

Modelos generalizados para a mortalidade de árvores de *Eucalyptus grandis* no Estado de São Paulo, Brasil

Generalized models for *Eucalyptus grandis* trees mortality in the São Paulo State, Brazil

María del Pilar Díaz
Hilton Thadeu Zarate do Couto

RESUMO: Na área florestal, a estimativa da mortalidade das árvores de uma espécie deve estar presente nos sistemas confiáveis de predição do crescimento. Porém, quando a soma de dados dessa variável é feita mediante um processo de classificação dicotômica, originam-se observações de contagem nas unidades experimentais, sendo muito freqüente a presença de superdispersão no modelo probabilístico. Este trabalho apresenta uma nova metodologia de ajuste de modelos para avaliar o comportamento da mortalidade de árvores de *Eucalyptus grandis* no Estado de São Paulo, incorporando informação de talhadia da espécie e da qualidade do sítio correspondente a três municípios. Através de técnicas formais e informais, identifica-se o modelo probabilístico e sob o enfoque dos modelos lineares generalizados, estimam-se os modelos Extra-Poisson de Breslow (1984) e Binomial Negativo (BN) pelos métodos de Quase-Verossimilhança e Máxima Verossimilhança, respectivamente, para a avaliação dos efeitos de interação entre os fatores. Os resultados mostraram superdispersão altamente significativa, indicando extra-variação quadrática, devida possivelmente à falta de independência entre as observações (processo ecológico de composição do *stand* florestal). Os modelos Extra-Poisson e BN evidenciaram que há interação significativa entre os fatores, indicando que nos municípios de Brotas e Ribeirão Bonito as plantas suportam melhor o estresse do primeiro corte.

PALAVRAS-CHAVE: Mortalidade, Superdispersão, Modelos lineares generalizados, *Eucalyptus grandis*

ABSTRACT: In the forestry area, a tree mortality estimation for a given species must be present in the reliable prediction systems for the species growth. However, the individual mortality study introduces a dichotomic process, arising to count data set for the experimental units and frequently the stochastic model has clustered behavior, called overdispersion. This work presents a new fitting approach models in order to know the behavior of trees mortality of *Eucalyptus grandis*, which is important genus in the Brazilian forestry production, in the São Paulo State, including information related to species rotation e quality site from three districts. Through formal and informal technics the probabilistic model were identified and under the generalized linear models approach the Extra-Poisson and Negative Binomial models were estimated by means of Quasi-Likelihood and Maximum-Likelihood methods, respectively, to evaluate the interaction between

the two factors. The results showed a highly significative overdispersion pointing out a quadratic extra-variation due to the dependence between the observations (composition process in the forestry stand). The Extra-Poisson and Negative Binomial models proved significative interaction effect showing that the Brotas and Ribeirão Bonito districts are the best places to bear the first court stress.

KEYWORDS: Mortality, Overdispersion, Generalized linear models, *Eucalyptus grandis*

INTRODUÇÃO

Na Ciência Florestal e dentro da área de manejo para fins de produção de madeira e definição dos regimes de manejo, segundo Somers et al. (1980), o estudo da combinação rendimento-densidade tornou-se o objetivo de intensas pesquisas durante as últimas décadas. O estudo dessa combinação, como processo, pode ser dividido em vários componentes interrelacionados, dos quais, o crescimento real (sobrevivência) e a mortalidade são os mais importantes visto que a informação do crescimento final constitui a parte essencial do manejo florestal.

Assim, o estudo do comportamento da mortalidade das árvores deve estar presente na construção dos sistemas de predição do crescimento da espécie. Quando em termos metodológicos a mortalidade de indivíduos é avaliada mediante um processo de classificação dicotômica (isto é, considerando a resposta como zero ou um, vivo ou morto conforme os planos de inventário contínuos para fins específicos), originando dados de contagem nas unidades experimentais dos ensaios experimentais ou estudos observacionais da área florestal.

Além disso, os sistemas atuais de inventário florestal estabelecem como necessária uma adequada estratificação do material, considerando para isso fatores como seleção da espécie, rotação, procedência, idade, produtividade, espaçamento e local, para a obtenção de regras gerais de manejo.

Dentro do enfoque dos modelos lineares generalizados (MLG), como uma extensão da teoria Normal para a modelação, dados de contagem podem ser assumidos com variação aleatória, isto é, seguindo uma distribuição Poisson (McCullagh e Nelder, 1989). A grande vantagem deste enfoque é tratar a variável em sua natureza estocástica original, evitando, portanto, trabalhar com proporções ou porcentagens para a utilização da metodologia estatística usual.

Porém, um fenômeno que ocorre nessa estrutura de dados é a superdispersão, ou seja, quando a $Var(Y_i)$, (variância empírica ou amostral) excede a variância nominal, (variância esperada conforme um modelo probabilístico estabelecido).

O objetivo deste trabalho é apresentar uma metodologia de análise estatística para o ajuste de modelos simples e interpretáveis que forneçam estimativas confiáveis do comportamento da mortalidade de árvores de *E. grandis* no Estado de São Paulo. Para isso, identifica-se o componente probabilístico dessa variável e, sob o enfoque dos MLG's, estabelecem-se os outros dois componentes, o sistemático e a função de ligação, incorporando também informação da situação experimental do *stand* florestal.

A metodologia de ajuste proposta, utiliza os modelos Extra-Poisson e Binomial Negativo e baseia-se na implementação de algoritmos para a obtenção das estimativas de quase e máxima verossimilhança dos parâmetros de interesse.

REVISÃO DE LITERATURA

Dentro do enfoque dos modelos lineares generalizados, um dos componentes que o definem é o aleatório. Esse componente refere-se à distribuição das observações, isto é, qual o modelo probabilístico a propor para gerenciar o comportamento da geração do conjunto de dados. Assim, quando um modelo linear generalizado é ajustado, pode acontecer que as medidas de discrepância criadas para observar a adequação de dito modelo superem o limite estabelecido pela teoria estatística. Um dos motivos da rejeição do modelo proposto é a presença de superdispersão. Díaz (1996) apresenta os diferentes origens desse fenômeno para o caso de observações de contagem.

Firth (1990), McCullagh e Nelder (1989), entre outros, observaram que a superdispersão está quase sempre presente nos dados reais, sendo uma das possíveis causas a dependência entre as observações. Se uma extra-variável estiver presente na distribuição significa que existe desconhecimento de informação, talvez devido à correlação positiva entre as respostas dentro de um grupo experimental.

Porém, se a superdispersão é ignorada, a eficiência nas estimativas pontuais não é garantida, pois os erros padrão terão grandeza menor (Díaz, 1996) e, portanto, as inferências podem ser extraídas em forma errada.

Para explorar a estrutura média-variância podem destacar-se os gráficos de algumas estatísticas (ou funções) de posição e dispersão, o teste do quociente média-variância de Elliot (1979), ou o teste para verificar alguma relação pré-estabelecida entre a média e a variância, como o teste de Dean e Lawless (1989) para o Modelo Poisson.

Um dos enfoques analíticos mais comuns na literatura para tratar dados de contagem com superdispersão é aquele apresentado por Breslow (1984), e denotado por QV/M (na sequência, Quase-Verossimilhança-Momentos). O mesmo consiste em usar somente a estrutu-

ra da média-variância estabelecida, estimando os b's do modelo de locação estabelecido (componente sistemático do MLG) pelo método de quase-verossimilhança (Wedderburn, 1974) e o parâmetro de dispersão pelo método dos momentos. Elder (1995); Bennett (1989); Breslow (1990) e Díaz (1994), entre outros, utilizaram esse enfoque para análise de Bioensaios.

O enfoque de Breslow (1984) estabelece uma distribuição marginal para a variável aleatória Z tal que o valor esperado é $\mu = \mu(x\beta)$ e a variância $V(\mu; \phi) = \mu(x, \beta) + \phi\mu^2(x, \beta)$, onde ϕ é o parâmetro de superdispersão. Díaz (1994, 1996) descreve alguns exemplos clássicos, apresenta a expressão do sistema de equações de máxima quase-verossimilhança e discute as propriedades assintóticas das estimativas, mostrando a eficiência relativa das mesmas em relação às de máxima verossimilhança para um modelo com variância quadrática como o Binomial Negativo.

O modelo Binomial Negativo pode ser construído para representar a dependência entre as respostas. Lindsey (1993) apresenta esse modelo gerado por uma distribuição composta, o que é chamado de processo de composição. Isto é, o parâmetro de locação m é uma variável aleatória, independentemente da distribuição das observações, que para o caso da mortalidade das árvores poderia quantificar heterogeneidade entre os indivíduos.

Em Díaz (1996) mostra-se a obtenção da distribuição conjunta das respostas e do parâmetro aleatório, propondo como candidata para a distribuição do m a distribuição Gama, como o sugerem Jorgensen (1987) e Lindsey (1993). Esse enfoque não é muito utilizado na literatura dado que a estimação dos parâmetros b's contidos no modelo de locação m através de máxima verossimilhança implica na resolução de um sistema de equações não lineares de complexa implementação nos pacotes estatísticos tradicionais.

METODOLOGIA

Os dados de inventário florestal contínuo com parcelas permanentes foram fornecidos pela Champion de Papel e Celulose Ltda. (São Paulo - Brasil), e correspondem à espécie *Eucalyptus grandis*, classificados conforme os regimes de alto fuste e de talhadia (primeira e segunda talhadia) e três locais ou qualidades de sítios, provenientes de três regiões geográficas diferentes no Estado de São Paulo. O sistema de inventário estabelece uma estratificação do material, ou seja, agrupamentos dos maciços em estratos que podem ser considerados como unidades uniformes para fins de manejo e produção.

O local foi escolhido de forma a refletir a diversidade ambiental (solo-clima). Esses sítios formam um triângulo onde o vértice superior está orientado ao norte. Aproximadamente em uma linha reta leste-oeste, estão os sítios chamados Regiões 1 (municípios de Mogi-Guaçu, Aguaí, Casa Branca e Espírito Santo de Pinhal, SP) e Região 2 (municípios de Brotas e Ribeirão Bonito, SP), e para o norte, os municípios de Altinópolis, São Simão, Santa Rita do Passa Quatro e Santa Rosa de Viterbo, SP) compõem a Região 3.

Para a determinação da área total relacionada ao experimento e às técnicas de amostragem, seguiram-se as normas de inventário florestal adotadas pela empresa. A localização das parcelas foi aleatória. O tamanho de parcela utilizado foi 20 x 20 metros (400 m²). O total de unidades de amostra foi n=2605 distribuídas em diferentes intensidades em função do fator que referencia a talhadia ou momento de observação, sendo a relação área ocupada - área amostrada de 0.32% para a primeira rotação e 0.64% para as restantes. Em cada parcela foram coletados o número de árvores mortas e a idade média das árvores.

Para estabelecer o modelo linear generalizado escolheram-se três componentes: o alea-

tório ou erro, o sistemático ou modelo de localização e a função de ligação como nexos entre os dois anteriores. Escolhidos esses componentes, o ajuste do MLG para os dados de mortalidade de *E. grandis* (Z_{ijk}) foi implementado como se descreve a seguir, mediante rotinas construídas no pacote GLIM versão 4.08 (Francis et al., 1994).

A parte sistemática do MLG, denotado por η , correspondeu a um modelo da análise da covariância com estrutura fatorial de tratamentos e idade das árvores como covariável, isto é $\eta_{ijk} = \mu + \alpha_i + \gamma_j + \alpha\gamma_{ij} + \beta\chi_{ijk}$ onde μ é uma constante inerente aos dados, α_i é o efeito da i -ésima região, γ_j é o efeito da talhadia, com $i, j=1,2,3$, $\alpha\gamma_{ij}$ é o efeito da interação entre os níveis i e j dos fatores, e β é o coeficiente de regressão para a variável idade, χ_{ijk} , com $k=1, \dots, K(i, j)$. A inclusão da variável explanatória idade foi decidida após o uso de ferramentas gráficas e testes apropriados para diagnósticos (Atkinson, 1982).

Para poder propor um modelo Poisson para as observações, os valores médios da mortalidade foram comparados com seus respectivos desvios padrões, para cada tratamento (combinação dos níveis dos dois fatores), para verificar a existência de uma relação linear entre esses pares de valores, como sugere Elliot (1979).

Foi construído também o gráfico $\ln(\bar{Z}_{ij} - \text{Var}(Z_{ijk})/n)$ versus $\ln(\text{Var}(Z_{ijk}) - \bar{Z}_{ij})$ para explorar o tipo de superdispersão na distribuição, isto é, $\sigma^2 > \mu$. Segundo Elder (1995), se o modelo de extra-variação for o adequado, a tendência nesse gráfico deveria ser próxima a uma reta.

Como teste de aderência para pesquisar a extra-variação no modelo Poisson, e propor

$\text{Var}(Z_i \setminus X_i) = \mu_i + \phi\mu_i^2$, foi utilizada a estatística $T = \frac{1}{2} \sum_{i,j,k} (Z_{ijk} - \hat{\mu}_{ij})^2 - Z_i$, onde

$\hat{\mu}_{ij} = \mu_{ij}(X_{ij}; \hat{\beta})$, com $\hat{\beta}$ o estimador de máxima verossimilhança do b sob o modelo Poisson, apresentada por Dean e Lawless (1989) para testar $H_0 : \phi = 0$ versus $H_1 : \phi > 0$. Os valores dessa estatística para cada tratamento foram calculados no pacote GLIM.

O algoritmo de estimação para o modelo de Breslow (1984) foi também construído nesse pacote estatístico, ajustando os parâmetros do componente sistemático pelo método da quase-verossimilhança (Wedderburn, 1974) e o parâmetro da dispersão, pelo método dos momentos, adaptando o proposto pelo autor para outras versões anteriores do pacote.

Para a obtenção da estimativa do parâmetro de dispersão foi utilizada, iterativamente, a estatística X^2 , mediante

$$\phi = (N - p)^{-1} \sum_{i=1}^n \left\{ \frac{(z_i - \hat{\mu}_i)^2}{\hat{\mu}(\hat{\mu}_i + \phi)} \right\}$$

Assim, dado o modelo linear generalizado para a mortalidade das árvores específica definido segundo:

- ♦ um modelo de análise da covariância como componente sistemático;
- ♦ erro Poisson com extra-variação, e
- ♦ a função de ligação logarítmica,

o algoritmo para o ajuste consiste no seguinte:

1. Ajustar o modelo Poisson conforme a estrutura do MLG. Se a *deviance* residual está próxima dos seus graus de liberdade, finaliza o processo e conclui-se que a variação residual está explicada satisfatoriamente.
2. Resolver (1) recursivamente para f.

3. Definir pesos novos $w_i = (1 + \phi \hat{\mu}_i)^{-1}$ e voltar ao passo [1].

Como análise de diagnóstico, foram observados os valores estimados para os componentes de *deviance* padronizados e estudentizados externamente para a aderência do modelo (Collett, 1991). No pacote, isto resulta operando com os vetores do sistema %di (componentes da *deviance*), %rs (resíduos de Pearson), %le (*leverage*), entre outros.

Para estimar o modelo Binomial Negativo adaptaram-se as macros construídas por Hinde (1996), cujo algoritmo permite a estimação do parâmetro de superdispersão, usando um processo iterativo dentro do processo de ajuste do MLG propriamente dito. O método escolhido para a estimação dos parâmetros do modelo $\eta_{ijk} = \mu + \alpha_i + \gamma_j + \alpha\gamma_{ij} + \beta\chi_{ijk}$ foi o da máxima verossimilhança, já que garante a eficiência das estimativas.

Para esse último ajuste, foram testadas duas funções de ligação, a logarítmica e a raiz quadrada, e selecionado o modelo de melhor performance. Foram construídas em GLIM, e anexadas à adaptação da macro de Hinde (1996), rotinas para cálculo das estatísticas resumidas, após o ajuste, dos resíduos padronizados e estudentizados, obtidos dos componentes da *deviance*, segundo sugere Collett (1991).

Como aplicação da obtenção do modelo apropriado para a mortalidade das árvores de *E. grandis* foram obtidos os tamanhos ótimos de amostras, segundo um experimento futuro com uma determinada precisão, em termos da porcentagem da média amostral.

RESULTADOS E DISCUSSÃO

A Figura 1 ilustra a dispersão das estatísticas propostas por Elder (1995), indicando como candidato um modelo que reflete magnitudes

de dispersão maiores do que as de posição. Portanto, rejeita-se o modelo de dispersão aleatória Poisson.

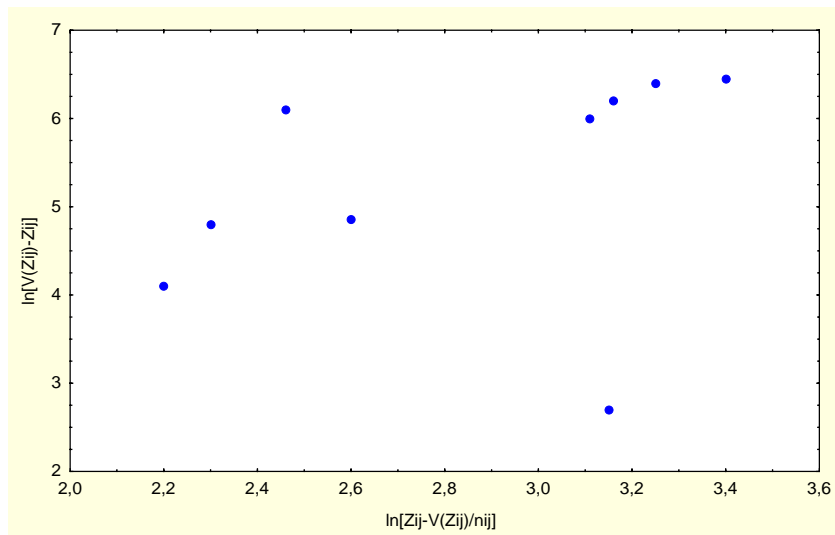


Figura 1. Diagrama de dispersão de $\ln(\bar{Z}_{ij} - V(Z_{ijk})/n_{ij})$ versus $\ln(V(Z_{ijk})\bar{Z}_{ij})$
 (Dispersion Diagram of $\ln(\bar{Z}_{ij} - V(Z_{ijk})/n_{ij})$ versus $\ln(V(Z_{ijk})\bar{Z}_{ij})$)

A Tabela 1 apresenta as médias e as variâncias amostrais para Z_{ijk} em cada tratamento, indicando uma possível relação $\bar{Z}_i = \beta S_i^2$ com coeficiente de regressão $\beta \neq 1$. Os valores das estatísticas propostas por Dean e Lawless (1989) para o teste de superdispersão são mostrados nessa tabela, indicando uma extra-variância aditiva significativa para o modelo Poisson.

A Tabela 2 mostra os resultados do algoritmo construído para o modelo Extra-

Poisson de Breslow (1984), constando os valores para as deviances iniciais (Dev. Ini.), isto é, sem ajustar a superdispersão (com componente aleatório segundo um modelo Poisson), as deviances em cada interação, ajustando a variação extra, com os respectivos graus de liberdade (g.l.), e as estimativas do parâmetro de superdispersão. Os valores estimados para os efeitos e covariável e os seus desvios padrões são mostrados nesta tabela.

Tabela 1. Valores médios, em porcentagem, desvios padrão da mortalidade das árvores de *Eucalyptus grandis* e valores obtidas para as estatísticas de superdispersão (D&L) para os tratamentos.

(Average percentage, standard errors of trees mortality of *Eucalyptus grandis* and overdispersion statistics values for each treatment).

Tratamento	Média da Mortalidade	Desvio Padrão	Estatística D&L	hline
1ª talhadia x Região 1	9.61	11.35	16955.13	(**)
1ª talhadia x Região 2	9.24	8.22	1435.48	(**)
1ª talhadia x Região 3	14.35	11.85	2730.14	(**)
2ª talhadia x Região 1	24.46	22.11	1763.42	(**)
2ª talhadia x Região 2	13.48	21.84	2452.97	(**)
2ª talhadia x Região 3	31.92	26.82	4622.08	(**)
3º momento x Região 1	23.75	20.93	865.02	(**)
3º momento x Região 2	28.68	25.55	234.71	(**)
3º momento x Região 3	41.69	7.71	0.69	(ns)

Tabela 2. Análise da *Deviance* para o modelo Extra-Poisson para a mortalidade das árvores de *Eucalyptus grandis*. Estimativas dos parâmetros, e seus respectivos erros padrão e estatísticas de Wald.

(Analysis of Deviance from Extra-Poisson model of trees mortality of *E. grandis* and parameters estimations, standard errors and Wald Statistics for each effect).

Estatísticas	Ajuste Inicial	1a. Inter	2a. Inter	3a. Inter.(g.l.)
<i>Deviance</i> Residual	33083.	2578.03	2597.54	2596.(2596)
$\hat{\lambda} = \phi$	-	0.8693	0.9299	0.9302
Parâmetro	Estimativa	Desvio Padrão	Valor de t	Efeito
μ	2.0200	0.0362	5.5801	(*)
β	0.1845	0.0096	19.2187	(*)
α_2	-0.1757	0.0658	2.6702	(*)
α_3	0.2567	0.0551	4.6588	(*)
γ_2	0.7369	0.0881	8.3644	(*)
γ_3	0.8145	0.1003	8.1206	(*)
$\alpha\gamma_{(2,2)}$	-0.2646	0.1352	-1.9571	(*)
$\alpha\gamma_{(2,3)}$	0.3552	0.2047	1.7352	(ns)
$\alpha\gamma_{(3,2)}$	0.0815	0.1093	0.7456	(ns)
$\alpha\gamma_{(3,3)}$	0.2232	0.7036	0.3174	(ns)

O índice de agregação foi maior do que 1 (1.0753), mostrando que a extra-variação não é desprezível, logo, a mortalidade de árvores de *E. grandis* segue um modelo Poisson com extra-variação. Ao nível 0.10 de significância, existiu efeito da interação entre os fatores estudados.

Essa dependência estaria evidenciada na combinação da região compreendida pelo município de Brotas e Ribeirão Bonito (Região 2) e a segunda talhadia da espécie. Embora tenha existido dependência entre os fatores, pode-se observar nas estatísticas de Wald que existiu, obviamente, um marcado efeito principal da talhadia.

Os resíduos componentes de deviance padronizados estiveram contidos, principalmente, no intervalo [-3.5364, 2.9045], embora existissem observações com valores de -7.53 e 10.40 para as quais seus baixos valores de leverage não indicaram influências no modelo. O valor da mediana desses resíduos padronizados estimados foi -0.0142 e a distribuição mostrou assimetria à esquerda, indicando que o modelo superestima com frequência.

A Tabela 3 apresenta os resultados do ajuste do Modelo Binomial Negativo para as funções de ligação logarítmica e raiz quadrada. Embora os dois modelos ajustados hajam fornecido resultados muitos semelhantes, foi escolhido o primeiro, o modelo não linear exponencial para o μ . Isto foi devido, entre outras coisas, ao fato de:

- encontrar diferenças no comportamento dos resíduos, onde para a função de ligação raiz quadrada, seus valores resúmenes indicam tendência;
- obter menor precisão nos intervalos de confiança para cada um dos parâmetros de interesse (coeficiente de regressão e efeitos dos tratamentos) com função raiz quadrada. A Tabela 4 apresenta o resumo das estatísticas para os parâmetros.
- ser fácil a interpretação do modelo para o m , i.e. $\hat{\mu} = e^{\hat{\eta}}$

Os resultados mostraram que o modelo da análise da variância com estrutura fatorial pode ser melhorado significativamente, incluindo a informação fornecida pela variável idade das árvores.

Tabela 3. Resultados do ajuste por máxima verossimilhança do modelo Binomial Negativo para a mortalidade de árvores de *E. grandis* para duas funções de ligação.

(Maximum likelihood fitting results from the Negative Binomial model for the trees mortality of *E. grandis*).

Função de Ligação Logarítmica			
Estatísticas	Valor(desvio padrão)	Estatísticas	Valor
$\hat{\lambda}$	0.4579(0.0158)	$\hat{\phi}$	2.1838
Deviance	2903.5600	Deviance/g.l.	1.1200
2xLog-Veros ¹ . (L.V.)	114053. (**)	g.l	2595
[Minr _(di) ² , Maxr _(di) ²]	[-2.6760; 2.0662]	Med _(di) ³ r	-0.0002
Função de Ligação Raiz Quadrada			
Estatísticas	Valor(desvio padrão)	Estatísticas	Valor
$\hat{\lambda}$	0.4547(0.0156)	$\hat{\phi}$	2.1996
Deviance	2900.9000	Deviance/g.l.	1.1178
2xLog-Veros. (L.V.)	114040. (**)	g.l	2595
[Minr _(di) ² , Maxr _(di) ²]	[-4.9683; 6.0315]	Medr _(di)	-0.0135

¹ estatística do teste do quociente de verossimilhança. ² resíduos componentes de deviance padronizados. ³ mediana dos resíduos.

Tabela 4. Estimativas dos parâmetros, seus respectivos erros padrão e estatísticas de Wald para hipótese $H_0: \beta=0$ a partir do modelo Binomial Negativo.

(Parameters estimations, standard erros and Wald Statistics for each effect from Negative Binomial model).

Função de Ligação Logarítmica			
Parâmetro	Estimativa	Erro Padrão	Valor do t
Constante	0.8608	0.0966	8.9109
Idade	0.2038	0.0156	13.0641
Região 2	0.0326	0.0981	0.3323
Região 3	0.5403	0.0838	6.4475
2a. talhadia	0.8566	0.1324	6.4698
3o. momento	0.7976	0.1513	5.2716
Região 2 x 2a. talhadia	-0.6373	0.2036	-3.1302
Região 2 x 3o. momento	0.0253	0.3097	0.0817
Região 3 x 2a. talhadia	-0.1464	0.1646	-0.8894
Região 3 x 3o. momento	0.6420	1.0670	0.6017
Função de Ligação Raiz Quadrada			
Parâmetro	Estimativa	Erro Padrão	Valor do t
Constante	1.213	0.1171	10.3587
Idade	0.2798	0.0234	11.9573
Região 2	0.0862	0.1169	6.3913
Região 3	0.7318	0.1145	5.6193
2a. talhadia	1.3610	0.2322	5.8613
3o. momento	1.1540	0.2705	4.2662
Região 2 x 2a. talhadia	-0.8748	0.3304	-2.6477
Região 2 x 3o. momento	0.0966	0.5922	0.1631
Região 3 x 2a. talhadia	0.0148	0.3101	0.0492
Região 3 x 3o. momento	1.4960	2.789	0.5364

A mudança no valor estimado do índice de agregação, $\hat{\phi}$, foi a seguinte: para o modelo sem covariável 2.411 e com covariável 2.184. Ambos foram significativos ($p < 0.05$). O valor do $\hat{\phi}$ para o modelo sem idade não pertence ao intervalo de confiança obtido para esse coeficiente quando é incorporada a covariável no modelo, com $\alpha = 0.05$.

O efeito da interação entre os fatores avaliados foi significativo para a mortalidade das árvores, sendo que a combinação da região do município de Brotas e Ribeirão Bonito com a segunda talhadia foi a responsável por essa dependência.

O menor número de falhas na segunda talhadia deve-se ao fato que as árvores submetidas ao corte na primeira talhadia estavam vivas, com o sistema radicular completo. Após o corte, o sistema radicular continua inalterado fazendo com que a planta suporte o estresse do corte. Na primeira talhadia, principalmente nos primeiros anos de vida, a planta é mais sensível a variações climáticas.

Comparando os dois modelos ajustados, o Extra-Poisson de Breslow (1984) é mais rápido e simples para implementar. Porém, a grande vantagem do modelo Binomial Negativo é a interpretação, no tocante a sua geração. Isto é, existe um modelo probabilístico em dois estágios (Hinde, 1996) para o qual o valor esperado do número de árvores mortas segue uma dis-

tribuição Gama, originando uma distribuição condicional Binomial Negativa para Z_{ijk} , conduzindo à relação média-variância seguinte:

$$Var(Z) = \mu + \mu^2 / \lambda \Rightarrow \lambda = \frac{\mu}{\phi^2 - \mu}$$

fornecendo assim elementos para o conhecimento da dispersão ou agregação possível das respostas, tornando-se importante nos planos de manejo florestal.

A mortalidade é um dos parâmetros chave que determinam mudanças nas populações arbóreas. Segundo Rogers (1974), o conhecimento da dispersão dos indivíduos é fundamental para o desenvolvimento de planos de amostragem confiáveis, e a obtenção de estimativas de parâmetros de fatores de interesse. Conhecendo a distribuição amostral, pode-se estabelecer relações úteis dessas contagens com alguma teoria de crescimento ou de dispersão populacional.

Se se conhece a distribuição amostral das observações, é possível estimar o tamanho amostral requerido para obter uma precisão desejada na estimação do parâmetro de localização (avaliada, por exemplo, como o coeficiente de variação da média amostral). A Tabela 5 apresenta as estimações do n para precisões usuais na pesquisa florestal, seguindo os dois modelos ajustados.

Tabela 5. Tamanhos ótimos de amostras segundo os enfoques Extra-Poisson e Binomial Negativo (Bin. Neg.).
(Sample size under Extra-Poisson and Negative Binomial models).

Precisão (Porcentagem da média)	Tamanho n [Extra-Poisson]	Tamanho n [Bin. Neg.]
1	63	95
1.5	30	42
2	16	24
2.5	10	15
3	9	11
3.5	8	8
4	7	6
4.5	6	5
5	3	3

Pode-se observar que para precisões menores, as estimações dos tamanhos de amostras são semelhantes para ambos os enfoques, resultando mais rigoroso o modelo Binomial Negativo na medida que a exigência da preci-

são se incrementa. Isto fornece elementos de decisão para os futuros planos de manejo florestal, uma vez que o delineamento contempla certos níveis de exigência desejados no que diz respeito às estimativas de posição.

CONCLUSÕES

Após a análise da variável aleatória “mortalidade das árvores” de *Eucalyptus grandis*, no Estado de São Paulo, foram obtidas as conclusões que seguem:

- Existe um processo ecológico de composição no *stand* para o qual o modelo probabilístico Binomial Negativo mostra boa performance.

- Nos municípios de Brotas e Ribeirão Bonito as plantas suportam melhor o estresse do primeiro corte.

- A distribuição de probabilidade da variável mortalidade das árvores de *Eucalyptus grandis* apresenta uma relação média-variância do tipo quadrática.

- modelo Binomial Negativo deve ser o preferido para a obtenção de inferências biparamétricas gerais e o Extra-Poisson pode ser recomendado como alternativa metodológica pela sua simples e rápida implementação.

AUTORES E AGRADECIMENTOS

MARÍA DEL PILAR DÍAZ é Dra. em Estatística e Experimentação Agronômica pela ESALQ/USP, e professora de Estadística y Biometría - Diseño de Experimentos no Departamento de Desarrollo Rural da Facultad de Ciencias Agropecuarias da Universidad Nacional de Córdoba. Ciudad Universitaria. C.C. 509. CP 5000. Córdoba. CBA. Argentina.

HILTON THADEU ZARATE DO COUTO é Professor Titular do Departamento de Ciências Florestais da Escola Superior de Agricultura Luiz

de Queiroz. Universidade de São Paulo. Caixa Postal 9 – 13400-970 – Piracicaba, SP. E-mail: htzcouto@carpa.ciagri.usp.br

Os autores agradecem à Champion S.A. de Papel e Celulose, pelo fornecimento do conjunto de dados de *Eucalyptus grandis* e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela subvenção Reserva Técnica para a apresentação deste trabalho na XVIIIth International Biometric Conference, 1-5 de julho de 1996 em Amsterdã, Holanda.

REFERÊNCIAS BIBLIOGRÁFICAS

ATKINSON, A.C. Regression diagnostics, transformations and constructed variables. **Journal of the Royal Statistics Society**, B, v. 44, n.1, p.1-36, 1982.

BENNETT, S. An extension of Williams' method for overdispersion models. **GLIM newsletter**, v.17, p.12-18, 1989.

BRESLOW, N.E. Extra-poisson variation in log-linear models. **Applied statistics**, v.33, p.38-44, 1984.

BRESLOW, N.E. Test of hypotheses in overdispersed poisson regression and other Quasi-Likelihood models. **Journal of the American Statistical Association**, v. 85, p.565-571, 1990.

COLLETT, D. **Modelling binary data**. London: Chapman and Hall, 1991. 369p.

DEAN, C.; LAWLESS, J.F. Tests for detecting overdispersion in poisson regression models. **Journal of the American Statistical Association**, v.84, p.406-415,1989.

- DÍAZ, M.P. **Ajustes de modelos lineares generalizados para dados de contagem com superdispersão.** Piracicaba: Departamento de Matemática e Estatística ESALQ/USP, 1994. 48p.
- DÍAZ, M.P. **Modelos lineares generalizados com superdispersão para mortalidade de árvores de *Eucalyptus grandis*.** Piracicaba, 1996. 151p. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz. Universidade de São Paulo.
- ELDER, L.F. Statistical analysis of genotoxicity assays. In: SIMEST, Santa Marta, 1995. **Anais.** 52p.
- ELLIOT, J.M. **Some methods for the statistical analysis of samples of benthic invertebrates.** 2.ed. Warehan: Freshwater Biological Association, 1979. 285p.
- FIRTH, D. **Overdispersion.** Piracicaba: Departamento de Matemática e Estatística da ESALQ/USP, 1990. (Minicurso)
- FRANCIS, B.; GREEN, M.; PAYNE, C. **The GLIM system.** Oxford: Clarendon Press, 1993. 821p.
- HINDE, J. Macros for fitting overdispersion models. **Glim newsletter**, v.4, p.12-18, 1996.
- JORGENSEN, B. Exponential dispersion models. **Journal of the Royal Statistical Society**, B, v. 49, p.127-162, 1987.
- LINDSEY, J.K. **Models for repeated measurements.** Oxford: Oxford Science Publications, 1993. 413p.
- MCCULLAGH, P.; NELDER, J.A. **Generalized linear models.** 2.ed. London: Chapman and Hall, 1989. 511p.
- ROGERS, A. **Statistical analysis of spatial dispersion.** London: Pion, 1974. 543p.
- SOMERS, R.G; ODERWALD, W.; LANGDON, O.G. Predicting mortality with a Weibull distribution. **Forest science**, v.26, p.291-300, 1980.
- WEDDERBURN, R.W.M. Quasi-likelihood functions, generalized linear models and the Gauss Newton Method. **Biometrika**, v.63, p.27-32, 1974.

