente, então sua função densidade de probabilidades  $f(x)$  e sua função de distribuição acumulada  $F(x)$  serão dadas, respectivamente, por:

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{para } a \leq x \leq b \\ \frac{2(c-x)}{(c-b)(c-a)} & \text{para } b < x \leq c \\ 0 & \text{caso contrário} \end{cases}; \quad F(x) = \begin{cases} 0 & \text{para } x < a \\ \frac{(x-a)^2}{(b-a)(c-a)} & \text{para } a \leq x < b \\ 1 - \left[ \frac{(c-x)^2}{(c-b)(c-a)} \right] & \text{para } b \leq x < c \\ 1 & \text{para } x \geq c \end{cases}$$

Usando, assim, o Método de Inversão (BUSTOS e ORGAMBIDE, 1992; CASELLA e BERGER, 2002), podem-se gerar valores amostrais de  $X$  com distribuição triangular de moda  $b$ , valor mínimo  $a$  e valor máximo  $c$ , a partir de um valor amostral  $u$  de uma variável uniformemente distribuída no intervalo  $[0, 1]$ , diga-se variável  $U$ , fazendo:

$$x = a + \sqrt{u(b-a)(c-a)}, \quad \text{para } 0 \leq u \leq \frac{b-a}{c-a}, \text{ e}$$

$$x = c - \sqrt{(1-u)(c-b)(c-a)}, \quad \text{para } \frac{b-a}{c-a} \leq u \leq 1.$$

Uma vez definidos  $a$ ,  $b$  e  $c$ , podem-se, portanto, facilmente gerar valores da variável aleatória  $X$ . Porém, caso se disponha apenas de valores extremos ( $a$  e  $c$ ) e do valor médio  $m$ , pode-se determinar o terceiro parâmetro  $b$ , fazendo:

$$b = 3\mu - a - c.$$

A expressão acima advém do fato de que, se  $X$  tem distribuição triangular com parâmetros  $a$ ,  $b$  e  $c$ , então a sua esperança matemática será dada por:

$$E(X) = \int_a^c x.f(x)dx = K_1 \left( \frac{b^3}{3} - \frac{ab^2}{2} + \frac{a^3}{6} \right) + K_2 \left( \frac{c^3}{6} - \frac{cb^2}{2} + \frac{b^3}{3} \right) = \frac{a+b+c}{3}$$

em que:

$$K_1 = \frac{2}{(b-a)(c-a)} \text{ e } K_2 = \frac{2}{(c-b)(c-a)}$$

Alternativamente, os valores para  $a$  e  $c$  poderão

ser determinados em função do valor médio definido pela tendência modelada, o que facilita o estudo de diversos cenários, mesmo em um futuro distante, desde que o modelo representando a tendência o permita.

## 2.2. Exemplo de aplicação

Conforme mencionado na parte introdutória do trabalho, um dos principais objetivos foi propor um novo método de geração de amostras aleatórias com vistas às análises de risco e geração de cenários em modelos de planejamento florestal, bem como avaliar a eficiência do método proposto em relação ao uso do valor médio  $\mu$ , ou da equação  $Y_k = \mu + \epsilon_k$ . Será demonstrado que, em situações em que a variável aleatória segue alguma tendência ao longo do tempo, utilizar apenas o valor médio  $\mu$ , ou a equação  $Y_k = \mu + \epsilon_k$ , para simular valores futuros pode levar a erros de avaliação significativos. Esses erros podem se tornar tanto mais significativos quanto mais bem definida for a tendência da variável aleatória em relação ao tempo.

Para verificar essa tese foram preparados diversos conjuntos de dados, de modo que houvesse valores (chamados aqui de “valores futuros”) seguindo diferentes tendências. Em seguida, assumindo que esses “valores futuros” gerados fossem reais, compararam-se os valores simulados por três diferentes métodos de simulação com os valores futuros gerados para as diferentes tendências ao longo do tempo.

### 2.2.1 Definição dos “valores futuros”

A variável aleatória escolhida para estudo foi o preço da celulose. A escolha dessa variável deveu-se à sua importância em processos de planejamento em um ramo importante do setor florestal, que é a produção de madeira para a indústria de papel e celulose. A partir de uma série histórica de preço de celulose<sup>4</sup> (período de 1963 a 1999), num total de 37 anos, ajustou-se o seguinte modelo de preço<sup>5</sup>:

$$P_{cel,k} = \beta_0 + \beta_1 T_k + \beta_2 P_{def,k} + \epsilon_k, \quad (k = 1 \text{ a } 37) \quad (m.1)$$

em que:

$P_{cel,k}$  = preço da celulose no tempo  $k$ ;

$T_k$  = variável tendência;

<sup>4</sup> Fonte: dados da FAO, retirados da home page <http://apps.fao.org>.

<sup>5</sup> O modelo m.1 ajustado não deve ser encarado como um modelo de projeção definitivo para obter valores futuros do preço da celulose. Esse modelo foi ajustado com o simples objetivo de produzir um conjunto inicial de estimativas para os parâmetros  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  que pudesse servir como referência para o processo de simulação proposto.

$P_{def,k}$  = preço defasado da celulose no tempo  $k$ , ou seja, o preço da celulose no tempo  $k-1$ ;

$\beta_j$  = parâmetros da regressão ( $j = 0$  a  $2$ ); e

$\varepsilon_k$  = erro aleatório associado ao preço da celulose no tempo  $k$ .

As estimativas dos parâmetros do modelo m.1, mais ou menos seus desvios-padrão, foram<sup>6</sup>:  $\hat{\beta}_0 = 34,7854 \pm 24,6296$ ;  $\hat{\beta}_1 = 3,1029 \pm 1,9289$ ;  $\hat{\beta}_2 = 0,7105 \pm 0,1306$ .

A equação apresentou ainda um coeficiente de determinação ( $R^2$ ) igual a 82,12% e um desvio padrão residual ( $s_{y,x} = \hat{\sigma} = \sqrt{QM Res.}$ ) igual a  $\pm$  US\$ 68,61.

A equação apresentada anteriormente representa a tendência para o preço da celulose em relação ao tempo e ao preço defasado, no período definido de 1963 a 1999. Para efeito deste trabalho, assumiu-se que essa tendência poderia se manter para o futuro, bem como se tornar mais forte ou mais fraca, ou mesmo nula. Foram definidos, então, diferentes tipos de tendências para testar a sensibilidade dos métodos de simulação, apresentados posteriormente, aos diferentes tipos de tendência. Assim, os parâmetros a serem usados para representar as tendências para estudo foram os seguintes:

$\beta_0^* = 35$ ;  $\beta_1^* = 0, 1, 3, 5$ ;  $\beta_2^* = 0; 0,5; 0,7; 0,9$ ; e  $\sigma^* = 30, 60, 90$ .

Esses valores propostos foram baseados nas estimativas dos parâmetros do modelo m.1. e nos respectivos desvios-padrão, além da tendência nula, representada pelo valor 0. Como exemplo, no caso  $\beta_1^* = 0, 1, 3, 5$  o valor 0 representa a tendência nula; o valor 3 representa um valor aproximado da estimativa obtida para o parâmetro  $\beta_1$ , e os valores 5 e 1 representam um valor aproximado da estimativa obtida para o parâmetro  $\beta_1$ , mais ou menos um desvio-padrão da estimativa, respectivamente. O mesmo raciocínio foi utilizado no caso de  $\beta_2^*$ . No caso de  $\beta_0^*$ , assumiu-se que ele seria o mesmo para todas as tendências testadas, sendo um valor próximo ao da estimativa do parâmetro  $\beta_0$ . Em relação ao erro-padrão ( $\sigma^*$ ), decidiu-se utilizar um valor próximo ao estimado, ou seja, de 60 unidades, variando-se 30 unidades para baixo e para cima, obtendo-se, assim,

$\sigma^* = 30, 60, 90$ . Com isso, pretendeu-se estudar a sensibilidade dos métodos de simulação sob o efeito de tendências em diferentes escalas de precisão. Cada combinação dos valores propostos para cada um dos parâmetros gera uma possível tendência futura, possibilitando, ao final, o estudo de 48 tipos diferentes de tendências (ou cenários).

Depois de definidos os 48 cenários possíveis, foram gerados 30 “valores futuros” para cada cenário, ou seja, 48 conjuntos de 30 “valores futuros”, por meio da equação que se segue:

$$P_{cel,k} = \beta_0^* + \beta_1^* T_k + \beta_2^* P_{def,k} + \varepsilon_k, \quad (k = 1 \text{ a } 30)$$

em que:

$P_{cel,k}$ ,  $P_{def,k}$ ,  $T_k$ ,  $\beta_0^*$ ,  $\beta_1^*$  e  $\beta_2^*$ , conforme definidos anteriormente; e

$\varepsilon_k \sim N(0, \sigma^{*2})$ , sendo  $\sigma^*$  conforme já definido.

Em seguida, dividiu-se cada um dos 48 grupos de 30 “valores futuros” ( $P_{cel,k}$ ,  $k = 1$  a  $30$ ) em dois grupos, conforme a seguir:

**Grupo 10 : 20**  $\Rightarrow$  neste grupo, foram utilizados 10 “valores iniciais” ( $P_{cel,i}$ ,  $i = 1$  a  $10$ ) para obtenção das estimativas dos parâmetros a serem usados nos métodos de simulação, simulando-se, então, os 20 “valores finais” ( $P_{cel,f}$ ,  $f = 11$  a  $30$ ).

**Grupo 20 : 10**  $\Rightarrow$  neste grupo, utilizaram-se 10 “valores iniciais” ( $P_{cel,i}$ ,  $i = 1$  a  $20$ ) para obtenção das estimativas dos parâmetros a serem usados nos métodos de simulação, simulando-se, então, os 10 “valores finais” ( $P_{cel,f}$ ,  $f = 21$  a  $30$ ).

Esses grupos foram criados com o intuito de mostrar que a simulação torna-se mais eficiente quando as estimativas dos parâmetros são obtidas de um conjunto de dados maior (uma série histórica maior, por exemplo) ou de modo equivalente, quando a informação prévia, ou subjetiva (com base na experiência do tomador de decisão da empresa), é boa no sentido de melhor especificar a tendência futura. Assim, a expectativa era de que, ao utilizar 10 dados iniciais para simular os 20 valores finais, os métodos de simulação seriam menos eficientes do que o contrário.

<sup>6</sup> As estimativas obtidas não devem ser consideradas para avaliar a significância ou não dos parâmetros do modelo m.1 utilizado. Essas estimativas devem ser vistas simplesmente como um ponto de partida para a definição dos novos valores para os parâmetros do modelo a serem usados nas simulações, conforme discutido no texto.

### 2.3. Métodos de simulação

Foram utilizados três métodos ( $M_1$ ,  $M_2$  e  $M_3$ ) para simulação dos “valores finais”. Os métodos  $M_1$  e  $M_2$  são aqueles encontrados na literatura e o método  $M_3$ , o proposto. Os métodos de simulação foram definidos como se segue:

$M_1 \Rightarrow$  uso da média dos “valores iniciais” ( $P_{cel,i}$ ,  $i = 1$  a  $10$  ou  $i = 1$  a  $20$ ), para simular os “valores finais”. Assim, os “valores finais” simulados ( $P_{cel,f}$ ,  $f = 11$  a  $30$  ou  $f = 21$  a  $30$ , respectivamente) são obtidos a partir do modelo:

$$P_{cel,f} = E(P_{cel,i}) = \mu \quad (m.2)$$

sendo  $\mu$  estimado pela média aritmética dos “valores iniciais”.

$M_2 \Rightarrow$  os “valores iniciais” ( $P_{cel,i}$ ,  $i = 1$  a  $I$ , e  $I = 10$  ou  $20$ ) fornecem média, valor máximo e valor mínimo. Assim, os “valores finais” simulados ( $P_{cel,f}$ ,  $f = I+1$  a  $30$ ) são obtidos a partir do modelo:

$$P_{cel,f} = \mu + \varepsilon_f \quad (m.3)$$

sendo  $\mu$  estimado por  $\bar{X}$  e  $\varepsilon_f$  gerado a partir de uma distribuição triangular com estimativas dos parâmetros, correspondendo aos valores máximo, médio e mínimo dos “valores iniciais”.

$M_3 \Rightarrow$  os “valores iniciais” ( $P_{cel,i}$ ,  $i = 1$  a  $I$ , e  $I = 10$  ou  $20$ ) fornecem: valor máximo e valor mínimo,  $\hat{\beta}_0^*$ ,  $\hat{\beta}_1^*$  e  $\hat{\beta}_2^*$ . Assim, os “valores finais” simulados ( $P_{cel,f}$ ,  $f = I+1$  a  $30$ ) são obtidos pela seguinte equação:

$$P_{cel,f} = \hat{\beta}_0^* + \hat{\beta}_1^* T_f + \hat{\beta}_2^* P_{def,f} + \varepsilon_f = \mu_f + \varepsilon_f \quad (m.4)$$

Nesse novo método,  $\mu_f$  é estimado por uma equação de regressão, que pode incorporar algum tipo de tendência ao processo de simulação de números aleatórios, e  $\varepsilon_f$  é gerado a partir de uma distribuição triangular, em que as estimativas dos parâmetros são obtidas como se segue:  $MAX(P_{cel,i}$ ,  $i = 1$  a  $I$ , e  $I = 10$  ou  $20$ ) = maior valor do preço da celulose dos 10 ou 20 “valores iniciais”;  $MIN(P_{cel,i}$ ,  $i = 1$  a  $I$ , e  $I = 10$  ou  $20$ ) = menor valor do preço da celulose dos 10 ou 20 “dados iniciais”; e  $I = 10$  para o grupo 10 : 20 e  $I = 20$  para o grupo 20 : 10, conforme item 2.1. Portanto, para a simulação dos “valores finais”  $P_{cel,f}$ , fez-se:

$$AT = MAX(P_{cel,i}, i = 1 a I, e I = 10 ou 20) - MIN(P_{cel,i}, i = 1 a I, e I = 10 ou 20),$$

em que AT equivale à amplitude total dos dados iniciais.

Em seguida, foram obtidos os valores mínimo e máximo e a moda, da seguinte forma: valor mínimo =  $\mu_f - AT/2$ ; moda =  $\mu_f$ ; e valor máximo =  $\mu_f + AT/2$ . Portanto, seguindo as expressões m.2, m.3, e m.4, os três métodos a serem avaliados poderiam ser resumidos nas seguintes equações:  $M_1 \Rightarrow P_{cel,f} = \mu$ ;  $M_2 \Rightarrow P_{cel,f} = \mu + \varepsilon_f$ ; e  $M_3 \Rightarrow P_{cel,f} = \mu_f + \varepsilon_f$ .

Após a simulação dos 10 ou 20 “valores finais”, foi calculada a diferença relativa média porcentual entre valores simulados e “valores futuros”, de acordo com a expressão:

$$DRM = \frac{\sum_j \left( \frac{VS_j - VF_j}{VF_j} \right)}{n} \cdot 100,$$

em que:

$DRM$  = diferença relativa média;

$VS_j$  = valor simulado  $j$ ;

$VF_j$  = valor futuro  $j$ ;

$n$  = número de valores simulados, 10 ou 20, dependendo dos grupos já definidos; e

$j = 1$  a  $10$  para o grupo 20:10, ou igual a  $1$  a  $20$  para o grupo 10:20.

Optou-se por fazer mil simulações para cada cenário. No final, foram coletadas:

Porcentagens de vezes em que a DRM do  $M_1$  foi maior do que a do  $M_3$ .

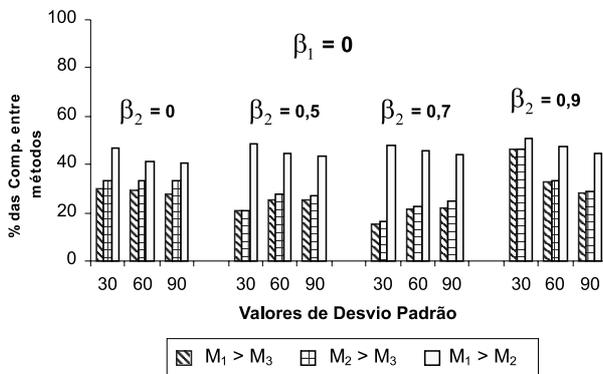
Porcentagens de vezes em que a DRM do  $M_2$  foi maior do que a do  $M_3$ .

Porcentagens de vezes em que a DRM do  $M_1$  foi maior do que a do  $M_2$ .

### 3. RESULTADOS E DISCUSSÃO

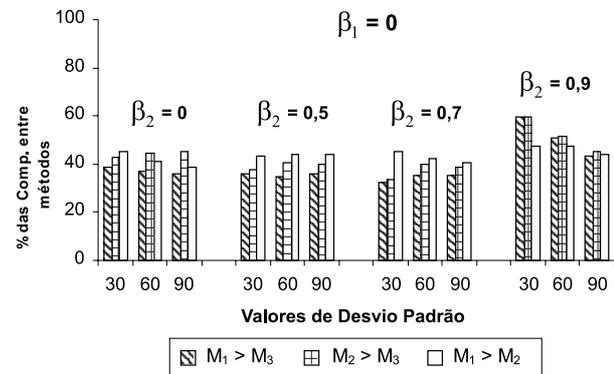
Nas Figuras 1 e 2, mostram-se os resultados das simulações para a situação em que  $\beta_1 = 0$  e  $\beta_2$  varia de 0 a 0,9; deve ser ressaltado que a Figura 1 se baseia nas simulações feitas para o grupo 10:20 e a Figura 2, para o grupo 20:10. O fato de  $\beta_1 = 0$  indica ausência de tendência dos valores futuros, o que faz que os métodos  $M_1$  e  $M_2$  tendam, em princípio, a ser mais eficientes do que o método  $M_3$ , por se basearem no valor médio sem tendência. De acordo com a Figura 1, pode-se observar, pela barra que compara os métodos

$M_1$  e  $M_3$ , no caso em que  $\beta_2=0$ , que em aproximadamente 30% das vezes o  $M_1$  apresentou diferença relativa média (DRM) superior ao  $M_3$  (e, portanto, o  $M_3$  exibiu resultados maiores para a DRM em cerca 70% das simulações). Isso significa que, em apenas 30% das simulações, os valores simulados pelo  $M_1$  ficaram mais distantes dos valores futuros, em média, quando comparados com o  $M_3$ , indicando que os valores simulados pelo  $M_1$ , nessa situação de ausência de tendência, aproximaram-se mais dos valores futuros do que os valores simulados pelo  $M_3$ . Esse resultado já era previsto, uma vez que, na ausência de qualquer tendência, é esperado que os valores simulados se situem em torno de um valor médio sem qualquer tendência, como pressupõem os métodos  $M_1$  e  $M_2$ . Pode-se observar também que, para o  $\beta_2=0,9$ , o que representa a máxima tendência possível para  $\beta_2$ , o desempenho do  $M_3$  melhora tanto em relação ao  $M_1$  quanto ao  $M_2$ . Isso ocorre porque, na presença de tendência, o método  $M_3$  tende a ser mais preciso em prever o futuro, principalmente quando a variabilidade é reduzida. Essa reduzida variabilidade, na prática, significa que existe maior controle da parte estocástica do processo, ou seja, o ruído. Portanto, quando a distribuição da variável em questão for bem conhecida e apresentar baixa variabilidade, melhor será o resultado das simulações com o uso do método proposto.



**Figura 1** – Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 0$ , no grupo 10:20.

**Figure 1** – Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 0$ , in the group 10:20.



**Figura 2** – Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 0$ , no grupo 20:10.

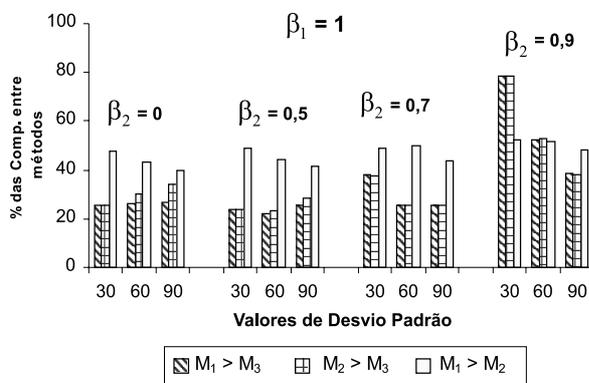
**Figure 2** – Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 0$ , in the group 20:10.

A comparação entre  $M_2$  e  $M_3$  é bastante semelhante àquela feita entre os métodos  $M_1$  e  $M_3$ , indicando haver semelhança de desempenho entre  $M_1$  e  $M_2$ . Tal fato se confirma quando se comparam os métodos  $M_1$  e  $M_2$ , ou seja, a barra que define essa comparação se situa sempre próxima aos 50%, indicando que não há vantagem de um método em relação ao outro. A aparente equivalência entre o  $M_1$  e o  $M_2$  se deve ao fato de que ambos os métodos se baseiam no valor médio sem tendência, embora o  $M_2$  considere distúrbios em torno da média e possa, portanto, ser usado nas análises de risco e geração de cenários em modelos de planejamento florestal, quando não existir forte presença de tendência na variação dos valores da variável em estudo.

Um fato importante a ser mencionado foi observado na Figura 2. Nessa figura, a diferença de desempenho entre os métodos  $M_1$  e  $M_3$  e  $M_2$  e  $M_3$  é bem menor do que aquela observada na Figura 1, ou seja, situando-se sempre próximo dos 40%, exceto para o maior valor de  $\beta_2$ . Isso significa que para o caso do grupo 20:10, representado pela Figura 2, são feitas melhores estimativas dos parâmetros, evidenciando-se que, ao utilizar 20 valores iniciais para prever 10 valores futuros, simulações mais confiáveis são esperadas. Para justificar

esse fato, considerou o modelo de regressão no qual se baseou o  $M_3$ :  $P_{cel,f} = \beta_0^* + \beta_1^* T_f + \beta_2^* P_{def,f} + \epsilon_f$ . Nesse modelo, caso  $\beta_1 = 0$  e  $\beta_2 = 0$  (como prevê a primeira situação das Figuras 1 e 2), o  $P_{cel,f}$  será igual ao  $P_{cel,i}$  médio sem qualquer tendência, fazendo que o  $M_3$  se iguale teoricamente ao  $M_1$  e  $M_2$ . Essa proximidade teórica está mais bem estabelecida na Figura 2, que representa o grupo 20:10, justificando, portanto, que nesse caso foram feitas melhores estimativas dos parâmetros. O fato de no grupo 10:20 os resultados terem ficado mais distantes do teoricamente esperado provavelmente se deve ao pequeno número de dados utilizados na definição da tendência, fazendo que esta se afastasse do valor médio sem qualquer tendência, que seria a verdadeira tendência esperada. Esse resultado, apesar de ser teoricamente esperado, foi obtido para mostrar que, na prática, poder-se-ia inferir que, ao conhecer bem a variável em estudo de modo a modelar da melhor maneira possível sua tendência, poderiam ser obtidos bons resultados usando o método proposto. Como será discutido nas figuras seguintes, se a tendência for mais definida ainda, melhor será o resultado com o uso do método proposto.

Nas Figuras 3 e 4, o que muda é que  $\beta_1$  passa a ser igual a 1, havendo, portanto, acréscimo de tendência

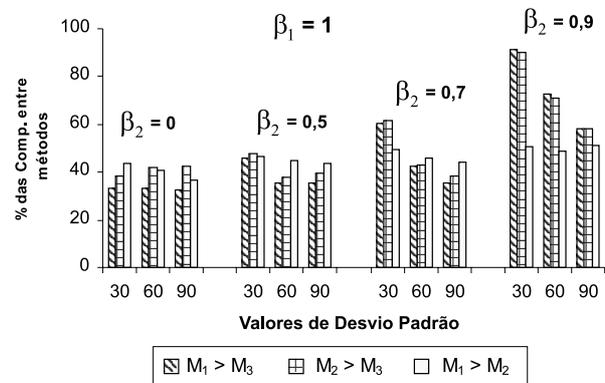


**Figura 3** – Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 1$ , no grupo 10:20.

**Figure 3** – Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 1$ , in the group 10:20.

nos valores a serem previstos pelos métodos de simulação. Pode-se notar, claramente, aumento no desempenho do método  $M_3$  em relação ao  $M_1$  e  $M_2$ , à medida que  $\beta_2$  aumenta. Essa situação fica ainda mais clara na Figura 4, que representa o grupo 20:10, mostrando que nesse grupo as estimativas dos parâmetros são realmente mais confiáveis. Os métodos  $M_1$  e  $M_2$  voltam a mostrar que são realmente equivalentes, apresentando sempre valores próximos a 50% de vantagem de um em relação ao outro.

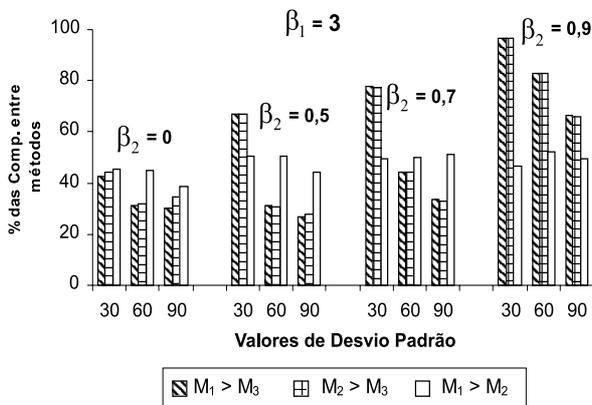
Nas Figuras 5 e 6, em que  $\beta_1 = 3$ , ou seja, já há tendência bem definida dos valores futuros, fica mais clara ainda a superioridade do  $M_3$ , principalmente, para maiores valores de  $\beta_2$ . Nesse momento, cabe retornar a discussão sobre a influência dos desvios-padrão (30, 60 e 90) no desempenho do  $M_3$ . Fica evidente que, para menores valores de desvio-padrão, o desempenho do  $M_3$  é bem superior. Visto que valores menores de desvio-padrão indicam tendências mais bem definidas, quanto maior a variabilidade dos dados que fornecerão estimativas dos parâmetros a serem usados no  $M_3$ , ou quanto menos preciso for o conhecimento subjetivo do tomador de decisão na definição dos valores máximo, mínimo e médio necessários para o uso do método  $M_3$ , este método fica prejudicado, e o método  $M_2$  poderia, então, ser usado.



**Figura 4** – Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 1$ , no grupo 20:10.

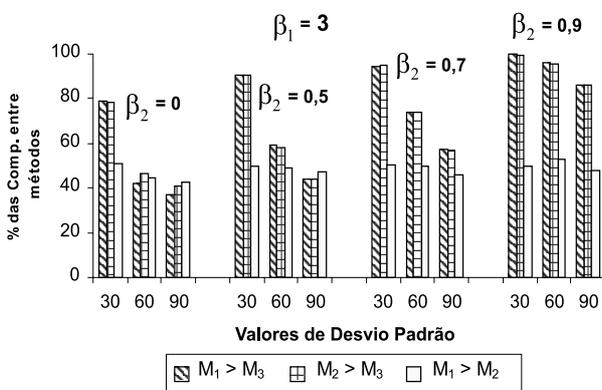
**Figure 4** – Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 1$ , in the group 20:10.





**Figura 5**—Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 3$ , no grupo 10:20.

**Figure 5**—Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 3$ , in the group 10:20.

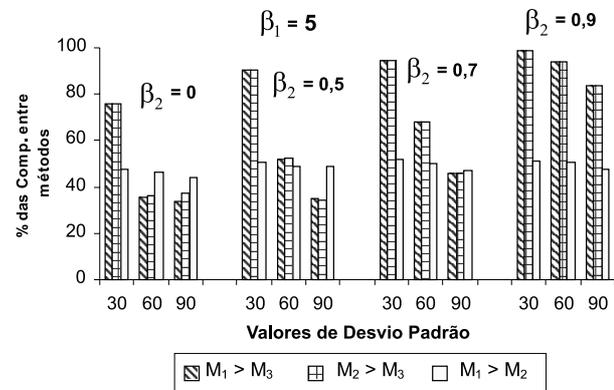


**Figura 6**—Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 3$ , no grupo 20:10.

**Figure 6**—Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 3$ , in the group 20:10.

No entanto, como o  $M_3$  se baseia na existência de alguma tendência, quanto mais evidenciada ela for, melhor será o seu desempenho, mesmo com a maior variabilidade. Caso exista boa definição da tendência futura, ainda que ela seja nula, o método  $M_3$  se torna superior e deveria ser usado nas análises de risco e geração de cenários em modelos de planejamento florestal.

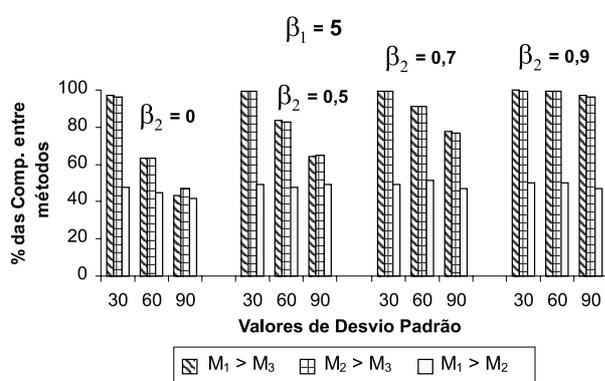
As Figuras 7 e 8, em que  $\beta_1 = 5$ , representam a situação de tendência mais forte para os valores futuros. Como ocorreu nas Figuras 5 e 6, o  $M_3$  teve seu desempenho bem superior aos métodos  $M_1$  e  $M_2$ , principalmente para maiores valores de  $\beta_2$  e menores de desvio-padrão, chegando praticamente a 100% de superioridade. Essa superioridade fica mais bem evidenciada na Figura 8, que representa o grupo 20:10, pelos mesmos motivos já apresentados anteriormente. Da maneira geral, alguns resultados importantes foram encontrados em todas as situações testadas. Dentre esses, cabe destacar a importância de se considerarem tendências, caso estas existam, em processos de simulação, o que ficou comprovado nas situações de maiores valores de  $\beta_1$  e  $\beta_2$  e menores desvios-padrão.



**Figura 7**—Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 5$ , no grupo 10:20.

**Figure 7**—Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 5$ , in the group 10:20.

Nesse sentido, o  $M_3$  levou vantagem em relação aos métodos  $M_1$  e  $M_2$ , que não fazem nenhuma consideração a tendências futuras. Outro fato verificado foi a importância de se utilizarem bons métodos de determinação de tendências para que as simulações sejam confiáveis. Isso ficou demonstrado nas comparações entre os grupos 10:20 e 20:10 e também pelos menores valores de desvio-padrão, em que ficou evidenciado melhor desempenho do  $M_3$ . Entretanto, cabe destacar que novos estudos precisam ser feitos, buscando alternativas apropriadas para considerar possíveis tendências em processos de simulação.



**Figura 8** – Comparações entre os métodos  $M_1$ ,  $M_2$  e  $M_3$ , conforme legenda, em que  $M_1 > M_3$  significa a porcentagem de vezes em que a Diferença Relativa Média do  $M_1$  foi maior do que a do  $M_3$ , e assim sucessivamente, nas mil simulações, em função dos desvios-padrão,  $\beta_2$ , e para  $\beta_1 = 5$ , no grupo 20:10.

**Figure 8** – Comparisons among methods  $M_1$ ,  $M_2$  and  $M_3$ , according to the legend.  $M_1 > M_3$  means the percentage of times the Average Relative Difference of  $M_1$  was greater than that of  $M_3$ , and then successively, after 1,000 simulations, as a function of the standard deviations,  $\beta_2$ , and for  $\beta_1 = 5$ , in the group 20:10.

#### 4. CONCLUSÕES

Ao final deste trabalho, pode-se concluir que:

Em processos de simulação de dados ao longo do tempo, caso estes apresentem algum tipo de tendência, é importante que esta seja considerada, sob pena de perda de confiabilidade do processo.

Na ausência de tendência, os métodos de simulação  $M_1$  ( $P_{cel,f} = \mu$ ) e  $M_2$  ( $P_{cel,f} = \mu + \varepsilon_f$ ) mostraram-se satisfatórios, embora o método  $M_2$  apresente a vantagem de considerar distúrbios em torno da média, o que pode ser útil para geração de cenários e análises de risco.

A precisa definição, ou modelagem, da tendência futura é fundamental para o sucesso do método  $M_3$ , em que  $M_3$  é igual a  $P_{cel,f} = \hat{\beta}_0^* + \hat{\beta}_1^* T_f + \hat{\beta}_2^* P_{def,f} + \varepsilon_f = \mu_f + \varepsilon_f$

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

BUSTOS, O.H.; ORGAMBIDE, A.C.F. Simulação estocástica: teoria e algoritmos. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 10., 1992, São Paulo, **Anais ...** São Paulo: 1992. 152 p.

CASELLA, G.; BERGER, R. L. **Statistical Inference**. 2. ed. Pacific Grove: Duxbury Press, 2002. 660 p.

HOF, J.; KENT, B.; PICKENS, J. Chance constraints and chance maximization with random yield coefficients in renewable resource optimization. **Forest Science**, v. 38, n. 1, p. 305-323, 1992.

LAW, A. M.; KELTON, W. D. **Simulation Modeling and Analysis**. 2. ed. New York: McGraw-Hill. 1991. 759 p.

PICKENS, J.; DRESS, P. Use of stochastic production coefficients in linear programming models: objective function distribution, feasibility and dual activities. **Forest Science**, v. 34, n. 3, p. 574-591, 1988.

PROTIL, R. M. Aplicação da técnica de simulação discreta estocástica na avaliação de risco em investimentos florestais. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 32, 2000, Viçosa. **Anais ...** Viçosa, MG: SOBRAPO, 2000. p. 953-961.

RODRIGUES, F.L. et al. Regulação de florestas equiâneas utilizando programação linear: uma aplicação da teoria do modelo II. **Revista Árvore**, v. 22, n. 2, p. 193-213, 1998.

RODRIGUES, F.L. et al. Determinação de estratégias de reforma, condução da brotação e compra de terras, utilizando programação linear. **Revista Árvore**, v. 23, n. 2, p. 169-186, 1999.

RODRIGUEZ, L.C.E. **Planejamento agropecuário através de um modelo de programação linear não determinista**. 1987. 83 f. Tese (Mestrado em Economia Agrária) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 1987.

RODRIGUEZ, L.C.E.; BORGES, J.G. Técnicas matemáticas para determinação de níveis sustentáveis de produção florestal: um exemplo em eucaliptal **Revista Florestal**, Lisboa, Portugal, v. 12, n. 1/2, p. 83-91, 1999.

SCOLFORO, J.R.S. Sistema integrado para predição e análise presente e futura do crescimento e predição, com otimização de remuneração de capitais, para *Pinus caribaea* var *hondurensis*. 1991. 290 f. Tese (Doutorado em Ciências Florestais) - Universidade Federal do Paraná, Curitiba, 1991.

SILVA, G.F.; RODRIGUES, F.L.; SANTOS, H.N. Um sistema de apoio a decisão para pequenas e médias empresas florestais. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 32., 2000, Viçosa. **Anais ...** Viçosa, MG: SOBRAPO, 2000. p. 953-961.

SOUZA, A.N. **Estudo econômico da reforma de povoamentos de *Eucalyptus* spp. - O caso do progresso tecnológico**. 1999. 139 f. Tese (Mestrado em Ciência Florestal) - Universidade Federal de Lavras, Lavras, 1999.

VOLPI, N.M.P.; CARNIERI, C.; SANQUETA, C.R. O impacto da estocasticidade das informações em um modelo de planejamento florestal. **Revista Árvore**, v. 12, n. 2, p. 100-110, 1999.

WAGNER, H.M. **Pesquisa operacional**. 2.ed. São Paulo: 1986. 851p.