

Comparação entre Modelos de Relação Hipsométrica com e sem
Componente Espacial para *Pinus* sp. na FLONA Ipanema - SPComparison between Models of Hypsometric Relationship with and
without Spatial Component for *Pinus* sp. at FLONA Ipanema - SPJúlio César Pereira¹, Débora Arantes Campos Mourão²,
Verônica Scalet² e Carlos Alberto Martinelli de Souza (*in memoriam*)**Resumo**

Visando reduzir custos e tempo, muitas vezes nas parcelas de um inventário florestal, apenas algumas árvores têm suas alturas medidas, sendo necessário estimar a altura das demais. Para isso é comum o uso de modelos de regressão relacionando altura e o diâmetro à altura do peito. No entanto, a suposição de independência dos erros em um modelo de regressão nem sempre é razoável, uma vez que para medidas feitas em pontos próximos entre si, espera-se que as alturas tenham valores parecidos (alta correlação espacial). Neste contexto, o presente trabalho teve por objetivo comparar modelos de regressão com e sem componente espacial, para um plantio de *Pinus* sp. da Floresta Nacional de Ipanema – SP e determinar a função de correlação que propicia o melhor desempenho para o modelo com componente espacial. Dos modelos ajustados o modelo que se mostrou com melhor desempenho foi o modelo com componente espacial, tendência quadrática, com as funções de correlação exponencial, gaussiana e esférica.

Palavras-chave: Relação hipsométrica; Geoestatística; Função de correlação; Comparação de modelos

Abstract

In order to reduce costs and time, often in the forest inventory's plots, only a few trees have their height measured, this is necessary to estimate the height of the other trees. So, regression models relating height and diameter at breast height are useful. However, the assumption of the independent errors in the regression model is not always accurate. If measurements are taken at points close to each other, it is expected that the heights have similar values (high spatial correlation). In this context, this paper aims at comparing regression models with and without spatial component, for *Pinus* sp. in the FLONA – SP and to determine the correlation function that gives the best performance for a model with the spatial component. The fitted model that showed the best performance was the model with spatial component, quadratic trend, with exponential, Gaussian and spherical correlation functions.

Keywords: Height-diameter relationship; Geostatistics; Correlation function; Models comparison

INTRODUÇÃO

De acordo com Péllico Netto e Brena (1997), inventário florestal é uma atividade que visa obter informações qualitativas e quantitativas dos recursos florestais existentes em uma área pré-definida.

Um dos principais objetivos de um inventário florestal é obter uma estimativa para o volume total de madeira existente em uma população florestal, sendo a estimativa efetuada a partir de uma amostra retirada de tal população (MELLO, 2004; SOARES *et al.*, 2007). Com essa

finalidade é comum o uso de equações volumétricas, correlacionando o volume com diâmetro à altura do peito (dap) e a altura total (SCHUMACHER; HALL, 1933). No entanto, visando reduzir custos e tempo para a realização de um inventário, muitas vezes apenas algumas alturas são medidas nas parcelas do inventário, sendo necessário estimar a altura das demais árvores (LEITE; ANDRADE, 2003).

Em inventários florestais, a variável diâmetro à altura do peito é facilmente medida para todas as árvores; a altura, no entanto obtida de modo

¹Professor Doutor da Universidade Federal de São Carlos, Campus Sorocaba – Rodovia João Leme dos Santos, km 110, SP 264, Sorocaba – SP - 18052-780 – Email: julio-pereira@ufscar.br

²Estudante de Engenharia Florestal na Universidade Federal de São Carlos, Campus Sorocaba – Rodovia João Leme dos Santos, km 110, SP 264, Sorocaba – SP - 18052-780 – Email: debyarantes@yahoo.com.br; ve.scalet@gmail.com

indireto através de instrumentos apropriados, fornece resultados acurados, porém não econômicos, devido ao tempo gasto na operação de coleta de dados. Por isso, é comum a medição do diâmetro de todas as árvores das parcelas e a altura apenas em algumas delas. O conjunto de dados de alturas das árvores medidas, com os respectivos diâmetros, é usado para estabelecer uma relação de regressão da altura em função do diâmetro (modelo de relação hipsométrica), a qual é empregada para estimar as alturas das demais árvores da parcela em função dos diâmetros já medidos. Logo, estima-se a variável de difícil acesso em função da mais fácil e de rápida medição, reduzindo-se, portanto, o tempo e os custos de mensuração (BARTOSZECK *et al.*, 2004).

Quando disponível, outras variáveis, além do dap, que influenciam a variável altura, podem ser incluídas no modelo hipsométrico, como observado no estudo de Guimarães *et al.* (2009), num plantio de *Eucalyptus* sp., em que os modelos foram ajustados levando-se em consideração a idade e o material genético. Os autores observaram que ao se acrescentar o material genético ou a idade ao modelo, o erro residual diminuiu em relação ao modelo sem a presença dessas variáveis. Além desse exemplo, recentemente diversos modelos vêm sendo testados, bem como diversas estratégias e metodologias vem sendo estudadas a fim de buscar maior precisão no ajuste dos modelos e na estimação da variável altura (RIBEIRO *et al.*, 2010; MISIR, 2010).

Segundo Mello (2004), os métodos tradicionais de estatística usados para a análise dos inventários florestais, utilizam-se de uma medida central (média) e uma dispersão (variância) para descrever um determinado fenômeno, dessa maneira, para os modelos tradicionalmente utilizados, assumem-se que os erros são independentes. Esta suposição nem sempre é válida em um modelo de regressão para altura em função do dap, principalmente quando se têm observações de árvores vizinhas. Nesse caso, espera-se que para medidas feitas em pontos próximos entre si as alturas tenham valores parecidos (alta correlação espacial) enquanto que, para pontos mais distantes entre si, as observações sejam mais independentes (SCHMIDT; SANSÓ, 2006).

Uma forma de explorar as relações que possam existir entre unidades amostrais é fazendo o uso da geoestatística (CRESSIE, 1993; DIGGLE; RIBEIRO JÚNIOR, 2007). A geoestatística, em-

pregada na modelagem da dependência espacial, em geral faz uso dos processos gaussianos. Considerando uma variável de interesse, Y , geralmente assume-se que o conjunto de observações $\{y(s): s \in G\}$ é uma realização de um processo gaussiano, sendo G um subconjunto fixo de \mathfrak{R}^p , em geral $p = 2$ ou 3 . Diz-se que uma quantidade aleatória, $Y(\cdot)$, que assume valores $y(s)$, $s \in G$ segue um processo gaussiano com média $\mu(\cdot)$ e função de covariância $c(\cdot, \cdot)$ denotado por $Y(\cdot) \sim PG(\mu(\cdot), c(\cdot, \cdot))$, se para quaisquer $s_1, \dots, s_n \in G$ e qualquer $n=1, 2, \dots$, a distribuição conjunta de $Y(s_1), \dots, Y(s_n)$ é uma normal multivariada com parâmetros dados por $E[Y(s_i)] = \mu(s_i)$ e $cov\{Y(s_i), Y(s_j)\} = c(s_i, s_j)$ (O'HANGAN, 1994).

O uso dos métodos geoestatísticos nas ciências florestais e em particular em estudos envolvendo características dendrométricas vem aumentando significativamente nos últimos anos (NANOS *et al.*, 2004; MELLO *et al.*, 2005; BOGNOLA *et al.*, 2008). Um exemplo da inclusão da dependência espacial em modelos de relação hipsométrica, pode ser encontrado em Meng *et al.* (2009). Estes autores usaram três tipos de modelos de regressão espacial para modelar a relação dap x altura. Primeiramente selecionaram dentre sete, um modelo linear clássico para a altura em função do dap. A seguir verificaram a correlação espacial no resíduo e a partir do modelo selecionado ajustaram três modelos de regressão espacial, comparando-os entre si e com o modelo linear clássico. Concluíram que os modelos de regressão espacial apresentaram os melhores resultados.

Ao se considerar a estrutura de dependência espacial em um modelo de relação hipsométrica, pode-se obter um modelo mais adequado para a descrição da variável altura. Ou seja, um modelo considerando o dap e a estrutura de correlação espacial, pode descrever melhor a altura do que um modelo que considera apenas o dap. Consequentemente, empregando-se um modelo que descreve melhor os dados de altura, pode-se obter estimativas mais realistas dessa variável em localizações não medidas.

O presente trabalho teve por objetivo comparar modelos de relação hipsométrica com e sem componente espacial, para dados de um plantio de *Pinus* sp. da Floresta Nacional de Ipanema e determinar, para os dados em estudo a função de correlação que propicia o melhor desempenho para o modelo com componente espacial.

MATERIAL E MÉTODOS

Caracterização da área de estudo

Os dados utilizados no presente estudo foram obtidos de um talhão de *Pinus* sp. pertencente à Floresta Nacional de Ipanema, localizada na Região Sudoeste do Estado de São Paulo, a 120 km da capital paulista, entre as coordenadas 23° 25' e 23° 27' latitude sul, 47° 35' e 47° 40' longitude oeste, com altitudes variando de 550 m a 968 m, fazendo divisa entre os Municípios de Iperó, Araçoiaba da Serra e Capela do Alto (REGALADO, 1999). O clima da região pertence ao tipo "Cfa", ou seja, subtropical, mesotérmico úmido, sem estiagens, com estações chuvosas e secas bem definidas (ALBUQUERQUE; RODRIGUES, 2000). Precipitação média regional de

1.400 mm, com mínimo de 800 mm e máximo de 2.200 mm (SOUZA; MARTOS, 2008).

Um total de 1133 árvores foram observadas para o talhão, cujos dados dendrométricos coletados foram: i) O dap (diâmetro à altura do peito - 1,30m), mensurado com fita diamétrica; ii) A altura - medida com o auxílio do hipsômetro de Haga, ajustado para a escala de distância com 15m. Para o posicionamento de cada árvore, estipulou-se um ponto próximo ao talhão definido como origem do sistema de coordenadas cartesianas. A partir da origem foram obtidos os pares ordenados $s=(x,y)$ indicando a localização das árvores em relação à origem estipulada (Figura 1). Desta forma, foi obtido o posicionamento de cada árvore dentro do talhão, permitindo uma análise espacial dos dados.

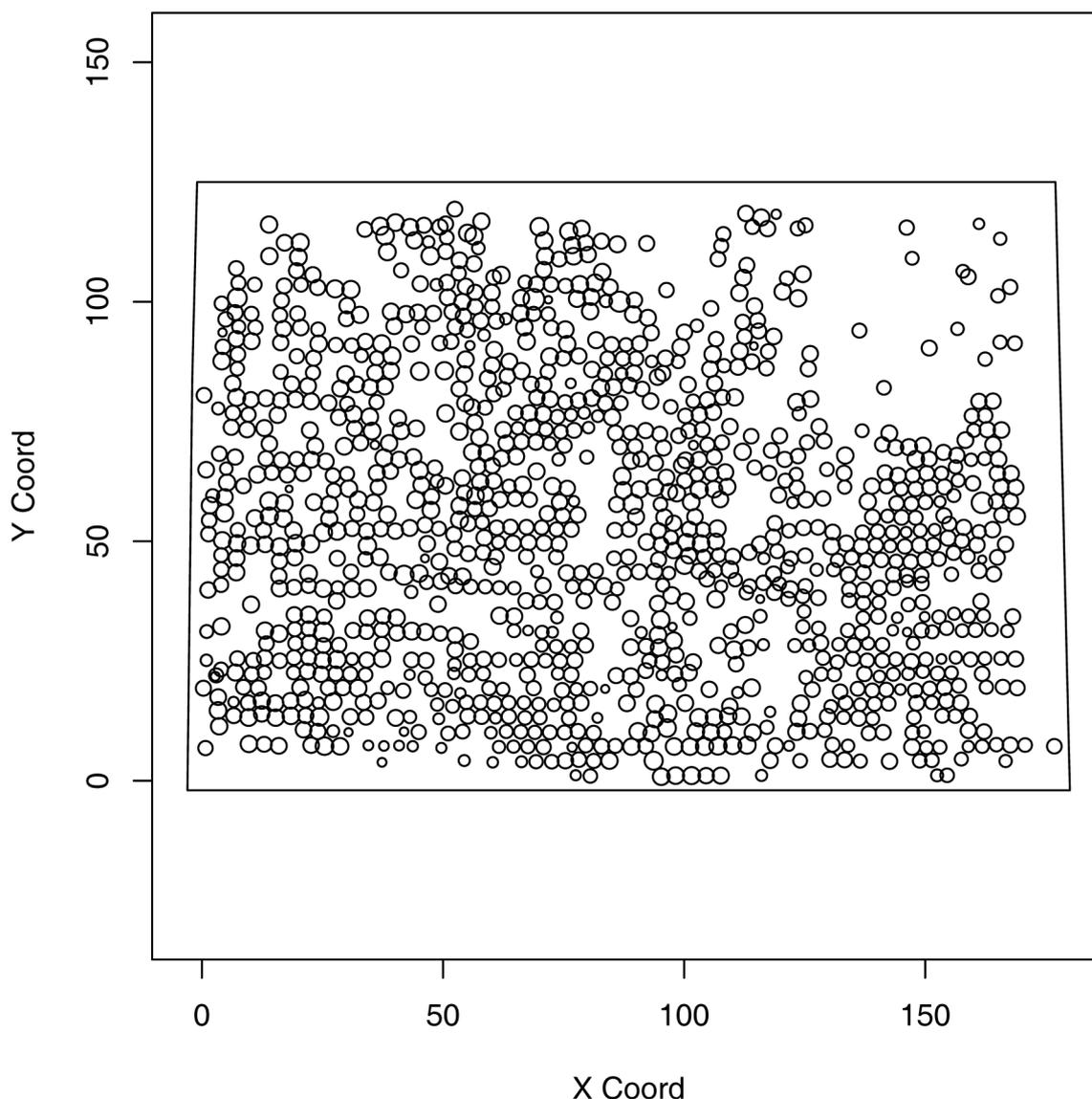


Figura 1. Posicionamento das árvores do talhão de *Pinus* sp. pertencente à Floresta Nacional de Ipanema. Cada círculo representa uma árvore, cuja altura é proporcional ao diâmetro do círculo.

Figure 1. Position of the trees in a plot of *Pinus* sp. belonging to the National Forest of Ipanema. Each circle represents a tree, whose height it is proportional to the diameter.

Métodos

Modelos de relação hipsométrica, com e sem componente espacial, os quais são descritos a seguir, foram ajustados aos dados citados na seção anterior.

Modelos com componente espacial

Considerando-se uma amostra da variável altura observada ao longo da área de estudo, $y(s_1)$, $y(s_2)$, ..., $y(s_n)$, e o dap como covariável, pretende-se encontrar um modelo candidato aos dados. Assumindo-se que a amostra observada forma uma realização parcial de um processo gaussiano, um modelo candidato é dado por:

$$Y(s) = X(s)\beta + v(s) + \varepsilon(s) \quad (1)$$

em que $X(s)\beta$ é a componente de tendência, $X(s) = [1 \text{ dap}(s)]$, $\beta = (\beta_0, \beta_1)'$, $\text{dap}(s)$ representa o diâmetro à altura do peito na localização s ; $v(\cdot)$ é um efeito aleatório com estrutura de correlação espacial tal que $(v(s)|\sigma, \phi)$ segue um processo gaussiano, $(v(s)|\sigma, \phi) \sim PG(0, \sigma^2 \rho(\|s-s'\|; \phi))$, sendo σ^2 a variância do processo e $\rho(\cdot; \phi)$ representa a função de correlação que depende do parâmetro ϕ ; $\varepsilon(s)$ é uma componente de ruído branco, normalmente distribuída, com média 0 e variância τ^2 (efeito pepita), que descreve a variação de microescala.

No modelo (1), a tendência do processo foi expressa por uma função linear do dap. Um segundo modelo candidato é especificado assumindo-se uma tendência quadrática em relação ao dap. O modelo é dado por:

$$Y(s) = X(s)\beta + v(s) + \varepsilon(s) \quad (2)$$

em que, $X(s) = [1 \text{ dap}(s) \text{ dap}^2(s)]$, $\beta = (\beta_0, \beta_1, \beta_2)'$, $\text{dap}(s)$ representa o diâmetro à altura do peito na localização s ; dap^2 denota o quadrado do dap e as componentes $v(s)$ e $\varepsilon(s)$ são as mesmas do modelo (1).

Segundo Diggle e Ribeiro Júnior (2007), a média (componente de tendência) e a estrutura de covariância, juntas definem um modelo para os dados. Para os modelos das Eq. (1) e Eq. (2)

diversas funções de correlação podem ser utilizadas. Foram testadas as funções apresentadas na Tabela 1, conhecidas na literatura geoestatística, a fim de eleger, dentre elas, aquela que melhor descreve a estrutura de dependência espacial das alturas das árvores vizinhas.

Modelos sem componente espacial

Modelos de regressão são amplamente utilizados para descrever a relação dap x altura. Em geral, os modelos de regressão utilizam-se de uma medida central (média) e uma dispersão (variância) para descrever um determinado fenômeno, dessa maneira, para os modelos tradicionalmente utilizados, assumem-se que os erros são independentes. Esses modelos são particularmente úteis e podem apresentar bom desempenho na ausência de correlação espacial nos dados. Neste sentido um terceiro modelo candidato, considerado neste trabalho é dado pela Eq. (3).

$$Y_i = X_i\beta + \varepsilon_i \quad (3)$$

em que, $X_i\beta$ é a componente de tendência, $X = [1 \text{ dap}_i]$, $\beta = (\beta_0, \beta_1)'$, dap_i representa o diâmetro à altura do peito para a i -ésima árvore observada e $\varepsilon_i \sim N(0, \sigma_y^2)$. O índice y no termo σ_y^2 é usado para evitar confusão com a variância do processo espacial, denotada por σ^2 . Este modelo, também foi ajustado aos dados, por ser um modelo mais simples que o modelo da Eq. (1), uma vez que não possui a componente espacial, consequentemente isso reduz o número de parâmetros a serem estimados.

Um modelo sem componente espacial, correspondente ao modelo da Eq. (2), isto é, com tendência quadrática em relação ao dap, também foi ajustado. O modelo é dado por:

$$Y_i = X_i\beta + \varepsilon_i \quad (4)$$

em que, $X = [1 \text{ dap}_i \text{ dap}_i^2]$, $\beta = (\beta_0, \beta_1, \beta_2)'$, dap_i representa o diâmetro à altura do peito para a i -ésima árvore observada, dap_i^2 denota o quadrado do dap e a componente ε_i é a mesma do modelo da Eq. (3).

Tabela 1. Funções de correlação utilizadas para os modelos das Eq. (1) e Eq. (2).

Table 1. Correlation function used in models of Eq. (1) and Eq. (2).

Função de correlação	$p(d)$
Exponencial	$p(d) = \exp\{-d/\phi\}$, $\phi > 0$
Gaussiana	$p(d) = \exp\{-d/\phi\}^2$, $\phi > 0$
Esférica	$p(d) = \exp \begin{cases} 1 - \frac{3}{2}(d/\phi) + \frac{1}{2}(d/\phi)^3, & 0 < d < \phi, \\ 0, & d > \phi. \end{cases}$
Cúbica	$p(d) = \exp \begin{cases} 7(d/\phi)^2 - 8,75(d/\phi)^3 - 3,5(d/\phi)^5 - 0,75(d/\phi)^7, & 0 < d < \phi, \\ 0, & d > \phi. \end{cases}$
Matérn, com $\kappa=0,3$ e $\kappa=0,2$	$p(d) = \exp \{2^{\kappa-1} \Gamma(\kappa)\}^{-1} (d/\phi)^{\kappa} K_{\kappa}(d/\phi)$, em que $K_{\kappa}(\cdot)$ denota uma função de Bessel modificada de ordem κ , $\phi > 0$ e $\kappa > 0$.

Método de ajuste dos modelos

Considerando um modelo candidato aos dados, o próximo passo é a estimação dos parâmetros do modelo. A estimação pelo método da máxima verossimilhança é frequentemente utilizada, por suas propriedades ótimas bem conhecidas para grandes amostras.

Nos modelos espaciais (Eq. (1) e Eq. (2)), os parâmetros a serem estimados são os de tendência, representados no vetor β , os parâmetros σ^2 e ϕ implícitos na componente espacial $v(s)$ e o efeito pepita τ^2 . Considerando-se que o vetor $Y = (Y(s_1), \dots, Y(s_n))$, forma uma amostra aleatória parcial do processo gaussiano $\{Y(s): s \in G\}$, em que $Y(s)$ representa a variável altura na localização s , a distribuição conjunta de $Y(s_1), \dots, Y(s_n)$ é normal n -variada, isto é, $Y \sim N_n(X\beta, \sigma^2 R + \tau^2 I_n)$, em que $R_{ii} = \rho(s_i - s_i)$, I_n é a matriz identidade de ordem n e a i -ésima linha do vetor $X\beta$ é dada por $\beta_0 + \beta_1 \text{dap}(s_i)$ para o modelo da Eq. (1) e $\beta_0 + \beta_1 \text{dap}(s_i) + \beta_2 \text{dap}^2(s_i)$ para o modelo da Eq. (2). A distribuição conjunta considerada como função dos parâmetros, dadas as observações, é a verossimilhança associada aos modelos das Eq. (1) ou Eq. (2), que pode ser expressa por:

$$L(\beta, \sigma^2, \phi, \tau^2 | y) = (2\pi)^{-n/2} |\sigma^2 R + \tau^2 I|^{-1/2} \exp\{-1/2 [y - X\beta]' [\sigma^2 R + \tau^2 I]^{-1} [y - X\beta]\} \quad (5)$$

A maximização da Eq. (5) produz estimativas de máxima verossimilhança dos parâmetros, e é obtida por procedimentos numéricos iterativos, para isso foi utilizado pacote geoR (RIBEIRO JÚNIOR; DIGGLE, 2001) do software estatístico R (R DEVELOPMENT CORE TEAM, 2009).

Os modelos sem componente espacial, apresentados nas Eq. (3) e Eq. (4), podem ser vistos como casos particulares dos modelos das Eq. (1) e Eq. (2) respectivamente. Nas Eq. (3) e Eq. (4) considera-se que as observações, componentes do vetor $Y = (Y(s_1), \dots, Y(s_n))$, são espacialmente independentes, podendo o vetor ser escrito como $Y = (Y_p, \dots, Y_n)$. Dessa forma, a distribuição conjunta de Y_p, \dots, Y_n é normal n -variada, com vetor de médias $X\beta$ e matriz de variâncias dada por $\sigma^2 I_n$. A função de verossimilhança associada é dada por:

$$L(\beta, \sigma^2 | y) = (2\pi)^{-n/2} |\sigma^2 I|^{-1/2} \exp\{-1/2 [y - X\beta]' [\sigma^2 I]^{-1} [y - X\beta]\} \quad (6)$$

A maximização da Eq. (4) produz estimativas de máxima verossimilhança dos parâmetros β e σ_y^2 .

Uma variante do método da máxima verossimilhança é o método da máxima verossimilhança restrita. No caso do modelo com componente

espacial, o método consiste em se trabalhar com transformações lineares dos dados originais, da forma $Y^* = PY$, tal que a distribuição de Y^* não depende de β (DIGGLE; RIBEIRO JÚNIOR, 2007; OLIVEIRA *et al.*, 2006), em que P é um projetor ortogonal, por exemplo $P = I - X(X'X)^{-1}X'$, de forma que um estimador para $\theta = (\tau^2, \sigma^2, \phi)$ pode ser obtido maximizando a seguinte expressão:

$$L^*(\theta) = (n-p/2) \ln(2\pi) + 1/2 \ln(X'X) - 1/2 \{ \ln |\tau^2 I + \sigma^2 R| + \ln |X'[\tau^2 I + \sigma^2 R]^{-1} X| + (y - X\tilde{\beta})' [\tau^2 I + \sigma^2 R]^{-1} (y - X\tilde{\beta}) \}$$

em que, p é o número de parâmetros do vetor β , $\tilde{\beta} = \hat{\beta}(\theta)$ e $R = R(\phi)$.

A predição ou interpolação de $Y(\cdot)$ em uma localização não observada (s_u) da região em estudo é chamada na literatura geoestatística de krigagem. Para a realização das predições, neste trabalho foram utilizadas as funções *krige.conv* e *krige.control* do pacote geoR (RIBEIRO JUNIOR; DIGGLE, 2001), onde a tendência do modelo é especificada usando os argumentos *trend.d*, para os valores de tendência nas localizações dos dados observados e *trend.l* para especificar os valores de tendência nas localizações deixadas para predição.

Crerios de comparação dos modelos

Considerando-se o modelo com componente espacial da Eq. (1), associado a cada uma das seis funções de correlação descritas na Tabela 1, além do modelo da Eq. (3), foram ajustados um total de sete modelos com tendência linear em relação ao dap. Do mesmo modo, considerando-se o modelo com componente espacial da Eq. (2), associado às funções de correlação da Tabela 1, além do modelo da Eq. (4), foram ajustados um total de sete modelos com tendência quadrática em relação ao dap. A seleção dos modelos foi realizada utilizando-se o critério de informação de Akaike (1974) e técnicas de validação cruzada.

O critério de informação de Akaike (AIC) é obtido de $AIC = -2\hat{l} + 2p$, em que p representa o número de parâmetros e \hat{l} o logaritmo do valor maximizado da função de verossimilhança. Um valor baixo para AIC é considerado como representativo de um melhor ajuste e os modelos são selecionados visando a obter um mínimo AIC. Na comparação de modelos pelo critério AIC, segundo Hardin e Hilbe (2007), dois modelos que apresentam uma diferença entre AICs maior que 2 são consideravelmente distintos.

Para a realização da validação cruzada, 150 localizações, s_u , foram deixadas de fora do processo de inferência, isto é, não foram utilizadas no ajuste dos modelos, para que depois pudessem ser estimadas pelos modelos ajustados. Dessa forma, foram realizadas previsões em tais localizações e comparações foram feitas entre as previsões de altura e os valores observados nessas localizações. O erro quadrático médio (EQM) e o erro absoluto médio foram utilizados na comparação dos valores preditos versus observados, para cada modelo ajustado, cujas expressões são dadas por:

$$EQM = 1/150 \sum_{u=1}^{150} (Y(s_u) - \hat{Y}(s_u))^2,$$

$$EAM = 1/150 \sum_{u=1}^{150} |Y(s_u) - \hat{Y}(s_u)|,$$

em que $Y(s_u)$ representa o valor observado e $\hat{Y}(s_u)$ o valor predito na localização s_u . Quanto menor forem os valores de EQM e EAM , melhor é a predição do modelo.

RESULTADOS E DISCUSSÃO

Uma análise descritiva dos dados mostrou que o coeficiente de correlação estimado entre dap e altura foi de 0,65, evidenciando, como era de se esperar, a dependência da altura em relação ao dap.

Na Figura 2 são apresentados os gráficos referentes a uma análise descritiva dos dados de altura retirando-se o efeito de tendência linear em relação ao dap. A Figura 2(a) e a Figura 2(b) permitem verificar que aparentemente não há evidências de tendência dos dados com relação aos eixos x e y do sistema de coordenadas. Pelo histograma da Figura 2(c) nota-se que a distribuição do resíduo é aproximadamente simétrica e aparentemente se distribui de acordo com uma distribuição normal. Comportamentos similares apresentaram os resíduos retirando-se o efeito de tendência quadrática em relação ao dap.

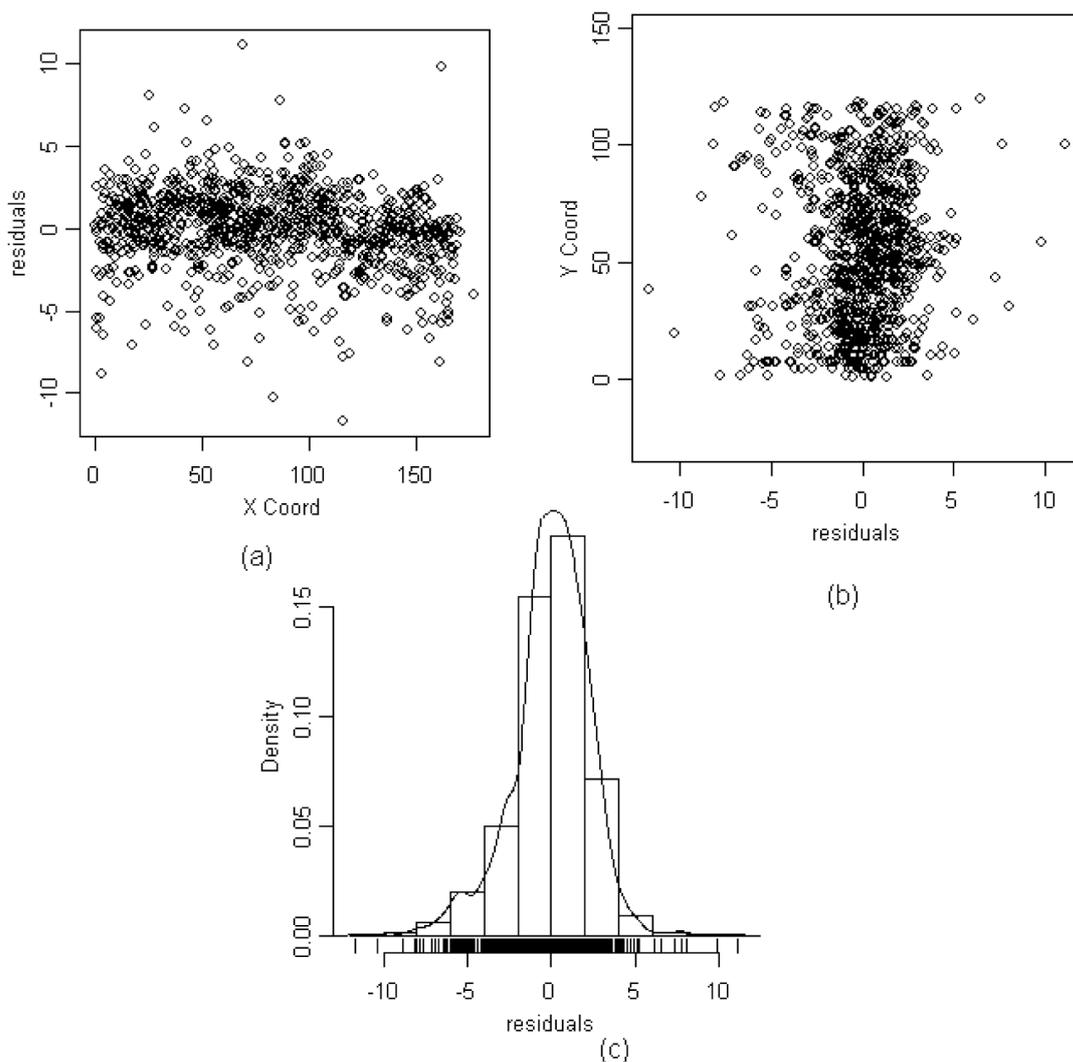


Figura 2. Gráficos exploratórios dos dados de altura retirando-se o efeito de tendência linear em relação ao dap, (a) relação dos dados com o eixo-x, (b) relação dos dados com o eixo-y e (c) histograma dos dados.

Figure 2. Exploratory graphs of height data without the linear trend effect due to the diameter at breast height, (a) data to x-axis relationship, (b) data to y-axis relationship and (c) histogram of the data.

A fim de verificar a existência de correlação espacial no resíduo após retirar as tendências linear e quadrática da altura em relação ao dap, construíram-se os semivariogramas juntamente com os gráficos de envelope, apresentados nas Figura 3(a) e (b) respectivamente. Nota-se a existência de correlação espacial, em ambos os casos, pois conforme há um aumento na distância, a semivariância aumenta até se aproximar de um valor constante. Além disso, a região do envelope representa a região sem dependência espacial e como alguns pontos dos semivariogramas estão fora do envelope, fica mais evidente a existência de correlação espacial.

Na Tabela 2 são apresentados os valores de AIC para o ajuste dos modelos em comparação, usando os métodos da máxima verossimilhança e máxima verossimilhança restrita. Nota-se que os modelos sem componente espacial foram os que apresentaram os maiores valores de AIC, indicando que um modelo com a componente espacial, com qualquer uma das funções de correlação utilizadas, é melhor que o correspondente modelo sem a componente espacial, segundo esse critério. Resultado semelhante foi encontrado por Meng *et al.* (2009), em que os modelos de regressão espacial utilizados apresentaram melhor performance que os modelos clássicos.

Comparando-se os modelos, ajustados com tendências linear e quadrática, pode-se constatar na Tabela 2 que o modelo que se utiliza do

dap juntamente com o dap^2 , isto é o modelo com tendência quadrática, descreve melhor os dados da população estudada, apresentado os menores valores de AIC. O uso de ambas as variáveis é às vezes útil na modelagem da variável altura. No trabalho de Meng *et al.* (2009), um modelo incluindo essas variáveis, também apresentou um bom ajuste.

A fim de se obter uma confirmação da necessidade da inclusão da variável dap^2 dado que a variável dap está no modelo, ou da necessidade da variável dap dado que dap^2 está no modelo, foi aplicado o algoritmo stepwise de seleção de variáveis. Para tanto utilizou-se o comando *step()* do software estatístico R (R DEVELOPMENT CORE TEAM, 2009), tomando-se como modelo de partida o modelo que inclui dap e dap^2 . O resultado evidenciou ser significativa a contribuição do dap^2 no modelo dada a presença do dap, apresentando uma redução de 152 no valor de AIC com a inclusão do dap^2 . Ficou evidenciada, também, a contribuição significativa do dap no modelo dada a presença do dap^2 , apresentando uma redução de 267 no valor de AIC com a inclusão do dap.

O método da máxima verossimilhança restrita se mostrou melhor para ajuste dos modelos que o método da máxima verossimilhança, isto é, apresentou os menores valores de AIC, para qualquer função de correlação, tanto com tendência quadrática e tendência linear em relação ao dap (Tabela 2).

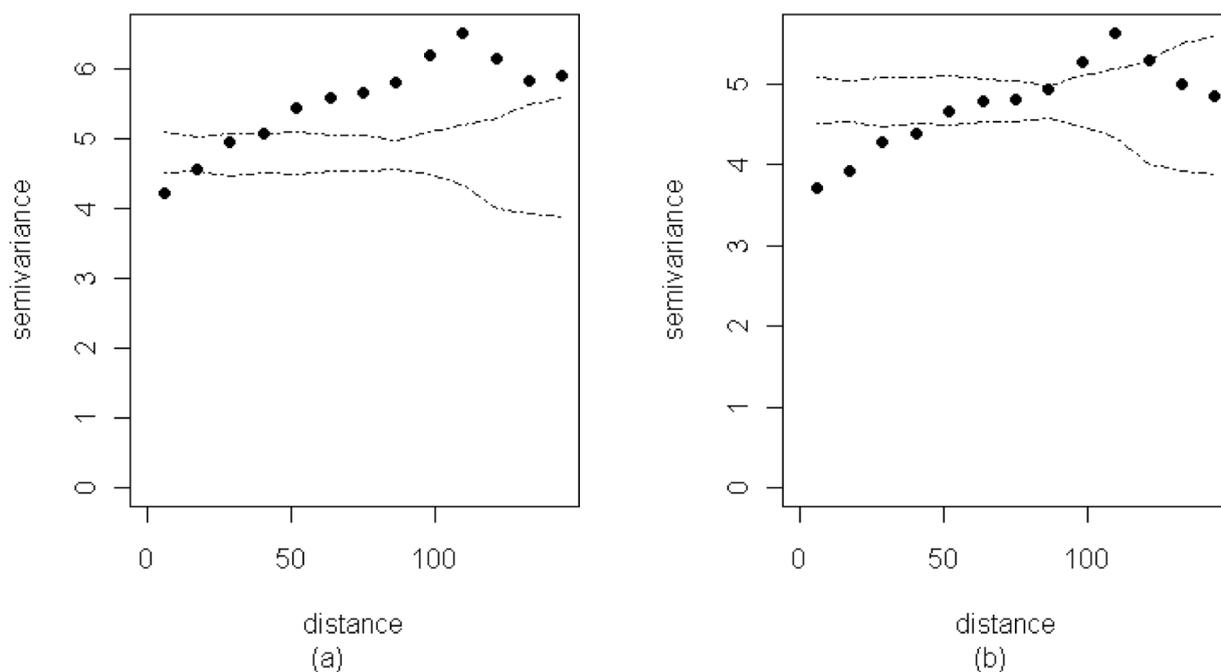


Figura 3. Semivariograma dos dados de altura após retirar a tendências linear (a) e quadrática (b) em relação ao dap, juntamente com o gráfico de envelope.

Figure 3. Semi-variogram of height data without the linear (a) and quadratic (b) trend effect due to the diameter at breast height, together with envelope plot.

Dentre os modelos com componente espacial, o modelo com função de correlação esférica foi o que apresentou os melhores resultados, considerando o ajuste por MV bem como por MVR, segundo o critério AIC. No entanto, o modelo com componente espacial, usando as funções de correlação gaussiana e exponencial, apresentaram valores de AIC muito próximos ao do modelo com função de correlação esférica, cujas diferenças entre os AICs desses três modelos são menores ou iguais a 2, para o método da MVR, como se observa na Tabela 2, ou sejam, diferenças não significativas, segundo Hardin e Hilbe (2007).

Na Tabela 3 são apresentados os valores do EQM e do EAM para as predições realizadas nas 150 localizações que não foram utilizadas no ajuste dos modelos. Segundo esses critérios, os modelos com componente espacial foram os que apresentaram os melhores resultados, isto é, apresentaram os menores valores de EQM e EAM. Esses resultados evidenciam novamente que os modelos com componente espacial produzem melhores predições independentemente da função de correlação espacial utilizada.

Comparando-se os modelos ajustados com tendência linear e quadrática em relação ao dap, segundo o EQM e EAM, observa-se na Tabela 3 que os valores obtidos para os modelos com tendência quadrática foram inferiores aos obtidos para os modelos com tendência linear. Esses resultados, mais uma vez evidenciam que ao se utilizar uma tendência quadrática o ajuste é feito com maior precisão, conseqüentemente obtêm-se melhores predições da variável altura. A tendência quadrática encontrada nos dados pode ser explicada pelo fato do povoamento analisa-

do estar em sua fase madura, onde o crescimento em dap e altura já não têm mais a mesma proporção, e por isso essa relação assemelha-se a uma parábola, diferentemente do que é observado para florestas em pleno crescimento, onde essa a tendência pode ser representada por uma reta.

Considerando a tendência quadrática e o ajuste pelo método da MVR, o modelo com componente espacial e função de correlação esférica, foi o que apresentou os menores valores de EQM e EAM (Tabela 3). Esse resultado está de acordo com o que foi observado em relação ao critério AIC, onde o modelo com função de correlação esférica foi um dos que apresentaram os melhores resultados. Dessa forma, dentre as funções de correlação analisadas, as funções esférica, gaussiana e exponencial são as que melhor se ajustam aos dados; e a função esférica é que proporciona melhor capacidade de predição.

Como já notado em outros trabalhos (BOGNOLA *et al.*, 2008; MELLO *et al.*, 2005; NANOS *et al.*, 2004) a presença de correlação espacial é uma característica comum para variáveis dendrométricas e em particular para a variável altura (MENG *et al.*, 2009). Essa característica também foi constatada neste trabalho, para a variável altura. Dessa forma, a presença da dependência espacial explica o fato dos modelos com componente espacial apresentarem melhor desempenho em relação ao modelo sem a componente espacial.

Ficou evidente, pelos critérios adotados neste trabalho, a superioridade de um modelo com componente espacial e tendência quadrática para a variável altura. Das funções de correlação espacial analisadas, observou-se que não há

Tabela 2. Valores de AIC obtidos para os modelos ajustados pelos métodos de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR), considerando-se tendência linear e quadrática em relação ao dap.

Table 2. AIC values achieved for filtered models with maximum likelihood and restricted maximum likelihood methods, considering linear and quadratic trend.

Modelo	Método de Ajuste	AIC (Tendência Linear)	AIC (Tendência Quadrática)
Exponencial	MV	4321	4181
	MVR	4308	4165
Gaussiano	MV	4320	4177
	MVR	4309	4164
Esférico	MV	4319	4177
	MVR	4308	4163
Cúbico	MV	4323	4179
	MVR	4312	4166
Matérn ($\kappa=0,3$)	MV	4324	4184
	MVR	4311	4169
Matérn ($\kappa=0,2$)	MV	4327	4187
	MVR	4314	4171
Sem componente espacial	MV	4486	4334
	MVR	4479	4343

Tabela 3. Erro absoluto médio e Erro quadrático médio para as predições de altura usando cada modelo ajustado, segundo os métodos da máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR), considerando-se tendência linear e quadrática para o dap.

Table 3. Mean absolute error and Mean square error for the predictions of height, using each fitted model, with maximum likelihood and restricted maximum likelihood methods, considering linear and quadratic trend.

Modelo	Método de Ajuste	Tendência Linear		Tendência Quadrática	
		EAM	EQM	EAM	EQM
Exponencial	MV	1,4596	4,2077	1,2695	3,3275
	MVR	1,4596	4,2090	1,2701	3,3348
Gaussiano	MV	1,4398	4,1967	1,2618	3,3297
	MVR	1,4399	4,2014	1,2623	3,3342
Esférico	MV	1,4561	4,1852	1,2604	3,3020
	MVR	1,4567	4,1866	1,2606	3,3042
Cúbico	MV	1,4455	4,2113	1,2648	3,3418
	MVR	1,4560	4,2446	1,2658	3,3469
Matérn ($\kappa=0,3$)	MV	1,4618	4,2134	1,2704	3,3316
	MVR	1,4622	4,2155	1,2715	3,3330
Matérn ($\kappa=0,2$)	MV	1,4629	4,2174	1,2703	3,3348
	MVR	1,4535	4,2201	1,2715	3,3366
Sem componente espacial	MV	1,6093	4,7089	1,4677	3,7629
	MVR	1,6093	4,7089	1,4677	3,7629

grandes diferenças entre as funções exponencial, gaussiana e esférica. No entanto, a função esférica pode ser preferível, pois apresentou melhor capacidade preditiva, além de ser amplamente empregada na prática. Esses resultados se revestem de importância, pois adotando-se um modelo mais adequado e com melhor capacidade preditiva, ganha-se em precisão na estimação da variável altura, a qual é fundamental para a estimativa do volume. Sendo assim, valores de altura com erros menores ocasionam em maior acurácia nas estimativas volumétricas.

CONCLUSÕES

Conforme os resultados e discussão apresentados, pode-se concluir que a adição de uma componente espacial no modelo de relação hipsométrica trouxe um ganho no ajuste do modelo, para os dados do plantio de *Pinus* sp. da Floresta Nacional de Ipanema, independentemente da função de correlação ajustada. A tendência quadrática em relação ao dap, explica melhor a variável altura do que a tendência linear. E dentre as funções de correlação analisadas as funções exponencial, gaussiana e esférica, apresentaram os melhores resultados, podendo ser indicada a função esférica, por ter apresentado melhor capacidade preditiva.

AGRADECIMENTOS

Aos professores Paulo Justiniano Ribeiro Junior (UFPR) e Clarice Garcia Borges Demétrio (ESALQ/USP), pelas discussões. À dire-

toria da FLONA de Ipanema, por permitir a coleta dos dados. Ao Cícero e aos alunos do curso de Engenharia Florestal da UFSCAR – Sorocaba: Estevam, Emily, Jéssica, Carlos Juliano e Danieli, pela colaboração na coleta e tabulação dos dados. Aos três revisores anônimos pelas valiosas sugestões.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBUQUERQUE, G.B; RODRIGUES, R.R. A vegetação do Morro de Araçoiaba, Floresta Nacional de Ipanema, Iperó (SP). *Scientia Forestalis*, Piracicaba, v.58, n.2, p.145-159, 2000.
- AKAIKE, H.A new look at statistical model identification. *IEEE Transactions on Automatic Control*, New York, v.19, n.6, p.716-723, 1974.
- BARTOSZECK, A.P.S.; MACHADO, S.A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E.B.A distribuição diamétrica para bracingais em diferentes idades, sítios e densidades na região metropolitana de Curitiba. *Floresta*, Curitiba, v.34, n.3, p.305-323, 2004.
- BOGNOLA, I.; RIBEIRO JÚNIOR, P.; SILVA, E.; LINGAU, C.; HIGA, A. Modelagem uni e bivariada da variabilidade espacial de rendimento de *Pinus taeda* L. *Floresta*, Curitiba, v.38, n.2, p. 373-385, 2008.
- CRESSIE, N.A.C. *Statistics for spatial data revised edition*. New York: John Wiley & Sons, 1993. 928 p.
- DIGGLE, P.J.; RIBEIRO JÚNIOR, P.J. *Model-based geostatistics*. Londres: Springer, 2007. 228 p.

- GUIMARÃES, M.A.M.; CALEGÁRIO, N.; CARVALHO, L.M.T.; TRUGILHO, P.F. Height-diameter models in forestry with inclusion of covariates. *Cerne*, Lavras, v.15, n.3, p.313-321, 2009.
- HARDIN, J.W.; HILBE, J.M. **Generalized linear models and extensions**. 2ed. Bryan: Stata Press, 2007. 387p.
- LEITE, H.G.; ANDRADE, V.C.L. Importância das variáveis altura dominante e altura total em equações hipsométricas e volumétricas. *Revista Árvore*, Viçosa, v.3, n.27, p.301-310, 2003.
- MELLO, J.M. **Geoestatística Aplicada ao Inventário Florestal**. 2004. 111p. Tese (Doutorado em Recursos Florestais) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2004.
- MELLO, J.M.; BATISTIA, J.L.F.; RIBEIRO JÚNIOR, P.J.; OLIVEIRA, M.S. Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de *Eucalyptus grandis*. *Scientia Forestalis*, Piracicaba, n.69, p.25-37, 2005.
- MENG, Q.; CIESZEWSKI, J.C.; STRUB, M.R.; BORDERS, B.E. Spatial regression modeling of tree height-diameter relationships. *Canadian Journal of Forest Research*, New Westminster, v.39, n.12, p.2283-2293, 2009.
- MISIR, N. Generalized height-diameter models for *Populus tremula* L. stands. *African Journal of Biotechnology*, Mubarak. v.9, n.28, p.4348-4355, 2010.
- NANOS, N.; CALAMA, R., MONTERO, G.; GIL, L. Geostatistical prediction of height/diameter models. *Forest Ecology and Management*. Amsterdã, v.195, n.1-2, p. 221-235, 2004.
- O'HANGAN, A. **Kendall's Advanced Theory of Statistics: Bayesian Inference**. London: John Wiley & Sons, 1994. v.2B, 448 p.
- OLIVEIRA, M.C.N.; BARBIN, D.; SILVA, E.F.; MACEDO, J.R.; CAPECHE, C.L.; HISSA, H.R. Estimação de parâmetros de uma função de covariância no estudo da variabilidade espacial do teor de cálcio no solo. *Revista Brasileira de Matemática e Estatística*, Jaboticabal, v.24, n.4. p.117-132, 2006.
- PÉLLICO NETTO, S.; BRENA, D. A. **Inventário Florestal**. Curitiba: UFPR, 1997. 316 p.
- REGALADO, L.B. **Composição e distribuição de aves passeriformes em uma parcela de mata do Morro de Araçoiaba (Floresta Nacional de Ipanema, Iperó/SP) utilizando um sistema de informação geográfica**. 1999. 118p. Dissertação (Mestrado em Engenharia Ambiental) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 1999.
- RIBEIRO, A.; FERRAZ FILHO, A.C.; MELLO, J.M.; FERREIRA, M.Z.; LISBOA, P.M.M.; SCOLFORO, J.R.S. Estratégias e metodologias de ajuste de modelos hipsométricos em plantios de *Eucalyptus* sp. *Cerne*, Lavras, v.16, n.1, p.22-31, 2010.
- RIBEIRO JÚNIOR, P.J.; DIGGLE, P.J. GeoR: a package for geostatistical analysis. *R-news*, Vienna, v.1, n. 2, p. 15-18, 2001.
- SCHMIDT, A.M.; SANSÓ, B. **Modelagem Bayesiana da Estrutura de Covariância de Processos Espaciais e Espaço-Temporais**. Caxambu: ABE, 2006. 151 p.
- SCHUMACHER, F.X.; HALL, F.S. Logarithmic expression of timber volume. *Journal of Agricultural Research*, Whashington, v.47, n.9, p.719-734, 1933.
- SOARES, C.P.B.; PAULA NETO, F.; SOUZA, A.L. **Dendrometria e Inventário Florestal**. Viçosa: UFV, 2007. 276p.
- SOUZA, P.C.; MARTOS, H.L. Estudo do uso público e análise ambiental das trilhas em uma unidade de conservação de uso sustentável: Floresta Nacional de Ipanema, Iperó – SP. *Revista Árvore*, Viçosa, v.32, n.1, p. 91-100, 2008.

Recebido em 24/05/2010
Aceito para publicação em 02/02/2011