

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA E MODELAGEM
QUANTITATIVA**

**REGRESSÃO LOGÍSTICA APLICADA AO ESTUDO
DA MORTALIDADE DE ÁRVORES EM FRAGMENTO
DA FLORESTA OMBRÓFILA MISTA NO RS**

MONOGRAFIA DE ESPECIALIZAÇÃO

Luiz Thiago Brondani Greff

**Santa Maria, RS, Brasil
2012**

**REGRESSÃO LOGÍSTICA APLICADA AO ESTUDO DA
MORTALIDADE DE ÁRVORES EM FRAGMENTO DA
FLORESTA OMBRÓFILA MISTA NO RS**

por

Luiz Thiago Brondani Greff

Monografia apresentada ao Curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM,RS),
como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa.

Orientadora: Prof. Dra. Anaelena Bragança de Moraes

**Santa Maria, RS, Brasil
2012**

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Curso de Especialização em Estatística e Modelagem Quantitativa**

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**REGRESSÃO LOGÍSTICA APLICADA AO ESTUDO DA
MORTALIDADE DE ÁRVORES EM FRAGMENTO DA FLORESTA
OMBRÓFILA MISTA NO RS**

elaborada por
Luiz Thiago Brondani Greff

como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa

COMISSÃO EXAMINADORA:

Anaelena Bragança de Moraes, Dra.
(Presidente/Orientadora)

Roselaine Ruviaro Zanini, Dra. (UFSM)

Luciane Flores Jacobi, Dra. (UFSM)

Santa Maria, 28 de fevereiro de 2012.

AGRADECIMENTOS

Agradeço a minha orientadora, professora Anaelena Bragança de Moraes, por todos os ensinamentos acerca da regressão logística; pela confiança depositada em mim e pela paciência e compreensão com o ritmo no qual esse estudo foi desenvolvido.

Agradeço a comissão examinadora, composta pelas professoras Luciane Jacobi Flores e Roselaine Ruviaro Zanini, pelas sugestões e correções propostas para a monografia.

Agradeço a professora Roselaine Ruviaro Zanini, coordenadora do curso, e a secretária, Rosa Maria Righi Licht, por todo o auxílio.

Agradeço ao professor Solon Jonas Longhi por ter disponibilizado os dados utilizados neste estudo e pelas contribuições para o desenvolvimento do mesmo.

Agradeço a minha mãe, Cecília Maria Brondani pelos conselhos e o incansável incentivo para chegar até o fim dessa caminhada, assim como meu pai, Luiz Antônio Greff, minha irmã, Giulia Brondani Greff e a todos os meus familiares que sempre me apoiaram.

Agradeço a minha namorada, Michelle Dullius, por estar ao meu lado, me dando carinho e força nas horas difíceis.

Agradeço a todos os meus amigos e colegas que de alguma maneira contribuíram para a concretização dessa etapa.

RESUMO

Monografia de Especialização
Curso de Especialização em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

REGRESSÃO LOGÍSTICA APLICADA AO ESTUDO DA MORTALIDADE DE ÁRVORES EM FRAGMENTO DA FLORESTA OMBRÓFILA MISTA NO RS

AUTOR: LUIZ THIAGO BRONDANI GREFF
ORIENTADORA: ANAELENA BRAGANÇA DE MORAES
Data e Local da Defesa: Santa Maria, 28 de fevereiro de 2012.

Este estudo foi realizado em fragmento da Floresta Ombrófila Mista, no estado do Rio Grande do Sul, tendo como objetivos construir um modelo logístico para a mortalidade de árvores, identificar os fatores associados e avaliar a qualidade do ajuste desse modelo. Para isso, fez-se uso de dois conglomerados que fazem parte da base de dados do Programa Brasileiro de Pesquisas Ecológicas de Longa Duração, referente ao ano de 2006. Foram testadas variáveis morfométricas: diâmetro a altura do peito, altura total, incremento em diâmetro, grau de esbeltez, classe de copa; variáveis de sítio: diâmetro máximo, altura média, área basal, densidade de árvores vivas; e uma variável de competição: *Basal area larger*. Por meio das análises de regressão logística univariada e múltipla identificaram-se as variáveis que influenciam a mortalidade e a magnitude dessa influência. Na modelagem foram testados três métodos de seleção de variáveis: *enter*, *forward* e *backward*. Para avaliar o modelo logístico utilizou-se o teste da razão de verossimilhança, R^2 de Nagelkerke, Teste de Hosmer e Lemeshow, entre outros. Foram utilizadas 1672 árvores, sendo que 4% dessas eram árvores mortas. Como resultado principal destaca-se que os três métodos de seleção de variáveis modelaram um modelo final idêntico, o qual foi significativo pelo teste da razão de verossimilhança (p -valor $< 0,001$), com R^2 de Nagelkerke de 0,144, e não significativo pelo Teste de Hosmer e Lemeshow (p -valor = 0,235), no entanto o modelo é de pequena capacidade preditiva. As variáveis associadas à mortalidade das árvores foram o incremento em diâmetro, a altura total e a densidade de árvores vivas, todas exercendo influência negativa na mortalidade das árvores.

Palavras-chave: modelagem da mortalidade, espécies arbóreas, modelo logístico, Floresta com Araucária.

ABSTRACT

Monograph of Specialization
Specialization Course in Statistics and Quantitative Modeling
Federal University of Santa Maria

LOGISTIC REGRESSION APPLIED ON THE STUDY OF MORTALITY OF TREES ON FRAGMENT OF THE MIXED RAIN FOREST ON THE RS

AUTHOR: LUIZ THIAGO BRONDANI GREFF
GUIDANCE: ANAELENA BRAGANÇA DE MORAES
Date and Location of Defense: Santa Maria, February 28, 2012

This study was conducted in a fragment of Araucaria Forest in the state of Rio Grande do Sul, aiming to build a logistic model for mortality of trees, identify the associated factors and assess the quality of the fit of this model. For this, use has been made of two groups that are part of the database of the Brazilian Program for Long-Term Ecological Research for the year 2006. Morphometric variables were tested: diameter at breast height, total height, diameter increment, degree of slimness, canopy class, site variables: maximum diameter, height, basal area, density of live trees, and a variable contest: basal area larger. Through the analysis of univariate and multiple logistic regression identified the variables that influence mortality and the magnitude of this influence. In the model tested three methods for selecting variables enter, forward and backward. To evaluate the logistic model used the likelihood ratio test, Nagelkerke R^2 , Hosmer and Lemeshow test, among others. 1672 trees were used, and 4% of these trees were dead. As a main result it is emphasized that the three methods of variable selection shaped a final model is identical, which was significant by the test of likelihood ratio (p-value <0.001), with R^2 of Nagelkerke of 0.144, not significant for the Test Hosmer and Lemeshow (p-value = 0.235), however the model is of little predictive ability. The variables associated with tree mortality were increasing in diameter, total height and density of living trees, all exerting a negative influence on tree mortality.

Keywords: modeling of mortality, tree species, the logistic model, Araucaria Forest.

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1 – Distribuição geográfica da Floresta com Araucária nas regiões Sul e Sudeste do Brasil, 1992..... | 16 |
|---|----|

LISTA DE QUADROS

| | |
|--|----|
| Quadro 1 - Categorias de posição sociológica | 24 |
| Quadro 2 - Classes de integridade da copa | 25 |
| Quadro 3 – Categorias de classe de copa | 25 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 – Frequência da variável dependente | 29 |
| Tabela 2 – Estatísticas descritivas das variáveis independentes quantitativas | 30 |
| Tabela 3 – Distribuição de frequência das variáveis independentes qualitativas | 32 |
| Tabela 4 – Estatísticas da regressão logística univariada..... | 33 |
| Tabela 5 – Estatística da multicolinearidade para as covariáveis significativas na análise de regressão logística univariada..... | 34 |
| Tabela 6 – Modelos encontrados pelos três métodos de seleção..... | 35 |
| Tabela 7 – Medidas de avaliação da qualidade do ajuste do modelo de RLM..... | 37 |
| Tabela 8 – Tabela de classificação. | 38 |

LISTA DE ABREVIATURAS E SIGLAS

ADAP – Amplitude do diâmetro
B – coeficiente estimado
BAL – Índice *basal area larger*
CAP – Circunferência à altura do peito (1,3 metros do solo)
CC – Classe de copa
CV – Coeficiente de variação
DAP – Diâmetro à altura do peito
DAP⁻¹ – Inverso do diâmetro
DAP² – Diâmetro ao quadrado
DAP_{máx} – Diâmetro máximo
DAP_{mín} – Diâmetro mínimo
DAP_{1cat} – Inverso do diâmetro categórico
dens. – Densidade de árvores vivas
E.P. – Erro padrão
FIV – Fator de Inflação da Variância
FLONA – Floresta Nacional
gabs – Área basal absoluta
GE – Grau de esbeltez
h_{copa} – Altura da copa
h_{méd} – Altura média
ht – Altura total
IDAP – Incremento em diâmetro
IC – Integridade da copa
IC-95% – Intervalo de confiança de 95%
OR – *Odds ratio* ou razão de chances
PELD – Pesquisas Ecológicas de Longa Duração
PS – Posição sociológica
Q1 – primeiro quartil
Q3 – terceiro quartil
R² – Coeficiente de determinação
RL – Regressão Logística
RLM – Regressão Logística Múltipla
RLU – Regressão Logística Univariada
RS – Rio Grande do Sul
s – Desvio padrão
SFP – São Francisco de Paula

LISTA DE APÊNDICES

| | |
|---|----|
| Apêndice A– Categorias de inverso do diâmetro e estatísticas da regressão logística univariada..... | 45 |
| Apêndice B – Pertinência prevista ao grupo para as 68 árvores mortas | 46 |

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO..... | 9 |
| 1.1 Objetivos..... | 10 |
| 1.1.1 Objetivo geral | 10 |
| 1.1.2 Objetivos específicos..... | 10 |
| 2 REVISÃO BIBLIOGRÁFICA..... | 11 |
| 2.1 Regressão Logística | 11 |
| 2.1.1 Modelo matemático da regressão logística | 12 |
| 2.1.2 Métodos de regressão logística | 13 |
| 2.1.3 Seleção de variáveis independentes e avaliação do modelo logístico | 13 |
| 2.2 Floresta Ombrófila Mista | 15 |
| 2.3 Mortalidade de árvores | 17 |
| 3 MATERIAIS E MÉTODOS..... | 19 |
| 3.1 Caracterização da área de estudo | 19 |
| 3.2 Base de dados do estudo..... | 19 |
| 3.2.1 Desenho amostral do PELD | 20 |
| 3.2.2 Levantamento da vegetação | 20 |
| 3.3 Estruturação do banco de dados..... | 21 |
| 3.3.1 Variável dependente..... | 22 |
| 3.3.2 Variáveis Independentes | 22 |
| 3.4 Tratamento estatístico | 26 |
| 3.5 Modelagem da mortalidade por meio da regressão logística | 26 |
| 4 RESULTADOS E DISCUSSÕES..... | 29 |
| 4.1 Análise de regressão logística univariada..... | 33 |
| 4.2 Análise de regressão logística múltipla | 35 |
| 5 CONCLUSÃO..... | 40 |
| REFERÊNCIAS..... | 41 |

1 INTRODUÇÃO

Em relatório divulgado pela Organização das Nações Unidas para Agricultura e Alimentação (2011) sobre o estado das florestas do mundo, a região da América Latina e o Caribe se destacam pela abundância em recursos florestais, sendo que cerca de 49% dessa região estava coberta por florestas no ano de 2010. No entanto, a sub-região da América do Sul mantém uma taxa de redução anual da área florestal de 0,45% causada pelo desmatamento e pela conversão de áreas florestais em áreas agrícolas ou urbanas.

Especificamente, o Brasil possui aproximadamente 61% do seu território ocupado por florestas, distribuídas nos biomas Amazônia, Mata Atlântica, Caatinga, Cerrado, Pantanal e Pampa. Destaca-se que a Mata Atlântica e o Pampa juntos respondem por aproximadamente 15% da área total dos biomas brasileiros, e são os principais biomas da região Sul do Brasil (SERVIÇO FLORESTAL BRASILEIRO, 2010).

A Mata Atlântica encontra-se em estado de intensa fragmentação, resultado da atual devastação que atinge cerca de 93% da sua formação original (FUNDAÇÃO SOS MATA ATLÂNTICA, 2011). Ainda assim, abriga em torno de 20 mil espécies de plantas, das quais em torno de 40% são endêmicas. Além da fragmentação, existe a lacuna de conhecimentos sobre essas regiões. A cada ano são descritas inúmeras novas espécies, porém ainda existem áreas pouco conhecidas, e os recursos humanos e investimentos para as pesquisas são limitados (MINISTÉRIO DO MEIO AMBIENTE/SECRETARIA DE BIODIVERSIDADE E FLORESTAS, 2000). Com essa situação de ameaça da flora, grande esforço da comunidade científica é direcionado para estes estudos, à medida que existam espécies desconhecidas e ainda diversas espécies de plantas em risco de extinção. Além da composição das espécies, existem outras lacunas no conhecimento a cerca das florestas nativas, principalmente nos processos que nela ocorrem.

Vários estudos realizados no estado do Rio Grande do Sul objetivaram modelar o crescimento das árvores (DELLA FLORA; DURLO; SPATHELF, 2004; ZANON, 2007; HESS, 2006; MATTOS, 2007; CHASSOT, 2009), no entanto, estudos sobre a mortalidade das árvores são escassos.

Segundo Rossi et al. (2007), a mortalidade é um componente importante nas estimativas de crescimento e produção florestal, no entanto é um processo pouco conhecido devido a sua complexidade. Hamilton (1986 apud ROSSI et al., 2007) salienta que poucos estudos foram desenvolvidos especificamente para aumentar o entendimento do processo de mortalidade, e por conta dessa carência de dados a mortalidade é tratada como uma suposição ou julgamento subjetivo.

As grandes limitações para a obtenção de dados para a modelagem da mortalidade das árvores estão na baixa frequência com que o evento ocorre e na necessidade de se realizar medições sucessivas para obter a informação de quando a árvore morreu, sendo que para supri-las há grande demanda de recursos materiais e humanos.

Nesse contexto, o presente estudo visa contribuir no entendimento dos fatores que levam à mortalidade de árvores e suas possíveis causas.

1.1 Objetivos

1.1.1 Objetivo geral

Construir um modelo de mortalidade para as espécies arbóreas, identificando fatores de risco associados, em fragmento da Floresta Ombrófila Mista, na Floresta Nacional de São Francisco de Paula, RS.

1.1.2 Objetivos específicos

- Identificar variáveis relacionadas com a mortalidade de espécies arbóreas utilizando o modelo de regressão logística;
- Avaliar a qualidade do ajuste de diferentes modelos de mortalidade.

2 REVISÃO BIBLIOGRÁFICA

2.1 Regressão Logística

Os métodos de regressão têm se tornado um componente integral de qualquer análise de dados relacionados com a descrição da relação entre uma variável resposta (dependente) e uma ou mais variáveis explanatórias (independentes) (HOSMER; LEMESHOW, 1989).

Na regressão linear a variável dependente é de natureza contínua, enquanto na regressão logística a variável dependente é qualitativa, admitindo duas ou mais categorias, necessitando assim, de métodos específicos de estimação dos coeficientes (FIGUEIRA, 2006). Para a regressão linear a estimação dos coeficientes é feita pelo método dos mínimos quadrados e para a regressão logística o método adequado é o da máxima verossimilhança (DIAS FILHO; CORRAR, 2007).

Segundo Figueira (2006), na regressão logística as categorias (ou valores) da variável dependente podem assumir valores de natureza ordinal (quando existe determinada ordem nas categorias) ou nominal (quando não existe ordem entre as categorias), sendo então, denominada de Regressão Logística Ordinal ou Regressão Logística Nominal, respectivamente.

Neste estudo aborda-se a Regressão Logística Nominal, sendo a variável dependente dicotômica ou binária. Esse tipo de variável dependente possui apenas duas categorias, com valores codificados como 0 ou 1; enquanto que as variáveis independentes, também chamadas de covariáveis, tanto podem ser categóricas ou contínuas.

Na Regressão Logística (RL) pode-se relacionar uma única covariável à variável dependente ou várias covariáveis, sendo denominadas de regressão logística univariada (RLU) e regressão logística múltipla (RLM), respectivamente (HOSMER; LEMESHOW, 1989).

Com a RL pode-se identificar variáveis independentes que influenciam no resultado (variável dependente) e a magnitude dessa influência (HOSMER; LEMESHOW, 1989). Essa técnica tem por objetivos classificar fenômenos ou

indivíduos em categorias específicas e ainda estimar a probabilidade de ocorrência de determinado evento (DIAS FILHO; CORRAR, 2007).

2.1.1 Modelo matemático da regressão logística

As estimativas da variável dependente geradas com o uso do modelo linear podem prever valores menores que zero e maiores que um. Para adequar as estimativas a esse intervalo de valores deve-se fazer alguns procedimentos matemáticos no modelo de regressão que estão detalhadamente descritos em Dias Filho e Corrar (2007) e assim obtém-se o modelo geral da regressão logística que é apresentado pela expressão:

$$\text{Probabilidade (evento)} = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + \dots + b_n X_n)}} \quad (1)$$

Onde: b_0, b_1, \dots, b_n são os coeficientes estimados das variáveis independentes; X_1, \dots, X_n são as variáveis independentes, e e é a base do logaritmo natural.

Os coeficientes são medidas das variações na proporção das probabilidades, chamadas de razão de desigualdade, ou razão de chance. São expressos em logaritmos, precisando ser transformados novamente (antilogaritmo) de forma que seu efeito relativo sobre as probabilidades seja avaliado mais facilmente. Um coeficiente positivo aumenta a probabilidade, ao passo que um valor negativo diminui a probabilidade prevista (HAIR et al., 2005).

Como exemplo, supomos que o coeficiente b_1 seja de 0,3, então a razão de chance será impactada em $e^{0,3} = 1,35$. Isso significa que cada unidade de variação que se registre na variável independente as chances da variável dependente aumentarão em 35% em relação à situação anterior, mantendo as demais variáveis constantes (DIAS FILHO; CORRAR, 2007).

Para a estimação de parâmetros β , o método dos mínimos quadrados não é recomendado para modelos com resultado dicotômico, pois os estimadores não apresentam propriedades estatísticas desejáveis para o ajuste do modelo. Nesse

caso, utiliza-se o método de máxima verossimilhança, que produz estimativas para os parâmetros desconhecidos que maximizam a probabilidade de obtenção do conjunto de dados observados (HOSMER; LEMESHOW, 1989).

2.1.2 Métodos de regressão logística

Segundo Field (2009), os principais métodos para inserir variáveis na Regressão Logística são o método da entrada forçada (*enter*) e os métodos passo a passo (*stepwise*). No método *enter* todas as covariáveis são inseridas no mesmo modelo de regressão em um único bloco, enquanto no *stepwise* pode-se iniciar com todas as covariáveis e removê-las uma a uma (*backward*) segundo um critério estatístico de significância dos coeficientes, ou iniciar apenas com a constante e ir incluindo variáveis uma a uma (*forward*) de acordo com sua contribuição para o modelo.

2.1.3 Seleção de variáveis independentes e avaliação do modelo logístico

Um importante critério para a seleção de covariáveis é o teste Wald, cuja finalidade é aferir o grau de significância de cada coeficiente da equação logística (DIAS FILHO; CORRAR, 2007). Esse teste é obtido pela razão entre o coeficiente estimado de uma covariável e seu erro padrão estimado. O valor resultante segue uma distribuição normal padrão, para o qual a hipótese da nulidade é testada, com um nível de significância estabelecido (HOSMER; LEMESHOW, 1989). A estatística do teste é dada pela equação:

$$W = \frac{\hat{\beta}_i}{EP(\hat{\beta}_i)} \quad (2)$$

Onde: $\hat{\beta}_i$ = estimativa do coeficiente de regressão; EP ($\hat{\beta}_i$) = erro padrão do coeficiente de regressão β_i .

Sendo as hipóteses: $H_0: \beta_i = 0$ e $H_1: \beta_i \neq 0$, e comparando-se a estatística W com um valor tabelado $Z_{\alpha/2}$ (bilateral) para o nível de significância (α) considerado, pode-se concluir pela aceitação ou rejeição de H_0 (HOSMER; LEMESHOW, 1989).

Outra medida para avaliação da significância dos coeficientes das covariáveis é o teste da razão de verossimilhança (G) que é usado para testar a hipótese de que todos os coeficientes do modelo logístico são iguais a zero, e para isso, é testada a diferença entre o valor do modelo, apenas com a constante, em relação ao modelo em que são incluídas todas as variáveis independentes, pelo teste do Qui-quadrado (DIAS FILHO; CORRAR, 2007). Segundo Hosmer e Lemeshow (1989) a estatística G pode ser obtida pela equação:

$$G = -2\ln(L_s) + 2\ln(L_c) \quad (3)$$

Em que: L_s é a verossimilhança do modelo sem a covariável, ou seja, apenas com o intercepto e L_c é a verossimilhança do modelo com a covariável.

Para a avaliação do modelo de regressão logística são utilizadas medidas do tipo R^2 e o Teste de Hosmer e Lemeshow. Segundo Dias Filho e Corrar (2007), no modelo logístico são utilizados indicadores semelhantes ao R^2 da regressão linear, e que são denominados de pseudo R^2 e podem ser utilizados para comparar o desempenho de modelos concorrentes. Ainda segundo os autores, o R^2 de Nagelkerke (R^2N) situa-se entre zero e um, e na comparação entre dois modelos logísticos deve-se preferir um modelo que possua o valor mais elevado dessa medida. Segundo Field (2009) a estatística R^2N é obtida pela equação:

$$R^2N = \frac{1 - e^{\left[-\frac{2}{n}(L_c - L_s)\right]}}{1 - e^{\left[\frac{2L_s}{n}\right]}} \quad (4)$$

Em que: $R^2_N = R^2$ de Nagelkerke; n = tamanho da amostra; L_s = verossimilhança do modelo com o intercepto; L_c = verossimilhança do modelo com a covariável.

O Teste de Hosmer e Lemeshow tem por objetivo verificar através de um teste do Qui-quadrado, a certo nível de significância, a hipótese de que não existem diferenças significativas entre as classificações previstas pelo modelo e as observadas; para isso, o teste consiste em dividir o número de observações em dez classes e comparar as frequências previstas com as observadas, caso haja diferença significativa entre as classificações o modelo é considerado pouco confiável (DIAS FILHO; CORRAR, 2007).

2.2 Floresta Ombrófila Mista

Dentre as fitofisionomias da Mata Atlântica na região Sul, destaca-se uma formação florestal denominada Floresta Ombrófila Mista ou Floresta com Araucária, cuja espécie predominante é a *Araucaria angustifolia*, ou o pinheiro brasileiro (MEDEIROS; SAVI; BRITO, 2005).

A denominação Floresta Ombrófila Mista decorre, segundo Leite e Klein (1990), da ocorrência da mistura de floras de diferentes origens, definindo padrões fitofisionômicos típicos, em zona climática caracteristicamente pluvial.

A área original de ocorrência da Floresta Ombrófila Mista era superior a 200.000 km², sendo uma fitofisionomia característica da região Sul do país (Figura 1), ocorrendo principalmente nos Estados do Paraná (40%), Santa Catarina (31%), Rio Grande do Sul (25%), apresentando manchas esparsas no sul de São Paulo (3%), sul de Minas Gerais e Rio de Janeiro (1%) (MEDEIROS; SAVI; BRITO, 2005).

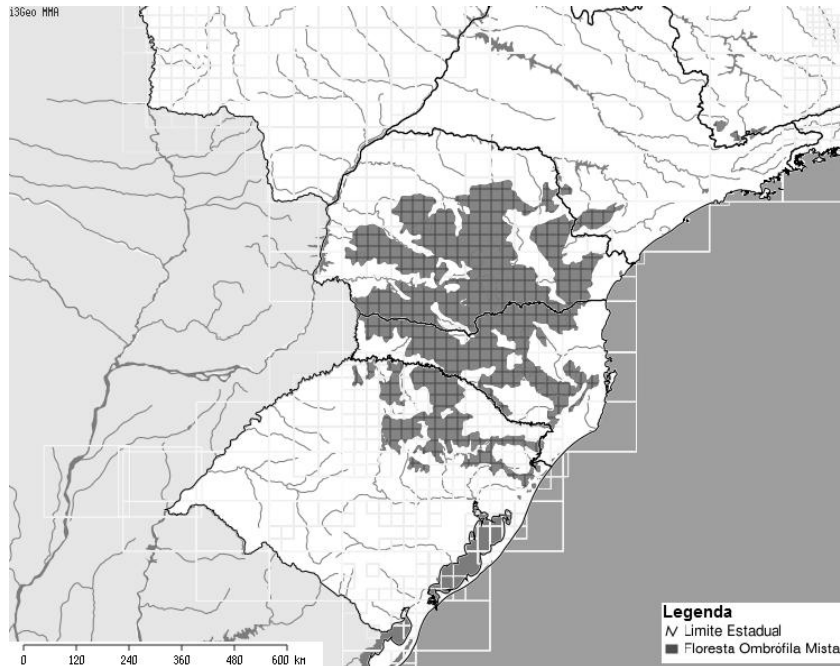


Figura 1 – Distribuição geográfica da Floresta com Araucária nas regiões Sul e Sudeste do Brasil, 1992.

Fonte: Adaptado do site <http://mapas.mma.gov.br/mostratema.php?temas=vegetacao>

Para Ivanauskas e Assis (2009), a Floresta Ombrófila Mista apresenta estrutura bem definida e é, via de regra, bem estratificada. O estrato emergente é formado exclusivamente pela araucária, que atinge em média 30 m de altura, podendo chegar até 40 m. O dossel apresenta cerca de 20 m e é formado predominantemente por espécies latifoliadas de Myrtaceae e Lauraceae e pelo pinheiro-bravo (*Podocarpus lambertii*).

A Floresta Ombrófila Mista, devido a qualidade e elevado valor econômico da madeira da araucária, foi alvo de intensa exploração madeireira, reduzindo-se drasticamente a área de ocorrência original de toda a formação florestal. Além disso, os desmatamentos e queimadas, a substituição da vegetação por pastagens, a agricultura, os reflorestamentos homogêneos com espécies exóticas e a ampliação das zonas urbanas no sul do Brasil, iniciados nos primeiros anos do século XX, provocaram uma dramática redução da área das florestas originais na região (MEDEIROS; SAVI; BRITO, 2005).

2.3 Mortalidade de árvores

As causas da mortalidade em árvores podem ser divididas em três grupos principais: catastrófica, antropogênica e regular. A catastrófica se refere a mortalidade em larga escala, causada por tempestades, insetos-praga, ou outros eventos extraordinários; a antropogênica ocorre devido a operações de colheita florestal, poluição industrial e outras alterações causadas pelas atividades humanas; e a regular é a que ocorre devido a idade das árvores e a competição entre elas, além de insetos, doenças e condições climáticas desfavoráveis (VANCLAY, 1994).

A mortalidade que ocorre por causas catastróficas ou antropogênicas é difícil de ser predita e modelada, pois requer longo tempo de mensuração de parcelas permanentes, por isso, a maioria dos estudos de mortalidade em árvores modela a mortalidade regular (SIMS et al. 2009).

Segundo Yang, Titus, Huang (2003) a função logística começou a ser utilizada na modelagem da mortalidade de árvores, no ano de 1967, e desde então, se tornou a principal função para a modelagem desse evento. Além da mortalidade, a sobrevivência da árvore pode ser o evento de interesse no modelo de regressão logística, no qual são utilizados praticamente os mesmos conjuntos de variáveis independentes para os dois eventos.

Existe um grande número de variáveis candidatas para a modelagem da mortalidade, e que podem ser classificados em quatro grupos de variáveis: medidas individuais do porte da árvore, medidas individuais da competição entre as árvores, medidas individuais da vitalidade das árvores e medidas da densidade do povoamento (HAMILTON, 1986 apud EID, TUHUS, 2001).

As principais variáveis explicativas de modelos logísticos, cuja variável dependente é o status da árvore (viva ou morta), são: diâmetro a altura do peito (DAP), transformação hiperbólica do diâmetro (DAP^{-1}), transformação quadrática do diâmetro (DAP^2), proporção de copa, área basal das árvores mais grossas, incremento em diâmetro (IDAP), índice de sitio, área basal, altura da árvore e idade da árvore (YAO; TITUS; MACDONALD, 2001; SIMS et al., 2009; YANG; TITUS; HUANG, 2003; EID; TUHUS, 2001; MONSERUD; STERBA, 1999; SCHNEIDER et al., 2005).

A grande diferença nos modelos de mortalidade de árvores está no número de espécies que compõem cada floresta. Em um estudo da mortalidade de árvores na Áustria, Monserud e Sterba (1999) modelaram a mortalidade das seis principais espécies que ocorrem nas florestas desse país. Na Noruega, Eid e Tuhus (2001) desenvolveram modelos de mortalidade para as três principais espécies de árvores do país e ainda para um grupo com as demais espécies. Em Alberta, no Canadá foram construídos modelos de mortalidade para três espécies de árvores (YAO; TITUS; MACDONALD, 2001) e também para uma única espécie (YANG; TITUS; HUANG, 2003).

No caso das florestas tropicais, o número de espécies é muito superior ao das florestas temperadas do hemisfério norte onde foram desenvolvidos os estudos supracitados. O que de certa forma dificulta a modelagem da mortalidade, como no estudo desenvolvido por Vanclay (1990) no norte de Queensland, na Austrália, para 100 espécies de árvores, em que o autor necessitou agrupar as espécies para a construção de 10 modelos de mortalidade, e mesmo assim, não obteve ótimos modelos.

3 MATERIAIS E MÉTODOS

3.1 Caracterização da área de estudo

A Floresta Nacional de São Francisco de Paula (FLONA – SFP), administrada pelo Instituto Chico Mendes de Conservação da Biodiversidade, constitui-se em uma Unidade de Conservação de Uso Sustentável. Sua área total é de 1.606 hectares distribuídos principalmente em floresta nativa (901 ha), reflorestamentos (605 ha) e estradas, barragens, campos, construções, entre outros (100 ha). A FLONA - SFP localiza-se no município de São Francisco de Paula, na região Nordeste do Estado do Rio Grande do Sul (BRASIL, 2000).

O clima da região, de acordo a classificação proposta por Rossato (2011) é o Subtropical muito úmido com inverno frio e verão fresco, que é caracterizado por chuvas abundantes que oscilam entre 1700 a 2000 mm anuais bem distribuídas, e que mensalmente ocorrem de 12 a 15 dias com precipitação. A temperatura média anual varia de 14 a 17 °C, a temperatura média do mês mais frio oscila entre 8 e 14 °C, e a temperatura média do mês mais quente varia entre 17 e 23 °C, sendo a região Nordeste do RS a mais fria e chuvosa do estado (ROSSATO, 2011).

A cobertura geológica dessa região é de rochas vulcânicas ácidas e o relevo varia de suave ondulado a forte ondulado. Os solos dominantes correspondem às classes de Cambissolos Húmicos, Cambissolos Háplicos, Neossolos Regolíticos e Neossolos Litólicos (STRECK et al., 2008).

3.2 Base de dados do estudo

Para a realização do presente estudo foi utilizado parte do banco de dados fornecido pelo Programa Brasileiro de Pesquisas Ecológicas de Longa Duração (PELD), intitulado: “Conservação e Manejo Sustentável de Ecossistemas Florestais – Bioma Floresta de Araucária e suas Transições”, que monitorou a floresta natural da FLONA - SFP por um período de 10 anos.

3.2.1 Desenho amostral do PELD

O universo amostral constitui-se da comunidade arbórea que compõe o fragmento florestal delimitado pela área da FLONA-SFP, o qual pertence à formação florestal denominada Floresta Ombrófila Mista.

A amostragem adotada foi em múltiplas ocasiões, em que foram realizadas sucessivas abordagens da população por um período de dez anos (2000 a 2009).

O método de amostragem utilizado foi o método da área fixa, enquanto que o processo de amostragem utilizado para selecionar as unidades amostrais foi misto, realizado em conglomerados e com repetição total. Os conglomerados foram caracterizados pela seleção aleatória das unidades primárias (primeiro estágio) e dentro dessas, a seleção sistemática das unidades secundárias (segundo estágio). A repetição total foi definida por sucessivas remedições de todas as unidades amostrais fixadas em um primeiro momento.

A amostra do PELD foi constituída de dez conglomerados, e a unidade amostral alocada foi uma parcela permanente da forma quadrada, com 100 m de lado (10.000 m² de superfície), a qual foi dividida em 10 faixas de 10 m de largura por 100 m de comprimento, e estas subdivididas em 10 subunidades de 10 m x 10 m (100 m²), totalizando 100 subunidades.

3.2.2 Levantamento da vegetação

Para o levantamento do componente arbóreo da vegetação, composto por árvores vivas, mortas em pé e cipós, foi utilizado como critério de inclusão a Circunferência à Altura do Peito (CAP) maior ou igual a 30 cm, e para cada indivíduo selecionado realizou-se o seguinte procedimento:

a) medição do CAP: foi medido com fita métrica retrátil de precisão em milímetros. As árvores com dois ou mais fustes (troncos) abaixo da altura do peito (1,30 m do solo) e que atingiram o critério de inclusão tiveram cada fuste considerado como uma árvore;

- b) medição da altura total e comercial: mediram-se a altura total e comercial das árvores com Hipsômetro Digital *Vertex Forestor*, com precisão de decímetros;
- c) numeração da árvore: cada indivíduo foi numerado em ordem sequencial de medição das árvores;
- d) identificação da espécie: foi registrado o nome científico da árvore quando identificada a campo. Em caso de dúvidas, foi coletado material botânico da espécie, o qual recebeu uma etiqueta com um código para posterior identificação no Herbário do Departamento de Ciências Florestais da Universidade Federal de Santa Maria;
- e) classificação da posição sociológica: as árvores foram classificadas de acordo com a posição que ocupam no estrato vertical com relação às suas vizinhas, em quatro classes: 1. Árvore dominante; 2. Árvore co-dominante; 3. Árvore dominada; e 4. Árvore suprimida;
- f) classificação da copa: as copas foram classificadas de acordo com a sua altura em: 1 = Copa longa, altura da copa maior que metade da altura total da árvore; 2 = Copa média, altura da copa entre 25 e 50% da altura total da árvore; 3 = Copa curta, altura da copa inferior a 25% da altura total da árvore; e 4 = Copa danificada ou quebrada.

O levantamento da vegetação iniciou no ano de 2000 e nos anos seguintes de remedições (2001 a 2009) foram registradas as árvores que alcançaram o critério mínimo de inclusão, sendo chamadas de árvore ingresso, e também as árvores que morreram durante o período de remedições, recebendo um código e o registro do ano de morte.

3.3 Estruturação do banco de dados

No presente estudo foram usados os dados referentes a dois conglomerados, selecionados pelo maior número de árvores mortas que os demais, no ano de 2006, o qual foi o ano de maior mortalidade.

O banco de dados original constituía-se em uma planilha eletrônica, na qual estavam dispostas as variáveis mensuradas ocupando as colunas e as árvores medidas sequencialmente por subunidade nas linhas. As variáveis utilizadas no presente estudo foram: CAP (cm), altura comercial (m), altura total (m), posição

sociológica, classe de copa e árvores mortas a cada ano. Com essas variáveis elaborou-se um novo banco de dados, utilizando as informações disponíveis para criar novas variáveis, que estão descritas a seguir.

3.3.1 Variável dependente

Com a informação da morte da árvore foi criada uma variável chamada de “estado”, que representa a condição da árvore quanto a sua sobrevivência, ou seja, se ela estava viva ou morta. Essa variável assume apenas dois valores, o que caracteriza um evento chamado de dicotômico (ou binário), definidos em: 0 para árvore viva, ou 1 para árvore morta; constituindo essa a variável dependente (Y) no presente estudo.

3.3.2 Variáveis Independentes

Foram criadas variáveis que representam cada indivíduo, o local onde ele se encontra, e ainda a competição entre os indivíduos. Essas variáveis foram denominadas de características dendrométricas, de sítio, e de competição, respectivamente, sendo descritas a seguir.

As características dendrométricas foram construídas através de medidas individuais das árvores, que originaram as seguintes variáveis: diâmetro à altura do peito (DAP), inverso do diâmetro (DAP^{-1}), diâmetro ao quadrado (DAP^2), altura da copa (hcopa), grau de esbeltez (GE), incremento em diâmetro (IDAP), altura total (ht), classe de copa (CC) e posição sociológica (PS).

As características do sítio foram criadas para representar a menor parte da unidade amostral, ou seja, cada uma das 200 subunidades que compõem o conjunto de dados do presente estudo. Essas características geraram as seguintes variáveis: diâmetro máximo (DAPmáx), diâmetro mínimo (DAPmín), amplitude do diâmetro (ADAP), área basal absoluta (gabs); densidade absoluta (dens.) e altura média (h méd).

A variável que representa a competição entre as árvores foi o índice *basal area larger* (BAL).

Abaixo estão descritas as variáveis construídas a partir das medidas existentes.

- Diâmetro à altura do peito (DAP): representa o tamanho ou a dimensão da árvore e sua unidade de medida é o centímetro. O DAP foi obtido pela medida da circunferência à altura do peito (CAP), e a partir dele foram criadas as variáveis diâmetro ao quadrado (DAP²) e inverso do diâmetro (DAP⁻¹). O diâmetro foi calculado pela seguinte fórmula:

$$DAP = \frac{CAP}{\pi} \quad (5)$$

Onde: DAP = diâmetro à altura do peito em centímetros; CAP = circunferência à altura do peito em centímetros; π = PI = 3,1415;

- Altura de copa (hcopa): obtida pela diferença entre a altura total da árvore e a altura comercial, serve como indicativo do tamanho da copa. Sua unidade de medida é o metro, sendo obtida pela expressão:

$$hcopa = ht - hc \quad (6)$$

Onde: hcopa = altura de copa em metros; ht = altura total da árvore em metros; hc = altura comercial em metros;

- Grau de esbeltez (GE): obtida pelo quociente entre a altura e diâmetro, representa o grau de instabilidade da árvore. O valor de GE inferior à unidade, indica que a árvore é considerada estável, enquanto que o valor superior indica o risco de instabilidade da árvore. Foi obtido pela expressão:

$$GE = \frac{ht}{DAP} \quad (7)$$

Onde: GE = grau de esbeltez; ht = altura total; DAP = diâmetro à altura do peito;

- Incremento em diâmetro (IDAP): representa o crescimento da árvore em diâmetro ocorrido durante duas medições, que no caso significa um ano. Sua unidade de medida é o centímetro, e o arredondamento dos valores dessa variável foi de uma casa decimal. Foi obtido pela expressão:

$$IDAP_i = DAP_{i2006} - DAP_{i2005} \quad (8)$$

Onde: $IDAP_i$ = incremento em diâmetro da árvore i em centímetros; DAP_{i2006} = diâmetro da árvore i em centímetros, no ano de 2006; DAP_{i2005} = diâmetro da árvore i em centímetros, no ano de 2005;

- Posição sociológica (PS): indica a qual estrato vertical a árvore pertence. As categorias dominada e suprimida foram agrupadas em uma única categoria, devido a baixa frequência de indivíduos ($n < 20$) na classe suprimida. Optou-se por usar as três categorias da variável em duas variáveis *dummy*, denominadas de posição sociológica 1 (PS1) e posição sociológica 2 (PS2) e apresentadas no quadro abaixo.

| Categoria | <i>dummy 1</i> | <i>dummy 2</i> |
|--------------------|-----------------------|-----------------------|
| Dominante | 0 | 0 |
| Codominante | 1 | 0 |
| Dominada/suprimida | 0 | 1 |

Quadro 1 - Categorias de posição sociológica

- Classe de copa: atribui uma classificação visual para o tamanho da copa de cada árvore. Houve a necessidade de criar-se uma nova variável com a classe danificada, uma vez que a copa danificada poderia ter qualquer tamanho de copa. Assim, foi criada a variável chamada integridade da copa (IC), a qual se usou *dummy*, apresentada no quadro 2.

| Classe | <i>dummy</i> |
|------------|--------------|
| Integra | 0 |
| Danificada | 1 |

Quadro 2 - Classes de integridade da copa

As demais categorias de classe de copa foram usadas em duas variáveis *dummy*, chamadas de classe de copa 1 (CC1) e classe de copa 2 (CC2), apresentadas no quadro 3.

| Categoria | <i>dummy 1</i> | <i>dummy 2</i> |
|-----------|----------------|----------------|
| Longa | 0 | 0 |
| Média | 1 | 0 |
| Curta | 0 | 1 |

Quadro 3 – Categorias de classe de copa

- Área basal absoluta (g_{abs}): representa quanto da superfície da subunidade está ocupada pelas árvores. Obtida pelo somatório da área basal de todos os indivíduos da subunidade, sua unidade de medida é o metro quadrado, e é expressa pela equação:

$$g_{abs_i} = \sum g \quad (9)$$

Onde: g_{abs_i} = área basal absoluta da subunidade i , em metros quadrados; g = área basal individual dos elementos da subunidade i em metros quadrados. Obtida pela expressão:

$$g = \frac{\pi * DAP^2}{40000} \quad (10)$$

- Índice *Basal area larger* (BAL): segundo Schneider (2008), este é um índice de competição entre as árvores, independente da distância entre as mesmas, obtido pelo quociente entre a área basal por hectare e a área basal da árvore em questão.

No entanto, optou-se por seguir a modificação utilizada por Ivanov (2008) em calcular o equivalente em área basal por hectare a partir da subunidade, sendo obtido pela equação:

$$BAL = \frac{\left(\frac{(\sum gi) \times 10000}{100} \right)}{g} \quad (11)$$

Onde: g_i = área basal da subunidade i ; g = área basal da árvore.

A unidade de medida dessa variável é o número de árvores por hectare (n° arv.ha⁻¹) e quanto maior o valor obtido no índice, maior é a competição que a árvore esta submetida.

3.4 Tratamento estatístico

Os dados gerados, inicialmente, foram analisados por meio da estatística descritiva, a fim de se conhecer a variabilidade de cada variável. Feito isso, utilizou-se a regressão logística univariada (RLU) e múltipla (RLM) para identificar as variáveis associadas à mortalidade das árvores construindo-se um modelo para explicar a relação entre elas.

As análises estatísticas foram realizadas com auxílio do aplicativo computacional PASW, versão 17.0, para *Windows*.

3.5 Modelagem da mortalidade por meio da regressão logística

O processo de modelagem inclui o teste da significância dos coeficientes das variáveis independentes, a seleção das covariáveis no modelo logístico, e a avaliação do ajuste deste modelo. Esse processo foi realizado similarmente aos

roteiros de análise utilizados por Mezzomo (2009) e Garcia (2010) e é descrito a seguir.

Inicialmente é feita a análise de regressão logística univariada, na qual a significância do coeficiente estimado de cada covariável é verificada pelo teste de Wald. Nessa etapa da análise, cada covariável foi testada com a variável dependente utilizando o método *enter*, onde se obtiveram a razão de chances (*Odds Ratio* bruto), o intervalo de confiança para a razão de chances, o p-valor do teste Wald e o erro padrão estimado de cada covariável. A interação entre as variáveis independentes não foi testada por haver um conjunto grande de covariáveis em relação ao número de observações que compõe o banco de dados. Utilizando o critério sugerido por Hosmer e Lemeshow (1989) a variável que apresentou p-valor $\leq 0,25$ no teste de Wald, foi considerada candidata a compor o modelo de regressão logística múltipla.

Antes da modelagem da RLM, procedeu-se o diagnóstico da multicolinearidade, em que são observados os valores de Fator de Inflação da Variância (FIV) das variáveis independentes, a fim de se detectar a multicolinearidade entre elas. Segundo Schneider, Schneider, Souza (2009) e Field (2009), valores de FIV maiores que 10 indicam sérios problemas de multicolinearidade.

Concluídas as etapas anteriores, foram utilizados três métodos para a seleção das covariáveis no modelo de RLM, que foram: *enter*, *backward* e *forward*. A significância dos coeficientes das variáveis independentes na RLM foi verificada pela razão de chances (*Odds Ratio* ajustado) e pelo teste de Wald. Utilizou-se como critério para a permanência da variável independente no modelo final de RLM p-valor menor que 5% pelo teste de Wald.

No método *enter* todas as covariáveis foram inseridas, simultaneamente, no modelo logístico, e a eliminação de variáveis não significativas foi feita manualmente, ou seja, uma de cada vez, e novamente o modelo era ajustado, até o momento que restassem apenas as variáveis significativas pelo teste de Wald, no modelo múltiplo.

Os outros dois métodos são automáticos e realizam o procedimento de modelagem passo a passo (*stepwise*). No método *forward* o modelo logístico inicial foi composto apenas com a constante e a seguir, a cada passo, foram inseridas, uma a uma, as variáveis independentes significativas pelo valor de *score* e

removidas baseado na probabilidade do teste de Wald. Já no método *backward* o modelo inicial foi composto por todas as variáveis independentes, e, passo-a-passo, cada variável não significativa no teste de Wald foi removida.

Dessa forma, foram obtidos três modelos de RLM para a mortalidade das árvores. A significância dos coeficientes também foi testada pelo teste da razão de verossimilhança, a qualidade do ajuste dos modelos foi verificada pelo R^2 de Nagelkerke, Teste de Hosmer e Lemeshow, tabela de classificação, e pela pertinência prevista ao grupo (DIAS FILHO; CORRAR, 2007; FIELD, 2009).

4 RESULTADOS E DISCUSSÕES

Inicialmente foram analisadas 1756 árvores correspondentes a duas unidades amostrais. Desse número de observações, 84 foram desconsideradas na análise por apresentarem dados faltantes ou valores incompatíveis com a realidade, restando assim 1672 árvores para o estudo.

Dessas, conforme a tabela 1, 68 árvores estavam mortas, que correspondem a uma taxa de mortalidade de 4,1 %, valor considerado alto se comparado com as taxas obtidas para 10 parcelas em 10 anos no mesmo local, apresentadas por Ebling (2012), em que a maior taxa foi de 2,6 % obtida no ano de 2006.

Tabela 1 – Frequência da variável dependente

| Variável | Classe | Frequência | Percentual |
|----------|--------|------------|------------|
| estado | Viva | 1604 | 95,9 |
| | Morta | 68 | 4,1 |
| | Total | 1672 | 100,0 |

Para descrever as variáveis independentes foram utilizadas as 1672 árvores distribuídas em 200 subunidades que perfazem o banco de dados. Na tabela 2 pode se observar os valores mínimo, máximo, médio, quartis e de dispersão (desvio padrão e coeficiente de variação) para as variáveis independentes quantitativas. Para as variáveis independentes qualitativas são apresentadas as distribuições de frequências na tabela 3.

Observa-se através dos valores apresentados na tabela 2, que a variável DAP possui grande variabilidade, caracterizada pelo valor do CV, de 57 %, e pela amplitude nos valores de diâmetro. Nota-se que o valor da média, primeiro e terceiro quartis dessa variável são próximos ao seu valor mínimo, o que significa que existem muitas árvores com pequeno diâmetro e esse número de árvores vai diminuindo com o aumento do diâmetro, até que restem poucas árvores com grandes dimensões. Esse é o comportamento esperado dessa variável em floresta natural, o que caracteriza um tipo de distribuição diamétrica que se enquadra no

modelo exponencial negativo, ou chamado de J-invertido (DURIGAN, 2009; SOARES; PAULA NETO; SOUZA, 2011).

A variável DAP² possui o mesmo comportamento do DAP quanto às estatísticas descritivas, exceto pela sua variabilidade, que é bastante elevada, com desvio padrão, de 870 cm², e CV, de 138 %. Isso se deve ao fato de que essa variável é uma transformação do diâmetro e por isso acompanha os seus valores.

Tabela 2 – Estatísticas descritivas das variáveis independentes quantitativas

| Variáveis | Sigla | Mínimo | Máximo | Média | s | CV | Q1 | Q3 |
|--|------------------|--------|--------|-------|-------|-----|-------|-------|
| diâmetro à altura do peito (cm) | DAP | 10 | 93 | 22 | 12 | 57 | 13 | 26 |
| diâmetro ao quadrado (cm ²) | DAP ² | 91 | 8686 | 630 | 870 | 138 | 164 | 696 |
| inverso do diâmetro (cm ⁻¹) | DAP-1 | 0,01 | 0,10 | 0,06 | 0,02 | 42 | 0,04 | 0,08 |
| incremento em diâmetro (cm) | IDAP | -2,01 | 3,57 | 0,09 | 0,62 | 707 | -0,22 | 0,35 |
| altura da copa (m) | h copa | 0,0 | 16,6 | 6,5 | 3,3 | 50 | 4,5 | 8,7 |
| altura total (m) | ht | 2 | 26 | 13 | 4 | 28 | 11 | 16 |
| grau de esbeltez | GE | 0,1 | 1,5 | 0,7 | 0,3 | 38 | 0,5 | 0,9 |
| área basal absoluta (m ² /sub.) | gabs | 0,038 | 1,886 | 0,474 | 0,230 | 48 | 0,315 | 0,581 |
| densidade absoluta (árv./sub.) | dens | 1 | 19 | 10 | 3 | 34 | 7 | 12 |
| altura média (m) | h méd | 6 | 18 | 13 | 2 | 13 | 12 | 14 |
| diâmetro mínimo (cm) | DAPmín | 10 | 18 | 11 | 1 | 11 | 10 | 11 |
| diâmetro máximo (cm) | DAPmáx | 14 | 93 | 44 | 14 | 32 | 34 | 53 |
| amplitude do diâmetro (cm) | ADAP | 3 | 83 | 34 | 14 | 42 | 24 | 43 |
| índice basal area larger (árv.ha) | BAL | 122 | 19574 | 2264 | 2117 | 94 | 685 | 3260 |

Em que: s = desvio padrão; CV % = coeficiente de variação; Q1= primeiro quartil; Q3 = terceiro quartil.

O inverso do diâmetro possui um comportamento diferente das variáveis anteriores, simetricamente distribuído (observar valores mínimos, máximo, média e quartis), com pequena dispersão absoluta (s = 0,02) e intermediária dispersão relativa (CV = 42 %).

O incremento em DAP exibiu grande amplitude de valores e extraordinário CV, de 707 %. A possível causa para os valores elevados é a mudança do local de medição do diâmetro de um ano para o outro, que originou dígitos irrealis de incremento tanto positivo (valor máximo) quanto negativo (mínimo). O incremento negativo em diâmetro ocorre frequentemente quando as árvores são remedidas em um intervalo de tempo curto, mas esses valores devem ser de pequena grandeza (SOARES; PAULA NETO; SOUZA, 2011).

A grande amplitude nos valores obtidos de incremento em diâmetro acarretou um valor de média pequeno. Mas, embora com esses problemas, essa variável foi mantida na análise pela sua importância para a mortalidade, pelo fato que incrementos próximos a zero podem ser indicativos da perda de vigor e sugerem a consequente morte da árvore (BUCHMAN, 1979 apud ROSSI et al., 2007).

Em relação à variável altura de copa o valor máximo corresponde a uma grande altura de copa (16,6 m) o valor mínimo (zero) é de árvores mortas ou morrendo que não possuem copa, e sua variabilidade é intermediária. Observando os valores de média e quartis nota-se a distribuição simétrica dos valores.

Quanto à altura total, suas estatísticas descritivas demonstram comportamento semelhante as da variável altura de copa, exceto pela menor dispersão relativa de ht, com CV de 28 %, considerada baixa. Os valores de mínimo e máximo são possíveis e esperados, já que existem na floresta espécies de pequeno porte como o Xaxim e outras de grande porte como a Araucária.

A variável grau de esbeltez possui grande amplitude de valores (de 0,1 até 1,5), e observando seu valor médio (de 0,7) e o terceiro quartil (0,9), constata-se que a maior parte das árvores (75 %) possuem valor inferior a unidade, e portanto podem ser consideradas árvores estáveis.

A variável área basal absoluta possui grande amplitude entre o valor mínimo e máximo, no entanto metade das subunidades possuem valores de gabs concentrados em torno da média (observar média, de 0,474, primeiro quartil, 0,315, e terceiro quartil, 0,581).

O comportamento dos dados referentes à densidade absoluta foi semelhante ao da variável gabs, diferindo apenas na menor dispersão dos valores de densidade, com CV, de 34 %, que pode ser considerado um valor intermediário de variação. Sua média foi de 10 árvores por subunidade, com valor mínimo de 1, e máximo de 19, que representam baixa e elevada densidade, respectivamente.

Quanto à altura média e diâmetro mínimo observam-se os dados concentrados em torno da média e a menor variabilidade do conjunto de variáveis com CV, de 13 e 11 %, respectivamente.

As variáveis DAPmáx e ADAP correspondem à amplitude de valores descrita para o DAP. Dessas variáveis destacam-se os valores mínimo e máximo de ADAP que possui grande magnitude, variando de uma pequeníssima amplitude (valor mínimo igual a 3) até grande amplitude (valor máximo igual a 83) e que

correspondem a uma alta e baixa competição potencial, respectivamente. Além disso, DAP_{máx} apresenta o mesmo comportamento dos valores de ADAP, porém com menor dispersão relativa (CV = 32 %).

A variável indicativa da competição entre as árvores, BAL elucidou a grande variabilidade existente nas subunidades e por consequência na competição que incide sobre as árvores. O valor médio de BAL foi de 2264 árvores que indica considerável nível de competição, já os valores mínimo (122) e máximo (19574) sugerem inexistência e altíssima competição, respectivamente.

As demais variáveis são apresentadas na tabela 3, onde constam as variáveis, suas categorias, frequência e percentual.

Tabela 3 – Distribuição de frequência das variáveis independentes qualitativas

| Variáveis | Sigla | Classe | Frequência | Percentual |
|-----------------------|--------------|--------------------|-------------------|-------------------|
| integridade da copa | IC | íntegra | 1594 | 95 |
| | | danificada | 78 | 5 |
| classe de copa 1 | CC1 | longa/curta | 754 | 45 |
| | | média | 918 | 55 |
| classe de copa 2 | CC2 | longa/média | 1082 | 65 |
| | | curta | 590 | 35 |
| posição sociológica 1 | PS1 | demais | 938 | 56 |
| | | codominante | 734 | 44 |
| posição sociológica 2 | PS2 | demais | 930 | 56 |
| | | dominada/suprimida | 742 | 44 |

Analisando os valores presentes na tabela 3, observa-se que em relação à integridade da copa, apenas 5% das árvores possuem copa danificada e a grande maioria das árvores possui copa íntegra (95%). Na variável classe de copa 1 destaca-se que 55% das copas são de tamanho médio, e em classe de copa 2, pode-se observar que 35% das copas são pequenas.

Quanto à posição sociológica as frequências foram bem distribuídas nas duas classes para as duas variáveis, PS1 e PS2. Observa-se que 44% das árvores são codominantes (PS1) e 44,4% das árvores pertencem a classe dominada/suprimida (PS2).

4.1 Análise de regressão logística univariada

Na tabela 4 são apresentadas as 19 covariáveis, os respectivos coeficientes estimados (B), *Odds Ratio* bruto (OR bruto), o intervalo de confiança para o OR (IC-95%), erro padrão, e significância no teste de Wald (p-valor).

Na RLU, a significância no teste de Wald é atingida com um p-valor $\leq 0,25$. Analisando os valores apresentados na tabela 4, pode-se observar que três covariáveis não são significativas. São elas: altura de copa, classe de copa 2 e área basal absoluta.

As demais variáveis obtiveram a significância exigida pelo teste de Wald, portanto seus coeficientes estimados diferem significativamente de zero, e assim, essas covariáveis serão testadas no modelo de regressão logística múltipla.

Tabela 4 – Estatísticas da regressão logística univariada

| Variáveis | B | OR bruto | IC -95% | E.P. | p-valor |
|----------------------------|---------|------------|--------------------|---------|---------|
| diâmetro à altura do peito | -0,035 | 0,966 | 0,939 - 0,993 | 0,014 | 0,013 |
| diâmetro ao quadrado | -0,001 | 0,999 | 0,999 - 1,00 | 0,000 | 0,018 |
| inverso do diâmetro | 12,964 | 426748,973 | 16,609 - 1,097E+10 | 5,181 | 0,012 |
| incremento em diâmetro | -1,111 | 0,329 | 0,216 - 0,502 | 0,215 | 0,000 |
| altura da copa | -73,381 | 1,352E-32 | 0,000 - 2,172E+257 | 339,762 | 0,829 |
| altura total | -0,216 | 0,806 | 0,760 - 0,854 | 0,030 | 0,000 |
| grau de esbeltez | -1,297 | 0,273 | 0,110 - 0,677 | 0,463 | 0,005 |
| integridade da capo | 0,722 | 2,059 | 0,862 - 4,918 | 0,444 | 0,104 |
| classe de copa 1 | -0,390 | 0,677 | 0,416 - 1,102 | 0,249 | 0,117 |
| classe de copa 2 | 0,261 | 1,298 | 0,792 - 2,126 | 0,252 | 0,301 |
| posição sociológica 1 | -0,806 | 0,446 | 0,258 - 0,772 | 0,279 | 0,004 |
| posição sociológica 2 | 1,003 | 2,728 | 1,626 - 4,577 | 0,264 | 0,000 |
| área basal absoluta | -0,617 | 0,539 | 0,168 - 1,730 | 0,595 | 0,299 |
| densidade absoluta | -0,076 | 0,927 | 0,857 - 1,002 | 0,040 | 0,056 |
| altura média | -0,144 | 0,866 | 0,764 - 0,980 | 0,064 | 0,023 |
| diâmetro mínimo | -0,169 | 0,845 | 0,661 - 1,079 | 0,125 | 0,177 |
| diâmetro máximo | -0,013 | 0,987 | 0,969 - 1,006 | 0,009 | 0,169 |
| amplitude do diâmetro | -0,012 | 0,988 | 0,971 - 1,007 | 0,009 | 0,209 |
| índice basal area larger | 0,000 | 1,000 | 1,000 - 1,000 | 0,000 | 0,045 |

Onde: B = coeficiente estimado; OR bruto = *Odds Ratio* bruto; IC-95% = intervalo de confiança para OR; E.P. = Erro padrão.

Quanto aos valores de OR bruto pode-se destacar o valor extremamente elevado da variável DAP⁻¹. Segundo Hosmer e Lemeshow (1989) valores irrealis do

coeficiente estimado e ou erro padrão demasiadamente grande são indicativos do chamado *overfitting*, que pode ocorrer quando o número de variáveis é grande em relação ao número de observações.

Na tentativa de superar esse problema, a variável DAP^{-1} foi categorizada (denominada de DAP1cat) e testada novamente na RLU.(Anexo A) O resultado foi que a covariável manteve a significância no teste de Wald e o valor de OR bruto e erro padrão foram semelhantes aos das demais variáveis independentes.

Antes de iniciar a modelagem da RLM, testou-se a multicolinearidade das 15 covariáveis selecionadas na RLU acrescidas da variável DAP1cat. As 16 covariáveis testadas foram: DAP, DAP^2 , IDAP, ht, GE, IC, CC1, PS1, PS2, dens., hméd, DAPmáx, DAPmín, ADAP, BAL e DAP1cat.

Na tabela 5 são apresentadas as covariáveis testadas e seus respectivos valores de fator de inflação da variância (FIV) inicial, com todas as covariáveis e final, com a eliminação de covariáveis.

Tabela 5 – Estatística da multicolinearidade para as covariáveis significativas na análise de regressão logística univariada

| Variáveis | Sigla | FIV inicial | FIV final |
|-------------------------------------|--------------|--------------------|------------------|
| diâmetro à altura do peito | DAP | 114 | 6 |
| diâmetro ao quadrado | DAP^2 | 44 | -- |
| incremento em diâmetro | IDAP | 1 | 1 |
| altura total | ht | 11 | 2 |
| grau de esbeltez | GE | 17 | -- |
| integridade da copa | IC | 1 | 1 |
| classe de copa 1 | CC1 | 1 | 1 |
| posição sociológica 1 | PS1 | 4 | 4 |
| posição sociológica 2 | PS2 | 5 | 5 |
| densidade absoluta | dens | 1 | 1 |
| altura média | h méd | 2 | 2 |
| diâmetro mínimo | DAP mín | 1 | 1 |
| diâmetro máximo | DAP máx | 2 | 2 |
| inverso do diâmetro categórico | DAP1cat | 11 | 4 |
| índice basal area larger categórico | BALcat | 4 | 4 |

Onde: FIV = fator de inflação da variância

No conjunto de variáveis testadas, apresentadas na tabela 5, observa-se que foram encontradas cinco covariáveis com valores de FIV acima de 10. Verificou-se que eliminando as variáveis DAP^2 e GE, que são variáveis secundárias

(transformações de variáveis medidas no campo), não havia mais valores de FIV acima de 10, portanto o problema severo de multicolinearidade foi resolvido.

Dessa forma, as variáveis restantes foram testadas no modelo de regressão logística múltipla, apresentado a seguir.

4.2 Análise de regressão logística múltipla

São descritos os modelos construídos por meio de cada método, os coeficientes estimados de cada covariável (B), a significância no teste de Wald (p-valor), o *Odds Ratio* ajustado (OR ajustado) e o intervalo de confiança de 95 % para o OR ajustado (IC-95 %).

Analisando a tabela 6, pode-se observar que o modelo final construído, utilizando o método *enter*, é composto por três variáveis independentes, que são: IDAP, ht e dens. Os coeficientes estimados das variáveis são negativos, o que significa que a influência das variáveis para a ocorrência da mortalidade é negativa.

Tabela 6 – Modelos encontrados pelos três métodos de seleção

| Método | Variáveis | B0 | B1 | B2 | B3 | OR ajust | IC 95% | p-valor |
|-----------------|-----------|---------|---------|---------|---------|----------|---------------|---------|
| <i>Enter</i> | IDAP | | -0,8894 | | | 0,411 | 0,263 - 0,643 | < 0,001 |
| | ht | | | -0,1941 | | 0,824 | 0,776 - 0,874 | < 0,001 |
| | dens. | | | | -0,0830 | 0,920 | 0,848 - 0,999 | 0,049 |
| | | -0,1881 | | | | 0,829 | | |
| <i>Forward</i> | IDAP | | -0,8894 | | | 0,411 | 0,263 - 0,643 | < 0,001 |
| | ht | | | -0,1941 | | 0,824 | 0,776 - 0,874 | < 0,001 |
| | dens. | | | | -0,0830 | 0,920 | 0,848 - 0,999 | 0,049 |
| | | -0,1881 | | | | 0,829 | | |
| <i>Backward</i> | IDAP | | -0,8894 | | | 0,411 | 0,263 - 0,643 | < 0,001 |
| | ht | | | -0,1941 | | 0,824 | 0,776 - 0,874 | < 0,001 |
| | dens. | | | | -0,0830 | 0,920 | 0,848 - 0,999 | 0,049 |
| | | -0,1881 | | | | 0,829 | | |

Pode-se interpretar diretamente a influência de cada covariável na mortalidade das árvores, por meio do valor de OR ajustado. Observa-se que o valor de OR ajustado para a variável IDAP é de 0,41. Como esse valor é inferior a um,

interpreta-se que o aumento em uma unidade de IDAP diminui em 59% a chance de que uma árvore morra. Em relação às demais variáveis, pode-se fazer a mesma interpretação, o que difere é a magnitude da influência de cada variável. Nota-se o valor de OR ajustado para altura total de 0,82, portanto, o aumento em um metro na altura da árvore acarreta 18% de chance de que a árvore não morra. Da mesma forma para variável dens., que seu aumento em uma unidade, diminui o risco de morte da árvore em 8%.

A influência significativa de características morfométricas (IDAP e ht) na mortalidade das árvores, também foi observado nos estudos da mortalidade de espécies florestais desenvolvidos por Monserud e Sterba (1999), Schneider et al. (2005) e Eid e Tuhus (2001).

Yao, Titus e MacDonald (2001) descreveram que a influência do sítio na mortalidade de árvores em Alberta, no Canadá, é dependente da espécie, podendo ser negativa ou positiva. Yang, Titus e Huang (2003), baseados no estudo supracitado, afirmam que não existe uma tendência óbvia entre o sítio e a mortalidade, por isso não consideraram essa variável em um modelo de mortalidade desenvolvido para uma espécie, em Alberta. No modelo de mortalidade desenvolvido no presente estudo, encontrou-se uma relação entre a mortalidade e o sítio, no entanto essa relação foi fraca, pois, se observou apenas uma variável de sítio (dens.) compondo o modelo e com pequena influência sobre a mortalidade.

O modelo final obtido pelo método *forward: Wald* foi atingido no quarto passo, e é idêntico ao modelado pelo método *enter*. Portanto, verificou-se a associação das mesmas variáveis com a mortalidade, e a maneira como cada variável a influencia, foi descrita anteriormente.

Da mesma maneira, a modelagem realizada com o método *backward: Wald* resultou em um modelo final após 12 passos, e esse modelo também foi idêntico aos modelos anteriores.

Assim, foi obtido pelos três métodos de seleção de variáveis, o mesmo modelo final de regressão logística múltipla para estimar a probabilidade de ocorrência de uma árvore morta na área do presente estudo, que é apresentado na seguinte equação:

$$\text{Probabilidade da árvore morrer} = \frac{1}{1 + e^{-(0,1881 - 0,8894IDAP - 0,1941ht - 0,0830dens.)}} \quad (12)$$

Na tabela 7, constam as medidas de avaliação do ajuste do modelo logístico, com os valores do teste da razão de verossimilhança, R² de Nagelkerke e o Teste de Hosmer e Lemeshow.

Tabela 7 – Medidas de avaliação da qualidade do ajuste do modelo de RLM.

| Razão de verossimilhança | | R ² Nagelkerke | Teste Hosmer e Lemeshow | |
|--------------------------|---------|---------------------------|-------------------------|---------|
| Qui-quadrado | p-valor | | Qui- quadrado | p-valor |
| 71,086 | < 0,001 | 0,144 | 10,453 | 0,235 |

Observando os valores apresentados na tabela 7, nota-se que o modelo foi significativo pelo teste da razão de verossimilhança, o que significa que ao menos um coeficiente de uma variável é diferente de zero.

O valor da medida de R² de Nagelkerke que representa a capacidade de explicação do modelo, é baixo, de apenas 0,144, o que significa que apenas 14,4 % das variações ocorridas na variável dependente são explicadas pelas variáveis independentes. Longhi et al. (2010) encontraram valores de R² de Nagelkerke de 0,149 e 0,208 para modelos logísticos de ocorrência das espécies *Araucaria angustifolia* e *Podocarpus lambertii* no mesmo local do presente estudo, evidenciando que de maneira geral, os modelos logísticos da área florestal construídos com a mesma base de dados, são de pouco poder de explicação.

A melhor avaliação referente à qualidade do ajuste do modelo foi a obtida no Teste de Hosmer e Lemeshow, em que, como o resultado para o modelo não foi significativo neste teste, indica que suas estimativas não diferem dos valores observados.

Entretanto, analisando a tabela 8, onde são apresentados os valores observados nas duas classes da variável dependente (árvores vivas e mortas) e também os estimados nessas duas classes, pode-se fazer a seguinte constatação: na diagonal principal das duas colunas de valores da tabela 7 estão as árvores observadas a campo que estavam vivas e foram estimadas pelo modelo como vivas,

e as árvores que foram observadas no campo como mortas e foram estimadas como mortas, portanto, é o que o modelo classifica corretamente. Na diagonal secundária está o número de árvores que o modelo classifica erroneamente.

Tabela 8 – Tabela de classificação.

| Observado | Estimado | | % correta |
|-----------|----------|-------|-----------|
| | viva | morta | |
| viva | 1604 | 0 | 95,9 |
| morta | 68 | 0 | |

Nota-se, a partir dos valores apresentados na tabela 8, que as 1604 árvores observadas como vivas foram corretamente estimadas pelo modelo logístico como árvores vivas, no entanto, nenhuma das 68 árvores observadas como mortas foram corretamente estimadas como mortas. Ou seja, 1604 árvores, que correspondem a 96 % do número total de árvores, foram corretamente classificadas. Esse é um alto percentual de classificação correta do modelo logístico, se comparado ao encontrado por Monserud e Sterba (1999) de 86 %, e os de Caldato (1998) de 76 % e 62 %. No entanto, o modelo não conseguiu estimar corretamente uma única árvore morta.

Essa limitação na capacidade preditiva do modelo também foi observada na análise da pertinência prevista ao grupo, na qual se evidencia que o modelo não estimou nenhuma das 68 árvores mortas, na classe correta (Apêndice B).

Ora, tem-se então, uma contradição entre as medidas que avaliam a qualidade do ajuste do modelo logístico. Por um lado o resultado no teste de Hosmer e Lemeshow nos leva a afirmar que os valores preditos não são significativamente diferentes dos valores observados, por outro lado a tabela de classificação e o resíduo denominado grupo previsto demonstram que o modelo não consegue estimar as árvores mortas.

Provavelmente, a causa desse problema esteja no banco de dados do presente estudo. A grande variabilidade que envolve o processo da mortalidade das árvores e de seus fatores associados, aliado a um número de observações relativamente pequeno, acarretam um modelo de baixa qualidade. Além disso, a tentativa de desenvolver um modelo global de mortalidade das árvores, agregando

todas as espécies, contribuiu para a baixa capacidade preditiva do modelo, uma vez que, cada espécie ou grupo de espécies possui características e necessidades específicas. Exemplo disso é a diferença entre o porte de duas espécies características da Floresta Ombrófila Mista, que são a Araucária e o Xaxim. Enquanto a primeira pode atingir 30 m de altura e dominar o estrato vertical da floresta, a segunda chega no máximo a cinco metros de altura, por isso, os fatores determinantes da mortalidade são de certa forma, específicos para cada espécie.

5 CONCLUSÃO

Os resultados obtidos em relação à modelagem da mortalidade de espécies arbóreas em fragmento da Floresta Ombrófila Mista, no RS, permitem concluir que:

- a) foram encontrados três fatores associados à mortalidade das árvores, e a influência que esses fatores exercem na sua ocorrência é negativa;
- b) as variáveis que influenciam a mortalidade são o incremento em diâmetro da árvore, a altura total da árvore e a densidade de árvores vivas na subunidade
- c) o modelo de regressão logística múltipla construído mostrou-se de pequena capacidade preditiva;
- d) o teste de Hosmer e Lemeshow mostrou-se pouco sensível a pequenas diferenças entre valores estimados e observados.

Limitações

A tentativa de construir um modelo de mortalidade para todas as espécies arbóreas do fragmento da Floresta Ombrófila Mista na FLONA – SFP, utilizando um número de observações pequeno, resultou em um modelo pouco ajustado; e por isso não foi feita a validação do modelo.

Sugestões

Utilizar uma amostra grande para construir um modelo de mortalidade para árvores, e modela-la para grupos de espécies, ou se possível, para cada espécie.

REFERÊNCIAS

BRASIL. Instituto Brasileiro do Meio Ambiente. **Floresta Nacional de São Francisco de Paula** - RS. Brasília: IBAMA, 2000. 6p.

CALDATO, S.L. **Dinâmica populacional da Ocotea porosa (Lauraceae) na floresta ombrófila mista em Caçador, SC.** 1998. 89 f. Dissertação (Mestrado em Engenharia Florestal) – Programa de Pós-graduação em Engenharia Florestal, Universidade Federal de Santa Maria, RS. 1998.

CHASSOT, T. **Modelos de crescimento em diâmetro de árvores individuais de Araucaria angustifolia (Bertol.) Kuntze na Floresta Ombrófila Mista.** Dissertação de Mestrado, Programa de Pós Graduação em Engenharia Florestal, Santa Maria, RS. 48 pp. 2009.

DELLA-FLORA, J. B.; DURLO, M. A.; SPATHELF, P. Modelo de incremento para árvores singulares - *Nectandra megapotamica* (Spreng.) Mez. **Ciência Florestal**, Santa Maria, v. 14, n. 1, p. 165-177, jun. 2004.

DIAS FILHO J. M; CORRAR, L. J. Regressão Logística. p.281-323 In: CORRAR; L. J. et al. (coord.) **Análise Multivariada: para os cursos de administração, ciências contábeis e economia.** Atlas, São Paulo, 2007. 540p.

DURIGAN, G. Estrutura e Diversidade de Comunidades Florestais. In: MARTINS, S. V. (Ed.) **Ecologia de florestas tropicais do Brasil.** Editora UFV. Universidade Federal de Viçosa. 2009. 261 p.

EBLING, A. A. **Dinâmica e projeção diamétrica em remanescente de Floresta com Araucária na Flona de São Francisco de Paula, RS** 127 p. Dissertação (Mestrado em Engenharia Florestal) – Universidade Estadual do Centro-Oeste, Irati, PR, 2012.

EID, T., TUHUS, E., Models for individual tree mortality in Norway. **Forest Ecology and Management** 154:69-84. 2001.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **State of the World's Forests.** Roma, Italy, 2011. 179p.

FIELD, A. **Descobrimos a estatística usando SPSS**. Tradução Lorí Viali. -2.ed. – Porto Alegre: Artmed, 2009. 688p.

FIGUEIRA, C. V. **Modelos de Regressão Logística**. Porto Alegre. Dissertação (mestrado em matemática) Programa de Pós-graduação em Matemática do Instituto de Matemática Universidade Federal do Rio Grande do Sul. 2006

FUNDAÇÃO SOS MATA ATLÂNTICA – Disponível em:
<<http://www.sosmatatlantica.org.br>> Acesso: 02 fev. 2011.

GARCIA, J. T. **Identificação de variáveis determinantes na seleção de candidatos, para os cursos de engenharia, no processo seletivo da Universidade Federal de Santa Maria, RS**. 65 p. Monografia (Especialização em Estatística e Modelagem quantitativa) – Universidade Federal de Santa Maria, Santa Maria, RS, 2010.

HAIR, J. F. et al. **Análise multivariada de dados**. 5.ed. Porto Alegre: Bookman, 2005. 593p.

HESS, A. F. **Inter-relações no crescimento de Araucaria angustifolia (Bertol.) Kuntze em diferentes locais do Rio Grande do Sul**. 177f. Tese (Doutorado em Engenharia Florestal) - Universidade Federal de Santa Maria, Santa Maria, 2006.

HOSMER, D. W. Jr.; LEMESHOW, S. **Applied logistic regression**. New York: Wiley, 1989. 382p.

IVANAUSKAS, N. M.; ASSIS, M. C. Formações Florestais Brasileiras. In: MARTINS, S. V. (Ed.) **Ecologia de florestas tropicais do Brasil**. Editora UFV. Universidade Federal de Viçosa. 2009. 261 p.

IVANOV, G. B. Modelagem do incremento periódico anual em área basal de *Eucalyptus dunnii* em função de variáveis dendrométricas e índices de competição. **Trabalho de Conclusão de curso**. Engenharia Florestal-UFSM 2008. 43p

LEITE, P. F.; KLEIN, R. M. Vegetação. In: IBGE. Fundação Instituto Brasileiro de Geografia e Estatística. **Geografia do Brasil: Região Sul**. Rio de Janeiro: 1990. p. 113-150.

LONGHI, S. J. et al. Fatores ecológicos determinantes na ocorrência de *Araucaria angustifolia* e *Podocarpus lambertii*, na Floresta Ombrófila Mista da FLONA de São

Francisco de Paula, RS, Brasil **Ciência Rural** [online]. 2010, vol.40, n.1, pp. 57-63. Disponível em <http://www.scielo.br/scielo.php?pid=S01034782010000100010&script=sci_abstract&tlng=pt> Acesso: 10 dez. 2010.

MATTOS, R. B. **Produtividade e incremento de *Cabralea canjerana* (Vell.) Mart., *Cedrela fissilis* Vell. e *Cordia trichotoma* (Vell.) Arrab. Ex Steud., em floresta nativa no Rio Grande do Sul. Santa Maria, 2007.** 105f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal de Santa Maria, 2007.

MEDEIROS, J. de D.; SAVI, M.; BRITO, B. F. A. Seleção de áreas para a criação de unidades de conservação na floresta ombrófila mista. **Biotemas**, 18 (2): 33 - 50, 2005

MEZZOMO, M. **Estudo da mortalidade infantil- um modelo de regressão logística múltipla** 59 p. Monografia (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, Santa Maria, RS, 2009.

MINISTÉRIO DO MEIO AMBIENTE / SERVIÇO FLORESTAL BRASILEIRO. **Avaliação e ações prioritárias para a conservação da biodiversidade da Mata Atlântica e Campos Sulinos.** Brasília: MMA/ SBF, 2000. 40p.

MONSERUD, R.A.; STERBA, H. Modeling individual tree mortality for Austrian forest species. **Forest Ecology and Management**, v.113, p.109-123, 1999.

ROSSI, L.M.B.; et al. Modelagem da mortalidade em florestas naturais. **Revista Floresta**, v.37, p.275-291, 2007.

ROSSATO, M. S. **Os climas do Rio Grande do Sul: variabilidade, tendências e tipologia.** 240f. Tese (doutorado). Programa de Pós-Graduação Geografia, Instituto de Geociências, UFRGS. Porto Alegre, 2011.

SCHNEIDER, P.R. et al. Análise da mortalidade de *Acacia mearnsii* De Wild. **Ciência Florestal**, v.15, p.137-143, 2005.

SCHNEIDER, P. R., SCHNEIDER, P. S. P., SOUZA, C. A. M. **Análise de regressão aplicada à Engenharia Florestal.** 2 ed. rev. e ampl. Santa Maria: FACOS, 2009. 294p.

SCHNEIDER, P. R. **Manejo Florestal: planejamento da produção florestal.** Santa Maria, 2008. 500 p.

SFB – SERVIÇO FLORESTAL BRASILEIRO. **Florestas do Brasil em resumo – 2010: dados de 2005 – 2010**. Brasília: SFB, 2010. 152p.

SIMS, A., et al. Estimating Tree Survival: a Study based on the Estonian Forest Research Plots Network. **Annales Botanici Fennici**. 46:336-352. 2009.

SOARES, C. P. B.; PAULA NETO, F.; SOUZA, A. L. **Dendrometria e inventário florestal**. Viçosa, MG: Universidade Federal de Viçosa, 2 ed. 272p. 2011.

STRECK, E. V. et al. **Solos do Rio Grande do Sul**. 2 ed. Porto Alegre: EMATER/RS, 2008. 222p.

VANCLAY, J .K. Mortality functions for north Queensland rainforests. **Journal of Tropical Forest Science** 4:15-36 1991

VANCLAY, J. K. **Modelling Forest Growth and Yield**: Applications to Mixed Tropical Forests. CAB International, Wallingford, U.K. 1994

ZANON, M. L. B. **Crescimento da *Araucaria angustifolia* (Bertol.) Kuntze diferenciado por dioiccia**. 2007. 107f. Tese (Doutorado em Engenharia Florestal) – Programa de Pósgraduação em Engenharia Florestal, Universidade Federal de Santa Maria, RS.

YAO, X., TITUS, S.J., MACDONALD, S.E. A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests. **Canadian Journal of Forest Research**, 31 (2), 283-291.2001

YANG, Y.; TITUS, S. J.; HUANG, S. Modeling individual tree mortality for white spruce in Alberta. **Ecological Modeling**, Amsterdam, v. 163, n. 3, p. 209-222, 2003.

Apêndice A– Categorias de inverso do diâmetro e estatísticas da regressão logística univariada

| Categorias de inverso do diâmetro (DAP⁻¹) | Código |
|---|---------------|
| $0,0107 \leq \text{DAP}^{-1} < 0,0342$ | 0 |
| $0,0342 \leq \text{DAP}^{-1} < 0,0577$ | 1 |
| $0,0577 \leq \text{DAP}^{-1} < 0,0812$ | 2 |
| $0,0812 \leq \text{DAP}^{-1} \leq 0,1047$ | 3 |

| Sigla | Coeficiente | OR bruto | Erro Padrão. | p-valor |
|--------------|--------------------|-----------------|---------------------|----------------|
| DAP1cat | | | | 0,090 |
| DAP1cat | 0,836 | 2,307 | 0,433 | 0,054 |
| DAP1cat | 0,546 | 1,727 | 0,464 | 0,239 |
| DAP1cat | 1,057 | 2,879 | 0,443 | 0,017 |

Apêndice B – Pertinência prevista ao grupo para as 68 árvores mortas

| Número do caso | Grupo predito | Número do caso | Grupo predito |
|----------------|---------------|----------------|---------------|
| 5 | Viva | 967 | Viva |
| 6 | Viva | 981 | Viva |
| 58 | Viva | 1.149 | Viva |
| 64 | Viva | 1.165 | Viva |
| 95 | Viva | 1.214 | Viva |
| 108 | Viva | 1.289 | Viva |
| 114 | Viva | 1.297 | Viva |
| 147 | Viva | 1.353 | Viva |
| 151 | Viva | 1.368 | Viva |
| 156 | Viva | 1.379 | Viva |
| 162 | Viva | 1.404 | Viva |
| 240 | Viva | 1.502 | Viva |
| 247 | Viva | 1.549 | Viva |
| 253 | Viva | 1.551 | Viva |
| 255 | Viva | 1.552 | Viva |
| 283 | Viva | 1.553 | Viva |
| 299 | Viva | 1.559 | Viva |
| 339 | Viva | 1.572 | Viva |
| 344 | Viva | 1.577 | Viva |
| 366 | Viva | 1.578 | Viva |
| 367 | Viva | | |
| 378 | Viva | | |
| 412 | Viva | | |
| 433 | Viva | | |
| 495 | Viva | | |
| 518 | Viva | | |
| 626 | Viva | | |
| 640 | Viva | | |
| 669 | Viva | | |
| 689 | Viva | | |
| 691 | Viva | | |
| 738 | Viva | | |
| 787 | Viva | | |
| 794 | Viva | | |
| 815 | Viva | | |
| 822 | Viva | | |
| 849 | Viva | | |
| 888 | Viva | | |