

ROGÉRIO CARLOS COSTA DO AMARAL

**PROPOSIÇÃO DE MÉTODO MULTIVARIADO DE ESTRATIFICAÇÃO
VERTICAL EM FLORESTA NATURAL**

RECIFE-PE – AGOSTO/2011



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**PROPOSIÇÃO DE MÉTODO MULTIVARIADO DE ESTRATIFICAÇÃO
VERTICAL EM FLORESTA NATURAL**

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

Área de Concentração:
Desenvolvimento de Métodos estatísticos e Computacionais

Orientador: Prof. Dr. Rinaldo Luiz Caraciolo Ferreira

Co-orientadores: Prof. Dr. José Antônio Aleixo da Silva
Prof. Dr. Kleber Régis Santoro

RECIFE-PE – AGOSTO/2011

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**PROPOSIÇÃO DE MÉTODO MULTIVARIADO DE ESTRATIFICAÇÃO VERTICAL
EM FLORESTA NATURAL**

ROGÉRIO CARLOS COSTA DO AMARAL

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 22/08/2011 pela Comissão Examinadora.

Orientador:

Prof. Dr. Rinaldo Luiz Caraciolo Ferreira
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Prof. Dr. Borko Stosic
Universidade Federal Rural de Pernambuco

Profa. Dra. Tatijana Stosic
Universidade Federal Rural de Pernambuco

Prof. Dr. Francisco Tarcísio Alves Júnior
Universidade Federal Rural de Pernambuco

Às pessoas que mais amo :
Meus Pais: Reginaldo Peixe e
Carmen Costa (in memoriam).
Meus filhos: Tiffany, Ryan e
Victor.

DEDICO

Agradecimentos

Primeiramente agradeço a inteligência suprema, causa primária de todas as coisas, ao nosso Deus, razão pela qual estou aqui, concluindo este trabalho, que após muito sacrifício, finalizo nesta data. Em segundo lugar ao Professor Doutor Rinaldo Luiz Caraciolo Ferreira que dedicou parte do seu tempo para me orientar neste trabalho, tendo muita paciência e muito empenho. Também ao Professor Doutor Kleber Régis Santoro e ao Professor Doutor José Antônio Aleixo, por fazerem parte deste trabalho como co-orientadores. Aos amigos que, direta ou indiretamente contribuíram para a minha longa e proveitosa caminhada, não citarei nomes para evitar a não inclusão merecida de alguns. Em especial, não posso deixar de reconhecer a garra, o profissionalismo e a dedicação do Professor Doutor Eufrázio a quem tenho uma admiração muito grande. A toda a equipe que forma o departamento de Biometria e com muito carinho a Zuleide que com o seu astral sempre nos motivou e passou um pouco de sua alegria.

PROPOSIÇÃO DE MÉTODO MULTIVARIADO DE ESTRATIFICAÇÃO VERTICAL EM FLORESTA NATURAL

Autor: Ms. Amaral, Rogério Carlos Costa do

Orientador: Dr. Rinaldo Luiz Caraciolo Ferreira

Resumo

Neste trabalho objetivou-se propor uma nova metodologia de estratificação vertical de florestas naturais, baseada na altura total da árvore, a partir de aplicação de técnicas multivariadas. A metodologia foi comparada com outros métodos usados em estratificação florestal. Foram utilizados dados estruturais de um fragmento de Floresta Atlântica, obtidos por Costa Júnior (2006), conhecido localmente por Mata das Caldeiras, com área aproximada de 38,56 ha, localizado no município de Catende, PE, a aproximadamente 11 km da sede municipal, nas coordenadas 8°40' S e 35°35' W, a uma altitude de 327 m. Foram utilizadas medidas de altura, diâmetro e área basimétrica de 766 indivíduos arbóreos com circunferência a altura do peito (CAP), a altura de 1,30 m do solo, maior ou igual a 15 cm. Foi proposto um método multivariado de estratificação vertical a partir da utilização conjunta das análises de componentes principais, de agrupamento e discriminante, bem como aplicados os métodos de Vega (1966), da IUFRO, de Longhi (1980), de Souza e Leite (1993), de Guedes-Bruni et al. (2006), da IUFRO-LOREY e de Sanquetta (1995). A eficiência do método aplicado foi considerada quando houve diminuição do CV dentro do estrato após a análise discriminante. Observou-se maiores ganhos de precisão em termos de coeficiente de variação para o método multivariado proposto em relação aos demais quando da aplicação da análise discriminante. No entanto, as dificuldades de estratificação vertical são claras já que os ganhos em coeficiente de variação variam muito conforme metodologia utilizada. O método multivariado proposto pode ser utilizado para estratificação vertical de florestas naturais e apresenta como vantagem principal a retirada de subjetividade do pesquisador.

Palavras-chave: análise de agrupamento, análise de componentes principais, análise discriminante

MULTIVARIATE METHOD PROPOSITION OF VERTICAL STRATIFICATION IN NATURAL FOREST

Author : Ms. Do Amaral, Rogério Carlos Costa

Tutor: Dr. Rinaldo Luiz Caraciolo Ferreira

Abstract

The objective of this research was to propose a new methodology for vertical stratification of natural forests, based on total tree height, from the application of multivariate techniques. The methodology was compared to others methods used in forest stratification. The structural data used in this work came from a research by Costa Junior (2006) of an Atlantic forest fragment known as 'Mata das Caldeiras', with 38.56 ha, located in Catende, PE. The measures used were the total tree height, diameter of breast height (DBH) and basal area of 766 trees with circumference at breast height (CBH) greater than or equal to 15 cm. The proposed methodology of multivariate vertical stratification resulting from the combination of principal component analysis, cluster and discriminant analysis was compared to the methods of Vega (1966), IUFRO, Longhi (1980), Sousa and Leite (1993), Guedes-Bruni et al. (2006), IUFRO-Lorey and Sanquetta (1995). The efficiency of the method used was considered when there was a decrease in the coefficient of variation (CV) within the stratum after the use of discriminant analysis. It was observed significant gains in accuracy of CV for the multivariate method proposed in relation to the others when the discriminant analysis was applied. However, the difficulties of vertical stratification are clear because the gains in CV vary widely depending on the methodology used. The new methodology can be used for vertical stratification of natural forests and presents as main advantage the removal of the researcher's subjectivity.

Keywords: Cluster Analysis, Principal Component Analysis, Discriminant Analysis.

LISTA DE FIGURAS

- Figura 1 - Relação alométrica entre diâmetro e altura, dos indivíduos com DAP ≥ 5 , amostrados na floresta aluvial, para definição dos limites do dossel (Fonte: GUEDES-BRUNI et al., 2006).8
- Figura 2. Localização da área de estudo, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE..... 14
- Figura 3: Histogramas das Alturas (A), dos Diâmetros a 1,30 m do solo - DAP (B) e da Área Basimétrica - G (C) para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE23

LISTA DE TABELAS

Tabela 1: Listagem das espécies arbóreas com respectiva média e coeficiente de variação do diâmetro a 1,30 m do solo (DAP), da altura e da área basimétrica para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE.....	15
Tabela 1: Continuação	16
Tabela 2 : Valor mínimo e máximo, média, desvio-padrão e coeficiente de variação da altura, do diâmetro a 1,30 m do solo (DAP) e da área basimétrica para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE.....	23
Tabela 3: Estimativas das variâncias (autovalores, λ_i) associadas aos componentes principais e suas importâncias relativas e acumuladas.....	24
Tabela 4: Número de indivíduos, valor mínimo e máximo, média, desvio-padrão e coeficiente de variação da altura, do diâmetro a 1,30 m do solo (DAP) e da área basimétrica dos estratos obtidos pelo algoritmo de Ward, com base na distância euclidiana quadrada, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE	25
Tabela 5: Número de árvores, amplitude de classe de altura, coeficiente de variação, classificação inicial por meio da análise de agrupamento, classificação final e correta conforme análise discriminante, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE.....	25
Tabela 6: Número de árvores, amplitude e coeficiente de variação na classificação inicial e final após análise discriminante conforme o método empregado	26
Tabela 6: Continuação	27

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE TABELAS	ix
SUMÁRIO.....	x
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA.....	3
2.1. Métodos de Estratificação Vertical.....	3
2.2. Análise Multivariada aplicada à Estratificação Vertical.....	8
3. MATERIAL E MÉTODOS.....	14
3.1. Proposta de Método Multivariado de Estratificação Vertical (MMEV).....	16
3.1.1. Análise de Componentes Principais	17
3.1.2. Análise de Agrupamento	18
3.1.3. Análise Discriminante	19
3.2. Método de Vega (1966)	20
3.3. Método de Souza e Leite (1993).....	20
3.4. Diagrama h-M (SANQUETTA, 1995)	20
3.5. Classificação da IUFRO (1958).....	21
3.6. Método de LONGHI (1980)	21
3.7. Método proposto por GUEDES-BRUNI et al. (2006).....	21
3.8. Classificação da IUFRO baseado na altura de Lorey	21
3.9. Comparação entre os métodos estudados	22
4. RESULTADOS E DISCUSSÃO	23
5. CONSIDERAÇÕES FINAIS	28
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	29

1. INTRODUÇÃO

Nas florestas tropicais é característica a presença de um dossel, formado por espécies capazes de atingir elevadas alturas, e de um sub-bosque, formado por espécies tolerantes a sombra, de baixa estatura (GOURLET-FLEURY et al., 2005). Por isso, existe a procura por métodos que possa simplificar a sua estrutura e facilitar as interpretações quanto ao funcionamento desses ecossistemas. Dentre estes métodos se pode destacar a estratificação da floresta quanto as variáveis relacionadas com o tamanho das árvores.

Em uma floresta natural é comum distribuir as suas plantas conforme uma arquitetura de estratificação, ou seja, por meio de uma divisão em estratos, pois segundo Roberts e Gilliam (1995), esta arquitetura é um fator importante para manter maior diversidade de espécies lenhosas, bem como subsidiar a gestão florestal. No entanto, o reconhecimento de estratos ainda é assunto muito controverso na literatura (VALE et al. 2009).

Para Calegário et al. (1994), o estudo da estratificação vertical é importante por possibilitar a produção de informações relativas a dominância que determinadas espécies exercem sobre as outras.

Cada estrato é definido pelo pesquisador por meio de uma regra de classificação baseada em uma ou várias variáveis de interesse. Dentre essas regras, uma das mais utilizada é a classificação dos indivíduos em relação à altura das árvores, desde que foi enfatizada por Richards (1952) no conceito de estratificação da floresta tropical.

Na fitossociologia ,essa estratificação em alturas, é conhecida como Posição Sociológica, e no seu estudo se objetiva informações sobre a composição florística dos distintos estratos da floresta em sentido vertical, e sobre o papel que representam as diferentes espécies em cada um deles (LAMPRECHT, 1990).

A maioria das regras utilizadas para a estratificação em alturas é subjetiva e baseada apenas nesta variável e na experiência do pesquisador. No entanto, o desenvolvimento de uma floresta depende de muitas variáveis, o que caracteriza a necessidade de um enfoque multivariado, pois se deve considerar simultaneamente um grande número de variáveis aleatórias relacionadas, onde cada variável tem a mesma importância no início da análise.

O objetivo do presente trabalho foi propor metodologia de estratificação baseada na altura a partir da aplicação de técnicas multivariadas e realizar a sua comparação com outros métodos encontrados na literatura.

2. REVISÃO DE LITERATURA

2.1. Métodos de Estratificação Vertical

Segundo LAMPRECHT (1964), a definição dos estratos que podem existir em uma floresta tropical, torna-se às vezes muito difícil de distinguir, devido principalmente à sua diversidade e a grande mistura de espécies em direção vertical.

Para HOSOKAWA (1998), estrato é definido como a porção de massa vegetal contida dentro de um limite de altura determinada, variando com a concepção pessoal. Geralmente distinguem-se os estratos superior, médio e inferior (VEGA, 1966; LONGHI, 1980), podendo-se acrescentar ainda, o sub-bosque (LAMPRECHT, 1964).

Na literatura existem vários métodos para a estratificação vertical da floresta baseado na altura dos indivíduos arbóreos, tais como, diagramas de perfil (DAVIS e RICHARDS, 1933), classificação da IUFRO (1958), método M-w (HOZUMI, 1975), método de Longhi (1980), método de Calegário et al. (1994), método h-M (SANQUETTA, 1995), Método proposto por SOUZA e LEITE (1993), Método proposto por SOUZA et al. (2003), Método proposto por GUEDES-BRUNI (2006).

a) Diagramas de Perfil

Os diagramas de perfil é a técnica mais antiga utilizada para a estratificação vertical de florestas. Eles foram propostos por Davis e Richards (1933) como ferramenta para a visualização da estrutura fisionômica das florestas e segundo Richards (1996), apesar de trabalhosa, é bastante útil para o estudo de estratificação e das diferenças na estrutura entre tipos de florestas. Esta técnica se caracteriza pela elaboração de histogramas de distribuição de frequência das classes de alturas máximas de copa, cuja amplitude é definida subjetivamente pelo pesquisador. É uma técnica bastante simples, no entanto, é muito dependente da tomada de decisão do pesquisador, principalmente na definição do número de estratos.

Lemos et al. (2001) em estudo de dois trechos de mata litorânea no município de Maricá - RJ, se utilizaram desta técnica e consideraram que foi um instrumento

útil na comparação de estruturas fisionômicas florestais. Jamir et al. (2006) também se utilizaram do diagrama de perfil para realizarem seus estudos em floresta úmida no nordeste da Índia.

b) Classificação da IUFRO (1958)

A classificação da IUFRO foi proposta por Leibundgut em 1958 (LAMPRECHT, 1990), e se baseia na altura dominante (h_{dom}) da floresta. Os estratos são definidos como inferior ($h_i < \frac{h_{dom}}{3}$), médio ($\frac{h_{dom}}{3} \leq h_i < \frac{2 \cdot h_{dom}}{3}$) e superior ($h_i \geq \frac{2 \cdot h_{dom}}{3}$) em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato.

c) Método de Vega (1966)

O método de estratificação proposto por Vega (1966) consiste em verificar a amplitude de variação em altura dos indivíduos amostrados e dividi-la em três partes para definição dos estratos inferior ($h_i < h_{mín} + \frac{A}{3}$), médio ($h_{mín} + \frac{A}{3} \leq h_i \leq h_{mín} + \frac{2A}{3}$) e superior ($h_i \geq h_{mín} + \frac{2A}{3}$), em que h_i = altura do i-ésimo indivíduo a ser classificado no estrato e A = diferença entre as alturas máxima ($h_{máx}$) e mínima ($h_{mín}$) amostradas.

d) Diagrama M-w (HOZUMI, 1975)

O diagrama M-w foi proposto por Hozumi (1975) e é baseado no método MNY (Hozumi, 1968). Para obtenção da estratificação se faz necessário a definição de: $f \rightarrow \phi(w)dw$: função de densidade da variável que caracteriza o tamanho individual das árvores (w) numa certa unidade de área ; $Y(w) = \int_w^{w_{máx}} w\phi(w)dw$: número cumulativo de árvores de tamanho w a $w_{máx}$ por unidade de área; $N(w) = \int_w^{w_{máx}} \phi(w)dw$: produção cumulativa de árvores de tamanho w a $w_{máx}$ por

unidade de área; e $M(w) = \frac{Y(w)}{N(w)} = \frac{\int_w^{w_{\max}} w \phi(w) dw}{\int_w^{w_{\max}} \phi(w) dw}$: valor médio de produção de uma população parcial de tamanho w a w_{\max} por unidade de área.

O relação entre M e w (abscissa) plotado em escala log-log é denominada de diagrama M - w . Os segmentos sobre o diagrama podem ser definidos por meio das seguintes expressões (HOZUMI, 1975):

$$M(w) = Aw + B$$

$$M(w) = Cw^b$$

em que: A , B , C e b são parâmetros das equações, considerando-se uma função de densidade do tipo beta. Assim, os limites dos estratos é definido como relação entre a primeira (S_1) e a segunda (S_2) derivadas de M em função de w sobre o diagrama M - w , ou seja:

$$S_1 = \frac{\partial \ln M}{\partial \ln w}$$

$$S_2 = \frac{\partial}{\partial \ln w} \left(\frac{\partial \ln M}{\partial \ln w} \right) = \frac{\partial S_1}{\partial \ln w}$$

Exemplos de aplicação do Diagrama M - w podem ser vistos em Feroz e Hagihara (2008), Feroz et al. (2006a,b; 2008; 2009) e Wu et al. (2010).

e) Método de LONGHI (1980)

Longhi (1980) propôs um critério para determinação dos estratos baseado nas freqüências relativas das alturas. Tal procedimento consiste na determinação da porcentagem das freqüências das alturas de todas as árvores encontradas e por meio das respectivas porcentagens acumuladas, estabelece-se uma curva onde se define a relação existente entre a freqüência das alturas acumuladas e a altura total. Finalmente, delimita-se os respectivos estratos considerando o critério de que cada estrato deverá abranger 1/3 das alturas encontradas, ou seja, a altura correspondente a 33,33% das freqüências acumuladas é o limite entre o estrato médio e o inferior, e a altura correspondente a 66,66% dessa freqüência é o limite entre os estratos médio e superior. Esta metodologia foi utilizada por Rivera Herrera et al. (2009).

f) Método Proposto por Calegário et al. (1994)

Proposto por Calegário et al. (1994) e utilizado por Coraiola (1997), este método tem a seguinte metodologia para estratificação de florestas naturais heterogêneas:

- 1) Ordenar da menor para a maior as alturas dos indivíduos amostrados;
- 2) Utilizar o teste χ^2 , a um nível de significância α , para a verificação da existência ou não de grupos de dados de altura estatisticamente homogêneos, conforme a expressão:

$$\chi_i^2 = \left(\frac{S_i^2}{\bar{H}_i} \right) \times (i - 1)$$

em que:

χ_i^2 = Valor do χ^2 calculado para o i-ésimo grupo de indivíduo;

S_i^2 = Variância das alturas dos indivíduos dentro do i-ésimo grupo;

\bar{H}_i = Média aritmética das alturas dos indivíduos dentro do i-ésimo grupo.

Com este procedimento são realizadas sucessivas estimativas dos valores de χ^2 para diversos grupos de indivíduos até que o valor calculado supere o valor tabelado, a um nível α de significância e com (i-1) graus de liberdade, concluindo-se que o i-ésimo grupo de indivíduos é estatisticamente homogêneo e discrimina-se como o estrato inferior. Assim, este estrato estará no intervalo definido pela menor altura amostrada e o valor da i-ésima altura.

- 3) Realizar o procedimento da letra b até que o número de estratos estabelecidos na pesquisa sejam determinados.

g) Diagrama h-M (SANQUETTA, 1995)

O Diagrama h-M foi proposto por Sanquetta (1995) e utilizado por Coraiola (1997) e Pizzato (1999) a partir de adaptações do Método MNY de Hozumi (1975) e possibilita o reconhecimento dos diferentes estratos em um povoamento como um todo e também para populações específicas. As adaptações realizadas foram:

- 1) Ao invés da variável peso, ou da sua correspondente do DAP elevado ao expoente 5/2, utilizou-se a variável altura (h);

- 2) A posição de M e h foi invertida em relação ao Diagrama M-w, ou seja, na abscissa grafou-se h e na ordenada M;
- 3) Não se utilizou a escala log-log, como originalmente concebido, mas sim se manteve a escala natural.

h) Método proposto por SOUZA e LEITE (1993)

No método proposto por Souza e Leite (1993) e utilizado por Mariscal - Flores (1993) se considera que as alturas dos indivíduos arbóreos têm função de densidade do tipo normal. Os estratos são definidos com base na altura média aritmética e no respectivo desvio-padrão. Assim, os estratos são definidos como inferior ($h_i < \bar{h} - s$), médio ($\bar{h} - s \leq h_i \leq \bar{h} + s$) e superior ($h_i > \bar{h} + s$) em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato; \bar{h} = altura média aritmética para a amostra; e s = estimativa do desvio-padrão da amostra.

j) Método proposto por GUEDES-BRUNI et al. (2006)

No método proposto por Guedes - Bruni et al. (2006) para definição de dossel são estabelecidas relações alométricas entre diâmetro e altura, transformados pelo respectivo logaritmo decimal, para todos os indivíduos coletados, identificando-se o dossel a partir da linha indentada surgida no diagrama de correlação. Com isto, a partir da verificação da primeira descontinuidade entre os pontos, são fixados como elementos de dossel todos os indivíduos que apresentarem diâmetros e alturas iguais ou superiores ao ponto de interseção dos eixos log dap e log altura (Figura 1).

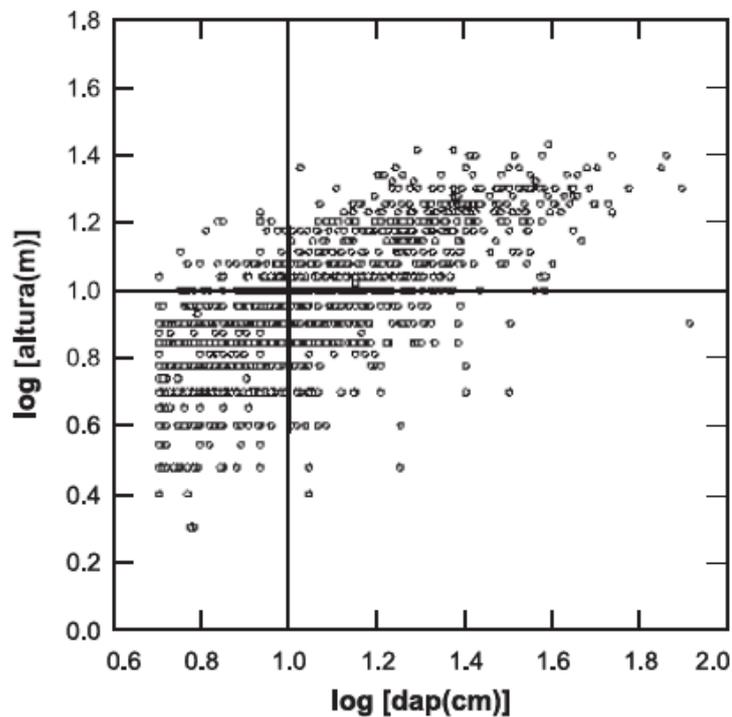


Figura 1 - Relação alométrica entre diâmetro e altura, dos indivíduos com DAP ≥ 5 , amostrados na floresta aluvial, para definição dos limites do dossel (Fonte: GUEDES-BRUNI et al., 2006).

Nesta pesquisa feita, não foram encontradas aplicações para os métodos propostos por: Vega (1966) , Guedes - Bruni (2006) e IUFRO-Lorey .

Para a maioria dos métodos descritos se observa que a estratificação é baseada apenas na variável altura. No entanto, o comportamento da floresta e de suas espécies é dependente de várias variáveis, tanto de natureza biótica como abiótica, que caracterizam a necessidade de um enfoque multivariado no estudo da estratificação vertical.

2.2. Análise Multivariada aplicada à Estratificação Vertical

A denominação “Análise Multivariada” corresponde a um grande número de técnicas que utilizam, simultaneamente, todas as variáveis na interpretação teórica do conjunto de dados obtidos.

Em quase todas as áreas de pesquisa várias variáveis são mensuradas e, em geral, essas devem ser analisadas conjuntamente. A análise multivariada é a área da estatística que trata desse tipo de estudo e existem várias técnicas que podem

ser aplicadas, sendo que, a utilização dessas depende do tipo de dado que se deseja analisar e dos objetivos do estudo.

Segundo Anderson (1984), basicamente, existe duas formas de classificar as análises multivariadas: as que permitem extrair informações a respeito da independência entre as variáveis que caracterizam cada elemento, tais como análise fatorial, análise de agrupamento, análise canônica, análise de ordenamento multidimensional e análise de componentes principais; e as que permitem extrair informações a respeito da dependência entre uma ou mais variáveis ou uma com relação à outra, tais como análise de regressão multivariada, análise de contingência múltipla, análise discriminante e análise de variância multivariada.

Na Ciência Florestal são utilizados vários métodos multivariados para análise de vários problemas, com destaque para as análises de componentes principais, de agrupamento e discriminante (SANTOS et al., 2004; SOUZA et al., 2004; ALBUQUERQUE et al. 2006; SOUZA et al. 2006; FERREIRA et al. 2008; LIMA JÚNIOR et al., 2009; BENITES et al., 2010). No entanto, especificamente para estratificação vertical da floresta há poucos relatos na literatura, como exemplo se pode citar o trabalho de Souza et al. (2003), que utilizaram as análises de agrupamento e discriminante para fazer uma proposição metodológica e Guilherme et al. (2004) que utilizaram a análise de correspondência retificada para gerar três estratos arbitrários em floresta atlântica no Sudeste do Brasil.

a) Análise de componentes principais

A análise de componentes principais é uma técnica estatística poderosa que pode ser utilizada para redução do número de variáveis e para fornecer uma visão estatisticamente privilegiada do conjunto de dados.

A análise por componentes principais consiste em transformar um conjunto original de variáveis (por exemplo, altura, produção etc.) em outro conjunto de dimensão equivalente, mas com propriedades importantes. Cada componente principal é uma combinação linear das variáveis originais. Além disso, são independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo da informação, em termos de variação total, contida nos dados iniciais (CRUZ REGAZZI, 1997).

A viabilidade da utilização desta técnica dependerá da possibilidade de reduzir o número de variáveis estudadas em poucos componentes principais, ou seja, reduzir de um espaço p -dimensional para um espaço bi ou tridimensional, com a menor perda de informação possível. Assim, é permitido descartar variáveis que contribuem pouco para a discriminação do material avaliado e agrupar os tratamentos similares, mediante exames visuais em dispersões gráficas. Esta técnica leva em consideração a distância euclidiana e pode ser obtida sem a realização de um experimento com repetições (RIBEIRO JÚNIOR, 2001).

A análise em componentes principais estabelece, com base em uma matriz de semelhança (correlações, variâncias-covariâncias ou até mesmo de similaridades), um conjunto de eixos (componentes ou fatores) perpendiculares.

Cada componente corresponde a um autovetor dessa matriz (VALENTIN, 2000).

Os objetivos desta análise são: i) redução da dimensão original; e ii) facilitação da interpretação das análises realizadas. Em geral, a explicação de toda a variabilidade do sistema determinado por p variáveis só pode ser efetuada por p componentes principais.

Os componentes principais são uma técnica de análise intermediária e, portanto não se constituem em um método final e conclusivo. Esse tipo de análise se presta fundamentalmente como um passo intermediário em grandes investigações científicas (FERREIRA, 1996).

b) Análise de Agrupamento

A análise de agrupamento é uma técnica multivariada que tem por objetivo proporcionar uma ou várias partições na base de dados, em grupos, por algum critério de classificação, de tal forma que exista homogeneidade dentro e heterogeneidade entre grupos.

Por meio dessa técnica se pode obter uma sumarização dos dados e assim facilitar a sua interpretação, já que se utiliza de métodos que procuram grupos excludentes, ascendentes, reduzindo a dimensão de um conjunto de n indivíduos para informações de um novo conjunto de g grupos, sendo g significativamente menor que n (MARDIA et al., 1997).

As técnicas de análise de agrupamento exigem do pesquisador decisões quanto à escolha da medida de dissimilaridade e a definição do número de grupos, bem como o conhecimento das propriedades dos algoritmos de agrupamento. A tomada dessas decisões pode representar diferentes agrupamentos (DUARTE et al., 1999).

Segundo Aaker et al. (2001), existem duas premissas importantes na análise de agrupamento. A primeira é a de que a medida utilizada seja uma medida válida de similaridade ou dissimilaridade entre os indivíduos. E a segunda premissa é a de que exista uma justificativa teórica para estruturar os indivíduos em grupos.

Geralmente, é difícil avaliar a qualidade do processo de agrupamento, pois não existem testes estatísticos padrões para garantir que o resultado seja puramente aleatório. Para Bussab et al. (1990), a definição da medida a ser utilizada, a legitimidade do resultado, a aparência de uma hierarquia natural (quando for empregado um método não hierárquico) e a confiabilidade de testes de divisão de amostra, oferecem informações úteis. No entanto, é difícil determinar quais os grupos são muito similares e quais objetos são difíceis de serem agrupados.

Segundo Aaker et al. (2001), a análise de agrupamento pode ser resumida em cinco etapas:

1. A seleção de indivíduos ou de uma amostra de indivíduos a serem agrupados;
2. A definição de um conjunto de variáveis a partir das quais serão obtidas informações necessárias ao agrupamento dos indivíduos;
3. A definição de uma medida de semelhança ou distância entre os indivíduos;
4. A escolha de um algoritmo de partição/classificação;
5. Por último, a validação dos resultados encontrados.

Vários são os tipos de técnicas de agrupamento encontradas na literatura (MARDIA et al., 1997), tendo o pesquisador que tomar a decisão de qual é a mais adequada ao objetivo do seu trabalho, já que as diferentes técnicas podem levar a diferentes soluções.

As próprias técnicas de agrupamento podem ser “classificadas” em grupos, e diferentes autores produzem diferentes classificações. Cormack (1971) propôs à seguinte: a) Técnica hierárquica de agrupamento (aglomerativas ou divisivas) - consiste em uma série de sucessivos agrupamentos ou sucessivas divisões de elementos, em que os elementos são agregados ou desagregados. b) Técnicas não-

hierárquicas, ou por particionamento - desenvolvidas para agrupar elementos em K grupos, em que K é a quantidade de grupos definida previamente.

c) Análise Discriminante

Para Johnson e Wichern (2007), a análise discriminante ou de classificação, trata dos problemas relacionados em separar conjuntos distintos de objetos (observações) e alocar novos objetos (observações) em grupos previamente definidos.

Segundo Mardia et al. (1997), pode-se também, por meio da análise discriminante, testar a suficiência de uma série de variáveis discriminantes. A suficiência é testada pela análise discriminante *stepwise*, de grande utilidade, principalmente, quando há mais variáveis do que fossem necessárias para se obter uma classificação satisfatória.

Como procedimento separatório, a análise discriminante é empregada com a finalidade de investigar as diferenças observadas, caso estas não sejam bem entendidas. Para Marriott (1974), a análise discriminante consiste em investigar como e quando é possível fazer distinções entre os membros de “g” agrupamentos, com base nas observações feitas sobre eles.

Para LACHENBRUCH (1979), os objetivos da análise discriminante são:

- Testar diferenças estatísticas, significantes a um dado nível de probabilidade entre “g” agrupamentos;
- Determinar o número de funções discriminantes;
- Construir regras de alocações para identificar um indivíduo como membro de um dos “g” agrupamentos e estimar as probabilidades de classificações corretas.

Uma das principais finalidades da análise discriminante é obter funções que permita classificar um indivíduo X, com base em medidas de um número p de características do mesmo, em uma de várias populações π_i , ($i=1,2,\dots,g$), distintas, buscando minimizar a probabilidade de classificar erroneamente um indivíduo, em uma população π_i , quando ele realmente pertence à população π_j , ($i \neq j=1,2,\dots,g$).

Para a classificação em uma de várias populações, recomenda-se o uso da função discriminante de Anderson (1984), no qual se tem uma observação x e deseja-se classificá-la em uma dentre g populações, sendo $g > 2$.

O resultado da classificação dos indivíduos levantado em cada grupo indica a precisão do agrupamento por meio de uma Tabela, o qual permite, além da identificação das árvores correta e incorretamente classificadas, também aponta os grupos com maior probabilidade de alocação, ou seja, aqueles de cujos centróides estão mais próximos (FERREIRA; LIMA, 1978).

3. MATERIAL E MÉTODOS

Este estudo foi realizado a partir de um banco de dados obtidos por Costa Júnior (2006) em um remanescente de Mata Atlântica denominado Mata das Caldeiras, localizado no município de Catende-PE, que está situado na mesorregião da Mata Pernambucana (Figura 2). A sede municipal está a 142 km da capital Recife, tendo seu acesso pela PE-126 e pela BR-101, estando a 169 m de altitude, nas coordenadas 8° 40' S e 35° 35' W. A superfície territorial do município é de 181 km², correspondendo a 1,6% da região fisiográfica, a 3,2% da microrregião homogênea na qual está inserido e 0,2% da área total do Estado de Pernambuco (CONDEPE, 1987).

Segundo a classificação de Köppen, o clima de Catende é do tipo As', tropical chuvoso com verão seco e estação chuvosa adiantada para o outono, antes do inverno. O período de maior umidade corresponde aos meses de abril a junho. A temperatura média anual ultrapassa 22°C e a precipitação média anual é de 1.414 mm. O relevo predominante varia de ondulado a forte ondulado. O solo predominantemente, encontrado na área geográfica em estudo, é do tipo Argilossolo Vermelho-amarelo distrófico.



Figura 2. Localização da área de estudo, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE.

Os dados foram obtidos a partir de 40 parcelas de 250 m² (10 x 25 m), alocadas sistematicamente ao longo de todo fragmento, distando 25 m entre si. Nas parcelas, foram amostrados apenas os indivíduos arbóreos vivos com CAP (circunferência a altura do peito – 1,30 m do solo) \geq 15 cm, que receberam placas

metálicas enumeradas e tiveram os seguintes dados anotados: o CAP, mensurado com fita métrica, e a altura, pela estimativa visual, utilizando como base as hastes do podão, as quais medem 2 m.

No presente estudo inicialmente foram analisados os 1.049 indivíduos arbóreos levantados por Costa Júnior (2006), distribuídos taxonomicamente em 90 espécies, 64 gêneros e 37 famílias botânicas e um táxon não identificado. A partir desta análise decidiu-se limitar a amostra a 766 indivíduos, 29 espécies, 27 gêneros e 18 famílias, pois se considerou que cada espécie deveria ter no mínimo 10 indivíduos amostrados (Tabela 1). Com a definição da amostra, foram realizadas as estimativas das estatísticas descritivas para as variáveis estudadas e suas correlações lineares por meio do coeficiente de Pearson.

Para este trabalho foram utilizados métodos de estratificação vertical identificados na revisão de literatura e feita uma proposição de um método multivariado a partir da utilização conjunta das análises de componentes principais, de agrupamento e discriminante.

Tabela 1: Listagem das espécies arbóreas com respectiva média e coeficiente de variação do diâmetro a 1,30 m do solo (DAP), da altura e da área basimétrica para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE

Família/Espécie	DAP(cm)		Altura(m)		Área basimétrica(m ²)	
	Média	CV (%)	Média	CV (%)	Média	CV (%)
Anacardiaceae						
<i>Tapirira guianensis</i> Aubl.	18,9	53,84	18,3	46,97	0,0360	101,54
<i>Thyrsodium spruceanum</i> Benth.	9,5	63,57	12,1	53,52	0,0099	151,48
Araliaceae						
<i>Schefflera morototoni</i> (Aubl.) Maguire, Steyer. & Frodin	16,1	74,81	16,4	37,75	0,0314	131,17
Bombaceae						
<i>Eriotheca gracilipes</i> (K. Schum.) A. Robyns	11,6	63,21	13,6	46,53	0,0145	151,06
Burseraceae						
<i>Protium heptaphyllum</i> (Aubl.) Marchand	8,0	50,97	10,8	48,54	0,0062	109,25
Caesalpiniaceae						
<i>Copaifera langsdorffii</i> Desf.	15,1	68,57	14,3	47,67	0,0257	130,44
<i>Dialium guianense</i> (Aubl.) Sandwith	14,2	64,10	14,9	52,91	0,0222	130,99
Cecropiaceae						
<i>Cecropia palmata</i> Willd.	13,8	61,50	12,4	40,40	0,0203	126,97
<i>Cedrela</i> sp. P. Browne	10,2	55,76	12,4	48,73	0,0106	129,08
Chrysobalanaceae						
<i>Licania rigida</i> Benth.	16,9	48,19	17,6	67,02	0,0271	79,13

Continua...

Tabela 1: Continuação

Família/Espécie	DAP (cm)		Altura(m)		Área basimétrica (m ²)	
	Média	CV (%)	Média	CV (%)	Média	CV (%)
Erythroxylaceae						
<i>Erythroxylum squamatum</i> Sw.	7,8	58,94	9,8	46,77	0,0064	144,22
Euphorbiaceae						
<i>Mabea occidentalis</i> Benth	13,8	78,57	16,3	39,28	0,0229	136,90
Fabaceae						
<i>Pterocarpus violaceus</i> Vogel	15,1	75,27	13,9	30,41	0,0271	139,95
Lauraceae						
<i>Nectandra cuspidata</i> (Nees et. Mart.) Nees	12,2	79,82	13,7	59,27	0,0189	169,96
<i>Ocotea gardneri</i> (Meisn.) Mez	12,0	73,57	14,6	47,25	0,0172	151,44
<i>Ocotea opifera</i> Mart.	11,4	74,46	11,2	41,65	0,0155	139,07
Lecythidaceae						
<i>Eschweilera ovata</i> (Cambess.) Miers	9,7	92,52	11,1	44,08	0,0134	290,78
Melastomataceae						
<i>Miconia albicans</i> (Sw.) Triana	7,5	29,88	10,2	28,04	0,0048	64,78
Mimosaceae						
<i>Inga thibaudiana</i> DC.	9,9	62,12	11,7	48,68	0,0105	133,36
<i>Parkia pendula</i> (Willd.) Benth. ex Walp.	33,4	64,04	16,9	40,64	0,1215	117,72
<i>Plathymentia foliolosa</i> Benth.	24,8	86,83	18,8	43,56	0,0837	190,20
<i>Stryphnodendron pulcherrimum</i> (Willd.) Hochr.	15,9	77,89	15,1	58,02	0,0307	150,71
Moraceae						
<i>Brosimum discolor</i> Schott	13,8	68,12	14,6	58,99	0,0218	136,14
<i>Helicostylis tomentosa</i> (Poepp. & Endl.) Rusby	10,0	75,79	12,7	47,70	0,0124	267,28
Ochnaceae						
<i>Ouratea hexasperma</i> (A. St.-Hil.) Baill.	8,6	63,37	11,1	50,87	0,0078	131,95
Sapotaceae						
<i>Chrysophyllum splendens</i> Spreng.	14,4	65,20	17,0	41,57	0,0225	127,54
<i>Pouteria grandiflora</i> (A. DC.) Baehni	9,9	51,97	10,8	39,29	0,0097	113,87
Sapindaceae						
<i>Cupania racemosa</i> (Vell.) Radlk.	8,0	48,62	9,3	52,15	0,0061	105,53
<i>Cupania revoluta</i> Rolfe	7,0	20,47	9,8	39,68	0,0040	42,83
Geral	13,2	80,10	13,8	52,50	0,0226	217,32

3.1. Proposta de Método Multivariado de Estratificação Vertical (MMEV)

Na presente proposição foram utilizadas as variáveis diâmetro a 1,30 m do solo (DAP), altura (H) e área basimétrica (G) dos indivíduos amostrados. Estas variáveis foram utilizadas em conjunto por meio da aplicação inicial da análise de componentes principais e obtenção dos escores das primeiras componentes que

explicaram pelo menos 80% da variação total. A partir da obtenção dos escores foi realizada análise de agrupamento, por meio da distância euclidiana quadrada e do método de Ward (WARD, 1963), obtendo-se um dendrograma e a classificação dos indivíduos em três grupos (estratos). Finalmente, realizou-se a análise discriminante para verificação da pertinência da classificação dos indivíduos dentro dos grupos.

3.1.1. Análise de Componentes Principais

A análise de componentes principais foi aplicada visando a redução do conjunto original de variáveis e de obtenção de um conjunto de combinações lineares. Os dados originais foram padronizados com vistas a minimizar os efeitos das diferentes escalas de mensuração. A padronização foi realizada utilizando-se a média X_{ij} da j -ésima variável ($j = 1, 2, \dots, m$) avaliada no i -ésimo indivíduo ($i = 1, 2, \dots, p$), gerando média zero e variância unitária (CRUZ e REGAZZI, 1997):

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_{X_j}}$$

Os componentes principais são combinações lineares de variáveis construídas de maneira a captar o máximo da variância, em que o primeiro componente explica a maior variação existente, o segundo a segunda maior variação, e assim sucessivamente. Diante dessa pressuposição, no estudo sobre classificação procurar-se-á estudar a dispersão das parcelas em sistemas de eixos cartesianos onde a variabilidade disponível fosse maximizada. A técnica consistiu na transformação de um conjunto de n variáveis padronizadas, $x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}$ em um novo conjunto $y_{i1}, y_{i2}, y_{i3}, \dots, y_{in}$, em que os y_{is} são funções lineares dos x_{is} e independentes entre si. As seguintes propriedades foram verificadas:

1 - Se Y_{ij} é um componente principal, então:

$$Y_{ij} = a_1x_{i1} + a_2x_{i2} + \dots + a_nx_{in}$$

2 - Se Y'_{ij} é outro componente principal, então:

$$Y'_{ij} = b_1x_{i1} + b_2x_{i2} + \dots + b_nx_{in}$$

Ou seja, os componentes são independentes.

3 - Os componentes principais foram obtidos pela solução do sistema:

$$\det (R - \lambda_i I)a = 0,$$

Em que: R = matriz de correlação entre as médias estimadas, λ_i = raízes características ou (autovalores) de R , I = matriz identidade de dimensão $p \times p$ e a = vetor característico (ou autovetor) associado aos autovalores (CRUZ e REGAZZI, 1997). Dessa forma, os autovalores de R correspondem às variâncias de cada componente e os autovetores normalizados, aos coeficientes de ponderação dos caracteres padronizados.

A importância relativa de um componente foi calculada por:

$$\text{Importância de } Y_j = \frac{\lambda_i}{\text{traço}(R)}$$

Após a determinação dos primeiros componentes principais, que explicaram um mínimo de 80% da variação disponível, foram estimados os respectivos escores associados a cada indivíduo.

3.1.2. Análise de Agrupamento

A análise de agrupamento foi realizada por meio da utilização da distância euclidiana quadrada e do método de Ward, considerando-se a divisão dos indivíduos em três classes de alturas, ou seja, estrato inferior, médio e superior.

A distância Euclidiana Quadrada foi utilizada como medida de distância para o Método de Ward, a qual foi obtida conforme a expressão:

$$d_{ij} = \sum_{j=1}^p (X_{ij} - X_{i'j})^2$$

Em que: d_{ij} = distância euclidiana quadrada entre o i -ésimo e o j -ésimo indivíduo; X_{ij} = escore obtido por meio da análise de componentes principais; $X_{i'j} = \dots$; p = número de indivíduos amostrados.

O método de Ward foi obtido a partir da execução das seguintes etapas (Orloci, 1978):

1 - Determinou-se a matriz de distâncias. Localizou-se os dois agrupamentos para os quais d_{ij} foi mínimo;

2 -. Reuniu-se estes agrupamentos, formando um novo agrupamento, e verificou-se se o número de agrupamentos (g), foi alcançado, senão seguiu-se à etapa 3, caso contrário, terminou-se a análise;

3 - Calculou-se o valor do aumento obtido na soma de quadrados pela reunião de qualquer dos agrupamentos: $I = \frac{1}{2} d_{pq}$

4 - Determinou-se os dois agrupamentos que apresentaram um menor incremento na matriz D , isto é, $Min(I_{ij})$ e voltou-se à etapa 2.

3.1.3. Análise Discriminante

A análise discriminante foi aplicada para verificar a pertinência dos grupos formados pela análise de agrupamento, de acordo com as funções discriminantes de Anderson (1984). Admitiu-se ainda, que a discriminação envolveu custos diferentes (pesos dos estratos) de má classificação e probabilidades conhecidas de várias populações.

Considerando que a função densidade de probabilidade associada à população p_i é normal multivariada, então a função discriminante foi dada por:

$$D_i(x) = -\frac{1}{2} \ln \left[\sum_i i \right] - \frac{1}{2} [x - \mu_i] \sum_i^{-1} [x - \mu_i] + \ln(p_i)$$

Pressupondo-se a igualdade das matrizes de covariâncias, então os componentes constantes de todo i podem ser retirados, e a função discriminante ficou assim:

$$D_i(x) = l'_i x - \frac{1}{2} l'_i \mu_i + \ln(p_i). \quad i = 1,2,3$$

Em que $l'_i = \sum x_i^{-1}$

A regra de decisão quanto à população de x foi a seguinte:

$$D(x) = \max(D_1(x), D_2(x), D_3(x))$$

Classificar x em Π_i se e somente se;

$$D_i(x) = l'_i x - \frac{1}{2} l'_i \bar{x}_i + \ln(p_i). \quad i = 1,2,3$$

Dessa forma, a partir das estimativas das matrizes de covariância e de médias foram obtidas regras de discriminação, objetivando-se verificar a probabilidade de má classificação de cada grupo formado na análise de agrupamento.

3.2. Método de Vega (1966)

O método de estratificação proposto por Vega (1966) foi aplicado da seguinte forma: estrato inferior ($h_i < h_{mín} + \frac{A}{3}$), médio ($h_{mín} + \frac{A}{3} \leq h_i \leq h_{mín} + \frac{2A}{3}$) e superior ($h_i \geq h_{mín} + \frac{2A}{3}$), em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato e A = diferença entre as alturas máxima ($h_{máx}$) e mínima ($h_{mín}$) amostradas.

3.3. Método de Souza e Leite (1993)

No método proposto por Souza e Leite (1993) os estratos foram definidos a partir do cálculo da altura média aritmética e do respectivo desvio-padrão. Assim, os estratos foram definidos como inferior ($h_i < \bar{h} - s$), médio ($\bar{h} - s \leq h_i \leq \bar{h} + s$) e superior ($h_i > \bar{h} + s$) em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato; \bar{h} = altura média aritmética para a amostra; e s = estimativa do desvio-padrão da amostra.

3.4. Diagrama h-M (SANQUETTA, 1995)

Para a obtenção desses estratos, as alturas foram colocadas em ordem crescente, em seguida foi criada uma coluna para o valor acumulado dessas alturas (Σh), e mais duas colunas para a freqüência relativa das alturas e o valor da freqüência acumulada (Σf). Foi calculado o valor de $M = \frac{\Sigma h}{\Sigma f}$ e foram plotados no gráfico, com a escala natural, os valores de h e os valores de M , e foram observados os pontos de descontinuidades onde percebemos os limites dos estratos.

3.5. Classificação da IUFRO (1958)

Inicialmente foi calculada a média aritmética entre as 100 árvores com maiores diâmetros (altura dominante) em seguida foram definidos os estratos seguindo a proposta por Leibundgut em 1958 (LAMPRECHT, 1990), como inferior ($h_i < \frac{h_{dom}}{3}$), médio ($\frac{h_{dom}}{3} \leq h_i < \frac{2 \cdot h_{dom}}{3}$) e superior ($h_i \geq \frac{2 \cdot h_{dom}}{3}$) em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato.

3.6. Método de LONGHI (1980)

Os estratos foram definidos considerando-se que a altura correspondente a 33,33% das freqüências acumuladas foi o limite entre o estrato médio e o inferior, e; a altura correspondente a 66,66% foi o limite entre os estratos médio e superior.

3.7. Método proposto por GUEDES-BRUNI et al. (2006)

No método proposto por Guedes - Bruni et al. (2006) ,para definição de dossel são estabelecidas relações alométricas entre diâmetro e altura, transformados pelo respectivo logaritmo decimal, desta forma, é percebido no gráfico de correlação , a formação de três estratos com os limites das alturas e dos diâmetros indicados pelas linhas indentadas.

3.8. Classificação da IUFRO baseado na altura de Lorey

Inicialmente foi calculada a altura de Lorey (h_L) em seguida foram definidos os estratos seguindo a regra: como inferior ($h_i < \frac{h_L}{3}$), médio ($\frac{h_L}{3} \leq h_i < \frac{2 \cdot h_L}{3}$) e superior ($h_i \geq \frac{2 \cdot h_L}{3}$) em que: h_i = altura do i-ésimo indivíduo a ser classificado no estrato. A altura de Lorey foi obtida conforme a seguinte expressão:

$$h_L = \frac{\sum_i^{766} h_i \cdot g_i}{\sum_i^{766} g_i}$$

Em que: h_L = altura de Lorey; h_i = altura do i-ésimo indivíduo; g_i = área basimétrica do i-ésimo indivíduo; $i = 1, 2, \dots, 766$.

Vale salientar que neste trabalho, também como proposta metodológica, a análise discriminante foi realizada para todos os métodos estudados.

3.9. Comparação entre os métodos estudados

Para comparação dos métodos foi calculado o coeficiente de variação dentro dos estratos para cada procedimento utilizado. Assim, a eficiência do método foi considerada quando houve diminuição do CV dentro do estrato após a análise discriminante. O CV foi estimado por meio da seguinte expressão:

$CV_i(\%) = \frac{s_i}{\bar{h}_i} \times 100$ em que: $CV(\%)$ = coeficiente de variação no i-ésimo estrato; s_i =

desvio-padrão no i-ésimo estrato; \bar{h}_i = média aritmética da altura no i-ésimo estrato.

As análises estatísticas foram realizadas por meio da utilização do software SPSS (*Statistical Package for Social Sciences*). versão 17.0 (SPSS, 2010).

4. RESULTADOS E DISCUSSÃO

Os dados estruturais do fragmento de Floresta Atlântica estudado apresentou uma concentração maior de árvores entre 5 e 20 m de altura (Figura 3 A), correspondendo a quase 80% da amostra, contudo tem-se mais de 50% das árvores com diâmetro abaixo de 10 cm (Figura 3 B), indicando assim, um número expressivo de árvores altas com diâmetros baixos.

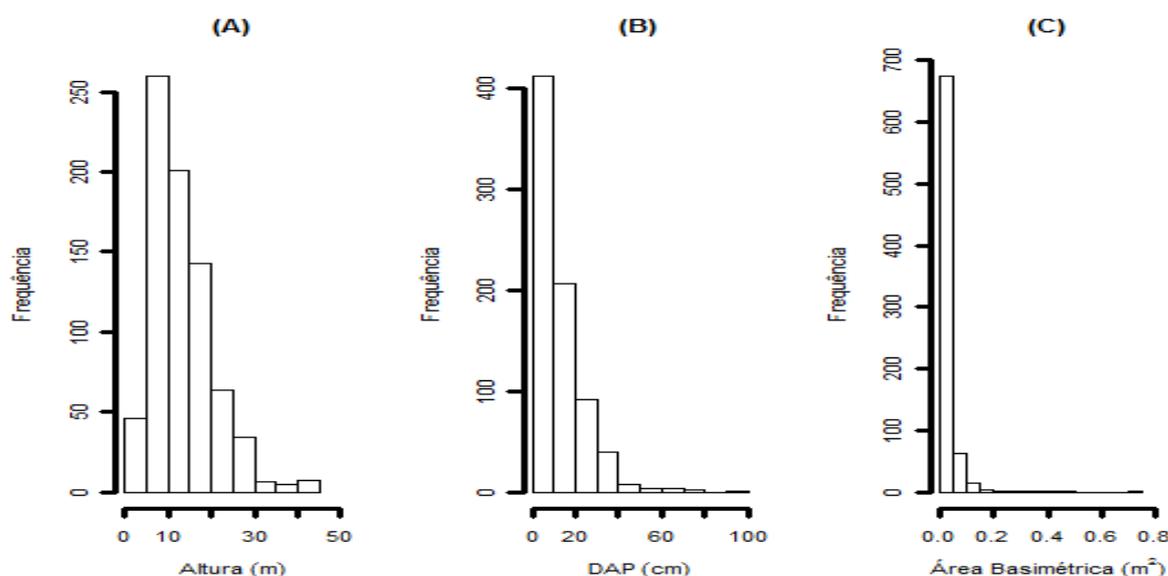


Figura 3: Histogramas das Alturas (A), dos Diâmetros a 1,30 m do solo - DAP (B) e da Área Basimétrica - G (C) para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE

Foram observadas as estatísticas descritivas das variáveis estudadas, com destaque para os coeficientes de variação cujos valores são normais em floresta natural e explicáveis pela diversidade de tamanhos encontrados, já que os indivíduos são inequidanos e de várias espécies, e também, possivelmente, pelos distúrbios antrópicos, com a ocorrência de poucos indivíduos com altura e diâmetros maiores (Tabela 2).

Tabela 2 : Valor mínimo e máximo, média, desvio-padrão e coeficiente de variação da altura, do diâmetro a 1,30 m do solo (DAP) e da área basimétrica para amostra de 766 indivíduos arbóreos, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE

Variável	Mínimo	Máximo	Média	Desvio Padrão	Coefficiente de Variação (%)
Altura (m)	3	45	13,82	7,26	52,53
Diâmetro (cm)	4,78	95,49	13,24	10,60	80,06
Área Basimétrica (m ²)	0,0018	0,7162	0,0226	0,0491	217,26

Observou-se correlação linear significativa ($P < 0,05$) entre as variáveis estudadas. Entre o diâmetro e a altura a correlação foi igual a 0,72; entre diâmetro e área basimétrica foi de 0,90 e entre altura e área basimétrica foi de 0,54. Bisquerra et al. (2004) mencionam que coeficientes de correlação entre 0,60 e 0,80 representam uma alta correlação e de 0,80 a 1, correlação muito alta. Vale salientar que a condição de correlação significativa entre as variáveis estudadas é necessária para a realização da análise de componentes principais (JOHNSON e WICHERN, 2007).

A primeira componente explica mais de 80%, no entanto utilizaram-se as duas primeiras componentes principais que explicam mais de 97% da variância total, tornando possível a redução para a segunda dimensão, já que o grau de distorção foi de apenas 2,16%, e considerado satisfatório (Tabela 3).

Tabela 3: Estimativas das variâncias (autovalores, λ_i) associadas aos componentes principais e suas importâncias relativas e acumuladas

Componente Principal	Autovalores (λ_i)	Proporção da Variância (%)	Proporção da Variância Acumulada (%)
1	2,453	81,77	81,77
2	0,482	16,07	97,84
3	0,065	2,16	100,00

Assim, justifica-se que apenas os escores obtidos a partir das duas componentes principais obtidas $Y_1 = 0,821H + 0,973DAP + 0,912G$ e $Y_2 = 0,568H - 1,22DAP - 0,381G$ podem ser utilizados nas análises de agrupamento e discriminante.

Por meio da análise de agrupamento, utilizando-se como medida de dissimilaridade a distância Euclidiana Quadrada e o algoritmo de agrupamento de Ward obteve-se a divisão em estratos inferior, médio e superior constantes na Tabela 4.

Tabela 4: Número de indivíduos, valor mínimo e máximo, média, desvio-padrão e coeficiente de variação da altura, do diâmetro a 1,30 m do solo (DAP) e da área basimétrica dos estratos obtidos pelo algoritmo de Ward, com base na distância euclidiana quadrada, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE

Estrato	Variável	Número de Indivíduos	Mínimo	Máximo	Média	Desvio Padrão	Coeficiente de Variação (%)
Inferior	Altura (m)		3	14	8,51	2,42	28,43
	Diâmetro(cm)	401	4,78	24,03	7,39	3,33	45,06
	Área Basimétrica (m ²)		0,0018	0,0454	0,0052	0,0061	117,31
Médio	Altura (m)		10	45	19,56	6,10	31,19
	Diâmetro(cm)	360	5,09	62,39	18,89	10,02	53,04
	Área Basimétrica (m ²)		0,0020	0,3057	0,0359	0,0402	111,98
Superior	Altura (m)		15	45	27,00	11,51	42,63
	Diâmetro(cm)	5	66,85	95,50	76,20	11,54	15,14
	Área Basimétrica (m ²)		0,3509	0,7162	0,4644	0,1481	31,89

Observa-se que existe sobreposição das classes de altura, ou seja, entre os estratos inferior, médio e superior, indicando que não há uma separação clara dos indivíduos. Tal resultado não é adequado para estratificação vertical de uma floresta. No entanto, observa-se que com aplicação da análise discriminante foi possível uma rearrumação dos grupos de forma que não houve interseção entre as amplitudes de classes (Tabela 5).

Tabela 5: Número de árvores, amplitude de classe de altura, coeficiente de variação, classificação inicial por meio da análise de agrupamento, classificação final e correta conforme análise discriminante, fragmento de Floresta Atlântica, Mata das Caldeiras, município de Catende, PE

Método Multivariado							
Estrato	Classificação Inicial			Classificação Final			Classificação Correta (%)
	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	
1	401	3-14	28,44	447	3-14	30,12	100
2	360	10-45	31,20	307	15-33	22,08	83,7
3	5	15-45	42,63	12	35-45	9,07	20

Comparando-se os métodos (Tabela 6), observa-se que o de Vega e Longhi, apenas no estrato 3 houve uma diminuição do coeficiente de variação, enquanto nos métodos da IUFRO, Souza e Leite e IUFRO-Lorey houve uma diminuição muito pequena nos estratos 1 e 3, no método de Sanquetta apesar de não haver um aumento no coeficiente de variação, observa-se uma má formação dos estratos, ficando o estrato 2 com 702 árvores de um total de 766. Guedes - Bruni e o método multivariado apresentou maiores ganhos de precisão em termos de coeficiente de

variação em relação aos demais quando da aplicação da análise discriminante. No entanto, as dificuldades de estratificação vertical são claras já que os ganhos em coeficiente de variação variam muito de método para método. Por outro lado, pode-se afirmar que o método multivariado proposto tem como vantagem a retirada de subjetividade do pesquisador que existe na maioria dos métodos.

Tabela 6: Número de árvores, amplitude e coeficiente de variação na classificação inicial e final após análise discriminante conforme o método empregado

Método Multivariado de Estratificação Vertical (MMEV)								
Classificação Inicial				Classificação Final				Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)		
1	401	3-14	28,44	447	3-14	30,12	100	
2	360	10-45	31,20	307	15-33	22,08	83,7	
3	5	15-45	42,63	12	35-45	9,07	20,00	
Método de Vega (1966)								
Classificação Inicial				Classificação Final				Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)		
1	561	3-17	35,09	561	3-17	35,09	100,00	
2	187	17,5-30	16,42	189	17,5-32	16,93	100,00	
3	18	32-45	11,49	16	33-45	9,94	88,89	
Método da IUFRO(1958)								
Classificação Inicial				Classificação Final				Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)		
1	148	3-7,5	20,37	86	3-6	18,29	58,10	
2	391	8-16	22,27	476	6,5-17,5	27,38	100,00	
3	227	16,5-45	26,60	204	18-45	25,86	89,90	
Método de Longhi (1980)								
Classificação Inicial				Classificação Final				Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)		
1	412	3-13	28,59	430	3-13	29,18	100,00	
2	217	13-20	12,04	238	14-21,5	12,75	91,70	
3	137	21-45	23,95	98	22-45	22,09	71,5	
Método de Souza e Leite (1993)								
Classificação Inicial				Classificação Final				Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coeficiente de Variação (%)		
1	93	3-6,5	18,64	24	3-4,5	12,72	25,80	
2	574	7-21	32,45	658	5-22	38,27	100,00	
3	99	21,5-45	22,15	84	22,5-45	21,29	84,80	

Continua...

Tabela 6: Continuação

Método de Sanquetta (1995)							
Classificação Inicial			Classificação Final			Classificação (%)	
Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)		
46	3-5	15,82	46	3-5	15,82	100	
702	5,5 -30	42,91	702	5,5 - 30	42,91	100	
18	32-45	11,49	18	32-45	11,49	100	

Método Guedes-Bruni et al. (2006)							
Classificação Inicial				Classificação Final			Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)	
1	235	3-9,5	22,54	311	3-10,5	25,31	100,00
2	178	3,5-24	25,51	119	11-13	6,29	38,80
3	353	10-45	35,82	336	14-45	30,25	80,20

Método IUFRO-LOREY							
Classificação Inicial				Classificação Final			Classificação (%)
Estrato	Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)	Número de Árvores	Amplitude (m)	Coefficiente de Variação (%)	
1	137	3-7	20,06	52	3-5,5	16,50	38,00
2	370	7,5-15	21,46	509	6-17	29,81	100,00
3	259	16-45	27,76	205	17,5-45	25,89	79,20

5. CONSIDERAÇÕES FINAIS

O método multivariado proposto neste trabalho pode ser utilizado para estratificação vertical de florestas naturais e apresenta como vantagem principal a retirada de subjetividade do pesquisador.

Foi percebido também, que as técnicas multivariadas são mais eficazes quando trabalhadas em conjunto, por isso recomenda-se a utilização sequencial das análises de componentes principais, de agrupamento e discriminante, na estratificação vertical de florestas naturais.

A aplicação de análise discriminante deve ser realizada após a classificação dos estratos a partir de um método de estratificação vertical com a finalidade de fazer um reagrupamento dos elementos e também para a formação das funções discriminantes, permitindo assim, que uma nova árvore seja alocada para o seu devido estrato.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AAKER, D. A.; KUMAR, V.; DAY, G. S. **Pesquisa de marketing**, São Paulo: Atlas, 2001. 745p.

ALBUQUERQUE, M.A. et al. Estabilidade em análise de agrupamento: estudo de caso em ciência florestal. **Revista Árvore**, v.30, n.2, p.257-265, 2006.

ANDERSON, T. W. **An introduction to multivariate statistical analysis**, New York: John Wiley & Sons, 1984, 675 p.

BENITES, V. de M. et al. Análise discriminante de solos sob diferentes usos da área de mata atlântica a partir de atributos da matéria orgânica. **Revista Árvore**, v.34, n.4, p.685-690, 2010.

BISQUERRA, R.; SARRIERA, J.; MARTÍNEZ, F. **Introdução à Estatística: Enfoque Informático com o Pacote Estatístico SPSS**. ARTMED: Porto Alegre, 2004.

BUSSAB, W. de O.; MIAZAKI, E. S; ANDRADE, D. **Introdução à análise de agrupamentos**. São Paulo: Associação Brasileira de Estatística, 1990. 105p.

CALEGÁRIO, N.; SCOLFORO, J. R. S.; SOUZA, A. L. de. Estratificação em alturas para floresta natural heterogênea: uma proposta metodológica. **Cerne**, v.1, n.1, p.58-63, 1994.

CONDEPE. **Catende**. Recife: CONDEPE, 1987. 62p. (Monografias municipais, 27).

CORAIOLA, M. **Caracterização estrutural de uma Floresta Estacional Semidecidual localizada no município de Cássia - MG**. 1997. 196f. Dissertação (Mestrado em Ciências Florestais) – Universidade Federal do Paraná, Curitiba, 1997.

CORMACK, R. A review of classification. **Journal of the Royal Statistical Society** (Series A), v.134, p.321-367, 1971.

COSTA JUNIOR, R. F. **Caracterização estrutural de um remanescente de Mata Atlântica do município de Catende-PE**. Recife: UFRPE, 2006. 52p. Dissertação (Mestrado em Ciências Florestais) – Universidade Federal Rural de Pernambuco, 2006.

CRUZ, C. D.; REGAZZI, A. J. **Métodos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 1997. 390p.

DAVIS, T.A.W.; RICHARDS, P.W. The vegetation of Moroballi Creek, British Guiana: an ecological study of a limited area of tropical rain forest. Part I. **Journal of Ecology**, v.21, n.2, p.350-384, 1933.

DUARTE, M. C.; SANTOS, J. B.; MELO, L. C. Comparison of similarity coefficients based on RAPD markers in the common bean. **Genetics and Molecular Biology**, v.22, n.3, p.427-432, 1999.

FEROZ, S. M.; YOSHIMURA, K.; HAGIHARA, A. Architectural stratification and woody species diversity of a subtropical forest grown in a limestone habitat in Okinawa Island, Japan. **World Academy of Science, Engineering and Technology**, v.23, n.1, p. 27-33, 2006a.

FEROZ, S.M.; HAGIHARA, A. Comparative Studies on Community Ecology of Two Types of Subtropical Forests Grown in Silicate and Limestone Habitats in the Northern Part of Okinawa Island, Japan. **Taiwania**, v.53, n.2, p.134-149, 2008.

FEROZ, S. M.; HAGIHARA, A.; YOKOTA, M. Stand structure and woody species diversity in relation to stand stratification in a subtropical evergreen broadleaf forest, Okinawa Island. **Journal of Plant Research**, v.119, n.4, p.293–301, 2006b.

FEROZ, S. M.; YOSHIMURA, K.; HAGIHARA, A. Stand stratification and woody species diversity of a subtropical forest in limestone habitat in the northern part of Okinawa Island. **Journal of Plant Research**, v.121, n.3, p.329–337, 2008.

FEROZ, S. M. et al. Floristic composition, woody species diversity, and spatial distribution of trees based on architectural stratification in a subtropical evergreen broadleaf forest on Ishigaki Island in the Ryukyu Archipelago, Japan. **Tropics**, v.18, n.3, p.103-114, 2009.

FERREIRA, M. L.; LIMA, O. M. B. Processo de classificação. In: FAISSOL, S. et al. (Ed). **Tendências atuais na geografia urbano regional: teorização e quantificação**. Rio de Janeiro: IBGE, 1978. p. 113-30.

FERREIRA, D. F. **Análise Multivariada**. Lavras: Universidade Federal de Lavras, 1996. 394p.

FERREIRA, R. L. C. et al. Comparação de duas metodologias multivariadas no estudo de similaridade entre fragmentos de Floresta Atlântica. **Revista Árvore**, v.32, n.3, p.511-521, 2008.

GOURLET-FLEURY, S. et al. Grouping species for predicting mixed tropical forest dynamics: looking for a strategy. **Annual Forest Science**, v.62, n.8, p.785-796, 2005.

GUEDES-BRUNI, R. R. et al. Composição florística e estrutura de trecho de Floresta Ombrófila Densa Atlântica aluvial na Reserva Biológica de Poço das Antas, Silva Jardim, Rio de Janeiro, Brasil. **Rodriguésia**, v. 57, n.3, p.413-428, 2006.

GUILHERME et al. Horizontal and vertical tree community structure in a lowland Atlantic Rain Forest, Southeastern Brazil. **Revista Brasileira de Botânica**, v.27, n.4, p. 725-737, 2004.

HOSOKAWA, R. T. **Introdução ao manejo e economia de florestas**. Curitiba: Editora UFPR, 1998. 162p.

HOZUMI, K., K. SHINOZAKI and Y. TADAKI : Studies on the frequency distribution of the weight of individual trees in a forest stand. I. A new approach toward the analysis of the distribution function and the $3/2$ th power distribution. **Japan. J. Ecol.** 18, 10-20 (1968).

HOZUMI, K. Studies on the frequency distribution of the weight of individual trees in a forest stand. V. The M-w diagram for various types of forest stands. **Japanese Journal of Ecology**, v.25, n.3, p.123-131, 1975.

JAMIR, S.A.; UPADHAYA, K.; PANDEY, H.N. Life form composition and stratification of montane humid forests in Meghalaya, northeast India. **Tropical Ecology**, v.47, n.2, p.183-190, 2006.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6 ed. Upper Saddle River: Prentice Hall, 2007. 773p.

LACHENBRUCH, P.A. **Discriminant analysis**. New York: Hatner Press, 1979. 128p.

LAMPRECHT, H. Ensayo sobre la estructura florística de la parte sur-oriental del Bosque Universitario "El Caimital" - Estado Barinas. **Revista Forestal Venezolana**, v.7, n.10-11, p.77-119, 1964.

LAMPRECHT, H. **Silvicultura nos trópicos**: ecossistemas florestais e respectivas espécies arbóreas – possibilidades e métodos de aproveitamento sustentado. Eschborn: Deutsche Gesellschaft für Technische Zusammenarbeit, 1990. 343p.

LEMOS, M.C.; PELLENS, R.; LEMOS, L.C. Perfil e florística de dois trechos de mata litorânea no município de Maricá - RJ. **Acta botânica brasileira**, v.15, n.3, p.321-324, 2001.

LIMA JÚNIOR, L.M. de et al. Utilização de técnicas multivariadas na classificação de fases de crescimentos de *Leucaena leucocephala* (Lam.) De Wit. **Floresta**, v.39, n.4, p.921-935, 2009.

LONGHI, S. J. **A estrutura de uma floresta natural de *Araucaria angustifolia* (Bert.) O. Ktze, sul do Brasil**. Curitiba. UFPR, 1980. 198 p. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal do Paraná, 1980.

MARDIA, A. K. V.; KENT. J. T.; BIBBY, J.M. **Multivariate analysis**. London: Academic Press, 1997. 518p.

MARISCAL-FLORES, E. J. **Potencial produtivo e alternativas de manejo sustentável de um fragmento de Mata Atlântica secundária, Município de Viçosa, Minas Gerais**. 1993. 165 f. Dissertação (Mestrado em Ciência Florestal) . Universidade Federal de Viçosa, Viçosa-MG, 1993.

MARRIOTT, F.H.C. **The interpretation of multivariate observation**. London: Academic Press, 1974. 117p.

ORLÓCI, L. **Multivariate analysis in vegetation research**. 2a ed. The Hague: DR W. Junk B.V., 1978. 451 p.

PIZZATO, W. **Avaliação biométrica da estrutura e da dinâmica de uma Floresta Ombrófila Mista em São João do Triunfo-PR:1995 a 1998**. Dissertação. Mestrado em Ciências Florestais. Curitiba: Universidade Federal do Paraná. 1999, 170p.

RIBEIRO JÚNIOR, J. I. **Análises Estatísticas no SAEG**. Viçosa: UFV, 2001. 301p.

RICHARDS, P. W. **The Tropical Rain Forest: an ecology study**. Cambridge: Cambridge University Press, 1952. 450 p.

RICHARDS, P. W. **The Tropical Rain Forest: an ecology study**. Cambridge: Cambridge University Press, 1996. 600p.

RIVERA HERRERA et al. Análise florística e fitossociológica do componente Arbóreo da floresta ombrófila mista presente na Reserva florestal Embrapa/Epagri, Caçador, SC – Brasil. **Floresta**, v.39, n.3, p.485-500, 2009.

ROBERTS, M. R.; GILLIAM, F. S. Patterns and mechanisms of plant diversity in forested ecosystems: implications for forest management. **Ecological Application**, v.5, n.4, p.969–977, 1995.

SANQUETTA, C.R. Análise da estrutura vertical de florestas através do diagrama h-M. **Ciência Florestal**, v.5, n.1, p.55-68, 1995.

SANTOS, J. H. et. al. Distinção de grupos ecológicos de espécies florestais por meio de técnicas multivariadas. **Revista Árvore**, v. 28, n. 3, p. 387-396, 2004.

SOUZA, A. L. de; LEITE, H. G. **Manejo florestal**. Viçosa: UFV, 1993. 147p.

SOUZA, D. R. de et. al. Emprego de análise multivariada para estratificação vertical de florestas inequidistantes. **Revista Árvore**, v.27, n.1, p.59-63, 2003.

SOUZA, D. et al. Ciclo de corte econômico ótimo em floresta ombrófila densa de terra firme sob manejo floresta sustentável, Amazônia Oriental. **Revista Árvore**, v.28, n.5, p.681-689, 2004.

SOUZA, D. et al. Análise estrutural em floresta ombrófila densa de terra firme não explorada, Amazônia Oriental. **Revista Árvore**, v.30, n.1, p.75-87, 2006.

SPSS. SPSS statistics Base 17.0. User's guide. Chicago: SPSS, 2007. 616p.

VALE et al. Composição florística e estrutura do componente arbóreo em um remanescente primário de floresta estacional semidecidual em Araguari, Minas Gerais, Brasil. **Hoehnea**, v.36, n.3, p.417-429, 2009.

VALENTIN, J. L. **Ecologia numérica**: uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Interciência, 2000. 117p.

VEGA, C. L. Observaciones ecológicas sobre los bosques de roble de la sierra Boyacá. Colombia. **Turrialba**, v.16, n. 3, p.286-296, 1966.

WARD, J. H.; Hierarchical grouping to optimize an objective function. **Journal of American Statistical Association**, v. 58, n.301, p. 236-244, 1963.

WU, M. et al. Comparative studies on vertical stratification, floristic composition, and woody species diversity of subtropical evergreen broadleaf forests between the Ryukyu Archipelago, Japan, and South China. **World Academy of Science, Engineering and Technology**, n.70, p. 934-939, 2010.