

UNIVERSIDADE FEDERAL DE VIÇOSA

JULIANA GALVÃO DE SOUSA MAGALHÃES

**PROGNOSE E REGULAÇÃO DE FLORESTAS DE EUCALIPTO
UTILIZANDO MODELO DE CRESCIMENTO E INTELIGÊNCIA
COMPUTACIONAL**

VIÇOSA
MINAS GERAIS – BRASIL
2014

JULIANA GALVÃO DE SOUSA MAGALHÃES

**PROGNOSE E REGULAÇÃO DE FLORESTAS DE EUCALIPTO
UTILIZANDO MODELO DE CRESCIMENTO E INTELIGÊNCIA
COMPUTACIONAL**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2014

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

M188p
2014

Magalhães, Juliana Galvão de Sousa, 1986-

Prognose e regulação de florestas de eucalipto utilizando modelo de crescimento e inteligência computacional / Juliana Galvão de Sousa Magalhães. – Viçosa, MG, 2014.

vii, 66f. : il. (algumas color.) ; 29 cm.

Orientador: Márcio Lopes da Silva.

Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Eucalipto - Cultivo. 2. Redes neurais artificiais.
3. Modelagem florestal. I. Universidade Federal de Viçosa.
Departamento de Engenharia Florestal. Programa de
Pós-graduação em Ciência Florestal. II. Título.

CDD 22. ed. 634.973766

JULIANA GALVÃO DE SOUSA MAGALHÃES

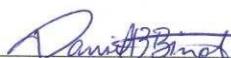
**PROGNOSE E REGULAÇÃO DE FLORESTAS DE EUCALIPTO
UTILIZANDO MODELO DE CRESCIMENTO E INTELIGÊNCIA
COMPUTACIONAL**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Magister Scientiae*.

APROVADA: 18 de fevereiro de 2014.



Prof. Helio Garcia Leite
(Coorientador)



Dr. Daniel Henrique Breda Binoti



Prof. Márcio Lopes da Silva
(Orientador)

*“A flor que desabrochou no jardim nos dá a paz
que precisamos em nosso dia a dia,
mas só podemos levar em prática a paz se quisermos,
pois o mundo será mais gigante no dia em que
a paz for a paz de VERDADE.”*

Chiquinho da Floresta

AGRADECIMENTOS

A minha família sergipana e a adquirida com bons amigos durante toda a minha estadia em Viçosa pelo apoio.

A oportunidade de cursar o mestrado proporcionada pelo professor Márcio Lopes e incentivada pelos professores Agostinho, Charles e Jacovine.

Ao desafio lançado pelo professor Helio Garcia com o tema da dissertação, além dos preciosos conselhos que sempre me disponibilizou.

Ao Eng. Florestal Daniel Binoti pelas considerações e suporte durante todas as etapas deste trabalho.

Ao conhecimento adquirido durante a graduação e mestrado no Departamento de Engenharia Florestal da Universidade Federal de Viçosa.

Ao apoio financeiro concebido pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

BIOGRAFIA

Juliana Galvão de Sousa Magalhães, filha de Johéline Magalhães do Nascimento e Marta Maria Galvão de Sousa Magalhães, nasceu em 30 de março de 1986, em Aracaju, Sergipe.

Em 2003, concluiu o 2º grau no COC - Colégio São Paulo, em Aracaju, Sergipe.

Em 2004, iniciou o curso de Engenharia Florestal na Universidade Federal de Viçosa, tendo o mesmo concluído em dezembro de 2008.

Entre 2006 e 2007 foi monitora da disciplina BVE230 – Organografia e Taxonomia das Espermatófitas.

No período de 2007 a 2008, foi bolsista de iniciação científica no Laboratório de Patologia Florestal, sendo mentora do projeto para teste de métodos alternativos para controle de oídio em minijardim clonal.

Durante janeiro de 2009 a maio de 2011, trabalhou como assessora técnica da Secretaria Municipal de Meio Ambiente da Prefeitura de Ponte Nova.

Neste mesmo intervalo, concluiu o curso de especialização em Plantas Ornamentais e Paisagismo pela Universidade Federal de Lavras.

Em janeiro de 2012, iniciou o programa de mestrado em Ciência Florestal na Universidade Federal de Viçosa, submetendo-se à defesa da dissertação em 17 de Fevereiro de 2014.

SUMÁRIO

RESUMO.....	vii
ABSTRACT.....	viii
1. INTRODUÇÃO	1
2. REFERÊNCIAS BIBLIOGRÁFICAS.....	4
CAPÍTULO 1: REVISÃO BIBLIOGRÁFICA.....	6
1. CRESCIMENTO E PRODUÇÃO DE POVOAMENTOS EQUIÂNEOS.....	6
2. MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO FLORESTAL	7
3. TIPOS DE MODELO DE CRESCIMENTO E PRODUÇÃO	8
3.1. Modelos de Clareira (<i>GAP Models</i>).....	9
3.2. Modelos Mecanísticos (<i>Process-Based Models</i>)	9
3.3. Modelos Empíricos	10
3.3.1. Povoamento Total (Tabelas de Produção e Densidade Variável)	10
3.3.2. Modelo de Distribuição de Diâmetros	12
3.3.3. Modelo de árvores individuais	12
4. DADOS E TÉCNICAS PARA MODELAGEM DO CRESCIMENTO E PRODUÇÃO ..	13
5. OUTRAS TÉCNICAS PARA MODELAGEM DO CRESCIMENTO E PRODUÇÃO FLORESTAL.....	16
5.1. Redes Neurais Artificiais	16
5.2. Máquina de Vetores Suporte.....	17
5.2.1. Teoria da MVS na Aproximação por Regressão.....	21
6. REGULAÇÃO FLORESTAL	23
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	25
CAPÍTULO 2: TÉCNICAS PARA PROJEÇÃO DA PRODUÇÃO FLORESTAL	32
RESUMO.....	32
1. INTRODUÇÃO	33
2. HIPÓTESES	34
3. METODOLOGIA	35
3.1. Descrição dos dados.....	35
3.2. Classificação da capacidade produtiva	37
3.3. Modelagem do crescimento e da produção.....	37
3.3.1. Modelo de Crescimento	37
3.3.2. Máquina de Vetores Suporte (MVS).....	38

3.3.3. Redes Neurais Artificiais (RNA)	39
3.4. Avaliação e aplicação das alternativas de modelagem	40
4. RESULTADOS E DISCUSSÃO	42
5. CONCLUSÕES	48
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	49
CAPÍTULO 3: EFEITOS DE DIFERENTES ALTERNATIVAS DE MODELAGEM DO	
CRESCIMENTO E DA PRODUÇÃO SOBRE A REGULAÇÃO FLORESTAL	
RESUMO	52
1. INTRODUÇÃO	53
2. HIPÓTESES	54
3. METODOLOGIA	55
3.1. Estudo de caso.....	55
3.2. Formulação e Resolução do Modelo de Regulação	55
3.2.1. Função Objetivo.....	56
3.2.2. Restrições.....	57
4. RESULTADOS E DISCUSSÃO	59
5. CONCLUSÕES	64
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	65

RESUMO

MAGALHÃES, Juliana Galvão de Sousa, M. Sc., Universidade Federal de Viçosa, Fevereiro de 2014. **Prognose e Regulação de Florestas de eucalipto utilizando Modelo de Crescimento e Inteligência Computacional.** Orientador: Márcio Lopes da Silva.

O intuito deste estudo foi avaliar o desempenho da técnica de inteligência artificial, Máquinas de Vetores Suporte (MVS), para a projeção da produção de povoamentos equiâneos, em relação ao emprego de um modelo de densidade variável e de uma rede neural artificial (RNA). Ademais, comparou-se o efeito destas metodologias na regulação da produção em uma fazenda modelo. A estrutura da dissertação é organizada em três capítulos. O capítulo 1 consiste na revisão bibliográfica dos diversos métodos de modelagem utilizados, e de uma breve abordagem sobre a influência destes e dos dados florestais para o estudo do crescimento e produção. Fundamenta-se também a proposta de utilização das técnicas de inteligência artificial como alternativa eficiente para modelagem florestal, abrangendo também o conceito de Máquina de Vetores Suporte e desta teoria na aproximação por regressão. No capítulo 2, diferentes Máquinas de Vetores Suporte são construídas para prognose da produção, sendo esses resultados comparados aos do modelo de crescimento e das Redes Neurais Artificiais. No capítulo 3, cada alternativa foi utilizada para elaborar tabelas de produção, as quais serviram como entrada (*inputs*) em um modelo de programação linear (PL) formulado com intuito de regular a produção de uma fazenda modelo.

ABSTRACT

MAGALHÃES, Juliana Galvão de Sousa, M. Sc., Universidade Federal de Viçosa, February, 2014. **Prognosis and Regulation of eucalyptus forest using Growth Model and Computational Intelligence.** Adviser: Márcio Lopes da Silva.

The purpose of this study is to evaluate the performance of an Artificial Intelligence technique, Support Vector Machines (SVM), on projecting the production of eucalyptus stands. Furthermore, it aimed to analyze the effects of these methodologies on regulating the production in a model farm. This dissertation is organized into three chapters: Chapter 1 consists of a forest modeling review, and a brief overview of these techniques and forest data influence on modeling the growth and yield. It also proposes the use of artificial intelligence techniques as an efficient alternative for forest modeling, covering the concept of Support Vector Machine and Least Square Support Vector Machine. In Chapter 2, different Support Vector Machines were built for eucalyptus production prognosis, and these results were compared to those by growth model and Artificial Neural Networks. In chapter 3, each alternative has been used to develop production tables, which served as input into a linear programming model (LP) formulated to regulate the production of a model farm.

1. INTRODUÇÃO

O setor florestal apresenta algumas características especiais que o diferem dos demais. A dependência com o meio ambiente, por exemplo, induz sérios riscos a produção, e é uma preocupação constante aos profissionais do setor florestal. Além disso, a produção em grande escala com retorno em longo prazo aumenta a responsabilidade do administrador ao tomar decisões (SILVA et al., 2005). Sendo assim, diferentes áreas da ciência florestal como economia, mensuração, silvicultura e, principalmente, o manejo devem proporcionar ao gestor ferramentas que o assegurem na tomada destas decisões (BUONGIORNO e GILLESS, 2003).

Diante da necessidade de informações mais precisas, o estudo dos processos e fatores que influenciam o crescimento e a produção de uma floresta torna-se um instrumento essencial para o seu gerenciamento. Além de sua importância na administração das florestas, o conhecimento da produção futura permite a ordenação, ou melhor, a regulação do uso dos recursos florestais.

Entretanto, o crescimento compreende um sistema complexo, uma vez que acontecem mudanças na magnitude de diversas características mensuráveis nas árvores, como diâmetro, altura, volume, peso e biomassa (ZEIDE, 1993). Ademais a influência e interligação de variáveis biológicas, econômicas e sociais também comprovam a complexa dinâmica do setor florestal (ARCE et al., 1998).

Todavia, uma maneira de representar este sistema é através dos modelos estatísticos e, ou matemáticos, devido a possibilidade de simplificação do sistema biológico (BUONGIORNO e GILLESS, 2003). A utilização de modelos de crescimento e produção, por exemplo, servem para estimar (predizer ou projetar) a produção em função de variáveis quantificáveis na floresta. Ressalta-se que um bom modelo de crescimento e produção necessita ser flexível o suficiente para permitir que um amplo número de alternativas de manejo possa ser simulado (VANCLAY, 1995). Além disso, é necessária uma visão holística para obter uma solução ótima do uso dos recursos florestais. Por isto que no manejo e no planejamento, uma das principais considerações é como realizar a prognose e a projeção de produções futuras a partir de informações correntes.

A escolha da melhor abordagem de modelagem é um desafio para o gestor florestal, haja vista sua implicação econômica. Informações precisas com relação ao

volume produzido e seu valor promovem uma auto-suficiência do setor. Vários modelos de sistemas de gestão dos recursos florestais combinam os métodos de pesquisa operacional e os de economia; pois é necessária para comparar os custos e os benefícios de abordagens alternativas (BUONGIORNO e GILLESS, 2003).

A adequabilidade de um modelo de crescimento e produção florestal deve ser realizada com base na sua aplicação e nos dados disponíveis; abrangendo também a facilidade de uso, a precisão das previsões e o realismo biológico (YOUNG e GIESE, 2003). Os modelos de crescimento e produção existentes empregam em sua maioria a técnica da análise de regressão (linear e não linear) com variáveis mensuradas nos povoamentos (diâmetro, altura e idade), considerando o espaçamento, a espécie ou clone e a capacidade produtiva para a estratificação do povoamento (CAMPOS E LEITE, 2013).

No entanto, o emprego deste método requer algumas condições específicas. Dependendo dos dados utilizados, alguns termos no sistema de equações referentes ao modelo de Clutter (CLUTTER, 1963), por exemplo, podem ser modificados ou até mesmo eliminados, devido a não significância, ou seja, inconsistência dos parâmetros (GARCÍA, 1988; VANCLAY, 1995; OLIVEIRA, 2012). Além disso, dependendo do modelo de regressão escolhido, a necessidade de estratificação do povoamento, e principalmente quando esta é muito intensa, pode proporcionar uma ineficiência ao modelo por conta da limitação de dados, (CAMPOS E LEITE, 2013).

Conquanto, o advento do computador causou uma revolução no processamento de dados, e tornou-se uma indispensável ferramenta de auxílio na modelagem. Atualmente, pode-se facilmente determinar a melhor solução para problemas com diversas variáveis e muitas restrições, sendo estes bastante comuns no setor florestal. A evolução das ferramentas computacionais possibilita também que novos processos matemáticos sejam incorporados na determinação dos modelos.

No setor florestal, a utilização de novos métodos computacionais, especialmente na etapa de planejamento, é impulsionada pela necessidade de se obter melhor produtividade e rentabilidade, sem desconsiderar a sustentabilidade da floresta. Recentemente, técnicas de inteligência artificial estão sendo incorporadas ao estudo da modelagem florestal, com intuito de fornecer estimativas mais precisas e alternativas de redução do custo (McQUEEN et al., 1995; RECKNAGEL, 2001; GORGENS, 2006;

SILVA et al., 2009; BINOTI, 2010; KING et al., 2000; CASTRO, 2011; BINOTI, 2012; CRISCI et al., 2012; ZHENG e LU, 2012).

Diante disso, observa-se que a busca pela otimização, isto é, não apenas uma solução, mas a melhor solução em menos tempo, é um tema recorrente para pesquisa florestal (OLIVEIRA, 1995; ARCE et al., 1998; KING et al., 2000; SILVA et al., 2009; BINOTI, 2010; CASTRO, 2011; BINOTI, 2012; OLIVEIRA, 2012). Sendo assim, o objeto de estudo desta dissertação é testar uma nova ferramenta para projeção da produção de um plantio florestal equiâneo, analisando a influência de seus resultados na regulação. Isto permitirá a ampliação do número de alternativas e melhoria da precisão das estimativas, tornando a tomada de decisão do gestor florestal mais eficaz. Ademais, cita-se o avanço nos estudos de modelagem de dados do setor florestal brasileiro.

2. REFERÊNCIAS BIBLIOGRÁFICAS

ARCE, J. E.; GOMES, F. dos S.; SANQUETTA, C. R.; CRUZ, E. C. da; Utilização dos processos de difusão na avaliação e simulação precoce do crescimento de povoamentos de *Pinus taeda* L. **Revista Cerne**, v.4, n.1, p.154-170, 1998.

BINOTI, M. L. M. S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. Viçosa: UFV, 2010. 54 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.

BINOTI, D. H. B. **Sistemas computacionais aplicados ao manejo florestal**. Viçosa: UFV, 2012. 127p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

BUONGIORNO, J.; GILLESS, J. K. **Decision Methods for Forest Resource Management**. Elsevier, 2003. 458p.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. Viçosa: Editora UFV, 2013. 605 p.

CASTRO, R. V. O. **Modelagem do crescimento em nível de árvores individuais utilizando redes neurais e autômatos celulares**. Viçosa: UFV, 2011. 90p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2011.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**, v. 9, n. 3, p. 354-371, 1963.

CRISCI, C.; GHATTASB, B.; PERERAC, G. A review of supervised machine learning algorithms and their applications to ecological data. **Ecological Modelling**, n.240, p.113– 122, 2012.

GARCÍA, O. Growth modeling: a (re)view. **New Zealand Forestry**, v.33, n.3, p.14-17, 1988.

GORGENS, E. B. **Estimação do volume de árvores utilizando Redes Neurais Artificiais**. Viçosa: UFV, 2006. 94p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2006.

KING, S. L.; BENNETT, K. P.; LIST, S. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. **Computers and Electronics in Agriculture**, n. 27, p.401–406, 2000.

McQUEEN, R. J.; GAMER, S. R.; NEVILL-MANNING, C. G.; WITTEN, I. H. Applying machine learning to agricultural data. **Computers and Electronics in Agriculture**, n. 12, p. 275-293, 1995.

OLIVEIRA, E. B. de. **Um sistema computadorizado de prognose do crescimento e produção de *Pinus taeda* L., com critérios quantitativos para a avaliação técnica e**

econômica de regimes de manejo. Curitiba: UFPR, 1995. 126p. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia Florestal, Universidade Federal do Paraná, Curitiba, 1995.

OLIVEIRA, J. de C. **Mínimos quadrados generalizados: uma proposta à modelagem do crescimento e produção em povoamentos não desbastados de eucalipto** Viçosa: UFV, 2012. 80p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

RECKNAGEL, F. Applications of machine learning to ecological modeling. **Ecological Modelling**, n. 146, p. 303-310, 2001.

SILVA, M. L.; JACOVINE, L. A. G.; VALVERDE, S. R. **Economia Florestal**, Viçosa: UFV, 2005. 178p.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de Redes Neurais Artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, v.33, n.6, p.1133-1139, 2009.

VANCLAY, J. K. Growth models for tropical forests: a synthesis of models and methods. **Forest Science**, v.41, p. 7-42, 1995.

ZEIDE, B. Analysis of Growth Equations. **Forest Science**, v. 39, n. 3, p. 594-616, 1993.

ZHENG, H.; LU, H. A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). **Computers and Electronics in Agriculture**, n.83, p. 47-51, 2012.

YOUNG, R. A.; GIESE, R. L. **Introduction to Forest Ecosystem Science and Management**. 2003, 589 p.

CAPÍTULO 1: REVISÃO BIBLIOGRÁFICA

1. Crescimento e Produção de Povoamentos Equiâneos

O crescimento de uma árvore pode ser entendido como o resultado da interação de um componente positivo, denominado potencial biótico, o qual predomina no início da vida manifestando a expansão do organismo e sua tendência natural de multiplicação; e outro negativo, que representa as restrições impostas por fatores externos (competição, estresse hídrico e respiração) e internos (mecanismos auto-regulatórios e envelhecimento), os quais determinarão o declínio do crescimento (ZEIDE, 1993).

O padrão de crescimento é impreciso e não deve ser analisado como uma relação linear. O seu declínio parece ser mais variável devido aos inúmeros fatores que o afetam, e pode ser obtido com precisão por várias expressões. Logo, é importante que os estudos sobre o crescimento florestal foquem nas características desses dois componentes: expansão e declínio (ARCE et al., 1998; BETTINGER et al., 2009).

O estudo e compreensão dos processos e fatores que influenciam o crescimento e a produção de povoamentos são essenciais para o gerenciamento da floresta; uma vez que o manejo florestal é decidido em grande parte pela prognose de produções futuras a partir de informações correntes (CAMPOS e LEITE, 2013). Além de sua importância no gerenciamento das florestas, essas informações permitem a ordenação do uso dos recursos florestais por meio da regulação.

Porém, os problemas envolvidos na gestão dos recursos florestais abrangem diversas variáveis, sendo que muitas vezes a variação numa delas pode influenciar as demais. Algumas variáveis são biológicas, como o potencial de crescimento das espécies; outras são econômicas, como o preço da madeira e o custo de mão de obra; outras, ainda, são sociais, como as legislações ambientais e trabalhistas. Todas estas variáveis e as relações entre elas constituem um sistema bastante complexo (BUONGIORNO et al., 2003).

2. Modelagem do Crescimento e da Produção Florestal

No processo de pensar sobre o problema, ponderar alternativas e tomar decisões, os gestores florestais utilizam os modelos. Modelos são representações abstratas (resumidas) do que é o problema na vida real; sendo extremamente úteis para fins de raciocínio, previsão e tomada de decisão (BUONGIORNO e GILLESS, 2003).

Torna-se mais fácil tomar uma decisão quando se analisa o problema por um tipo de modelo, mesmo que informal. Dentre os diversos tipos, os modelos matemáticos são os mais usados e muito poderosos. Sua eficiência decorre da utilização de uma linguagem mais concisa e menos ambígua do que a linguagem natural (VANCLAY, 1995). Neles, a realidade é captada por variáveis simbólicas e algébricas, abstraindo a analogia visual encontrada, por exemplo, nos mapas (BUONGIORNO e GILLESS, 2003).

Os modelos matemáticos não são novos na área florestal. Existem aqueles utilizados para expressar as relações biométricas entre o volume por unidade de área com a idade e a qualidade do local; além das fórmulas há muito tempo desenvolvidas para calcular o valor da terra em função da produção esperada, os preços dos produtos florestais (madeireiros e não madeireiros), os custos de gestão e as taxas de juros (BUONGIORNO e GILLESS, 2003). Estes modelos de investimento são fundamentais na tomada de decisão dos recursos florestais.

Com relação a aplicação da modelagem matemática no setor florestal, as duas principais são: a focada na predição e projeção do crescimento e da produção florestal, inclusive aquela que analisa os efeitos de práticas silviculturais e outra que se destina a estudar a dinâmica, a sucessão da floresta e os efeitos de fatores ambientais sobre a estrutura do povoamento (PORTÉ e BARTELINK, 2002).

A evolução do estudo de crescimento e produção foi iniciada no século XVIII, quando já se trabalhava com tabelas de produção, as quais forneciam o volume médio dos povoamentos em várias idades. Os avanços em estudos de novas metodologias estatísticas, como técnicas de regressão múltipla, possibilitaram a elaboração de tabelas de produção envolvendo variáveis como densidade do povoamento, idade e índice de local (SULLIVAN e CLUTTER, 1972).

Os trabalhos de Buckman em 1962 e Clutter em 1963 abriram uma nova fase nos estudos de modelagem do crescimento e produção. A introdução do conceito de modelos compatíveis propunha que o modelo de produção fosse obtido pela integração

do modelo de crescimento, reduzindo assim, os erros das estimativas. Os modelos baseados em teorias biológicas do crescimento, como o modelo de Chapman-Richards, marcaram outro passo importante nos estudos de modelagem de dados florestais (OLIVEIRA, 1995).

Atualmente, a evolução das ferramentas computacionais possibilita que novos processos matemáticos sejam incorporados na determinação dos coeficientes dos modelos. Os sistemas de prognose do crescimento e produção, por exemplo, tornaram-se integrados, o que demonstra as infinitas possibilidades de aplicação de técnicas mais refinadas nas atividades de planejamento florestal (BINOTI, 2012).

3. Tipos de Modelo de Crescimento e Produção

Independentemente do seu detalhe, um modelo pode ser determinístico ou estocástico. Um modelo de crescimento determinístico fornece uma estimativa do crescimento esperado de um povoamento florestal; por exemplo, dadas as mesmas condições iniciais, um modelo determinístico sempre prevê o mesmo resultado. No entanto, devido a variação natural no ambiente, povoamentos florestais reais não crescem exatamente a mesma quantidade a cada ano.

Entretanto, o modelo estocástico tentará ilustrar esta variação natural fornecendo previsões diferentes, cada uma com uma probabilidade específica de ocorrência (VANCLAY, 1995). Embora os modelos estocásticos possam fornecer algumas informações úteis que não são disponíveis a partir de modelos deterministas, a maior parte das informações necessárias para os planejadores e gestores florestais é obtida de forma mais eficiente com modelos determinísticos (PORTÉ e BARTELINK, 2002).

Muitos dos modelos de crescimento e produção existentes permitem ao usuário simular diferentes alternativas de manejo e compreender os seus impactos sobre a estrutura da floresta. Estes modelos são representados por três tipos: modelos de clareira (*Gap Models*), que estudam a dinâmica de regeneração em clareiras; modelos mecanísticos ou ecofisiológicos (*Process-based Models*), os quais simulam como os fatores (CO₂, água, evapotranspiração, precipitação) são convertidos em biomassa; e os modelos empíricos, onde são utilizados dados observados nas florestas para os ajustes (PORTÉ e BARTELINK, 2002).

3.1. Modelos de Clareira (*GAP Models*)

Nos modelos de clareira, o estabelecimento, crescimento e mortalidade dos indivíduos são simulados em função de fatores bióticos (competição) e fatores abióticos (clima e solo). São bem similares aos modelos de árvore individual, uma vez que também utilizam a árvore como unidade básica. Cada árvore é representada espacialmente no modelo de acordo com a área de ocupação da copa. Incluem elementos estocásticos e necessitam serem calculados múltiplas vezes para que seja desenvolvido um padrão de comportamento do crescimento da floresta (PORTÉ e BARTELINK, 2002).

A complexidade da floresta continua a ser um grande desafio para os que estão envolvidos no conhecimento dos processos de sucessão. Uma abordagem puramente estatística para a modelagem da sucessão florestal é inapropriada; ao mesmo tempo em que é impossível modelar o comportamento dos ecossistemas considerando apenas aspectos da fisiologia vegetal. Logo, é necessário empregar uma combinação de abordagens 'estatísticas' e 'mecanísticas', para assim, fazer projeções da dinâmica florestal em longo prazo, prever fenômenos e processos naturais (BUGMANN, 2001).

3.2. Modelos Mecanísticos (*Process-Based Models*)

Um modelo mecanístico, também conhecido como “modelo de processo”, tem a capacidade de imitar os processos físicos, químicos ou biológicos relevantes e descrever o como e o porquê de um determinado resultado (VANCLAY, 1995). Os modelos de processos tentam modelar o processo de crescimento, considerando como entrada a luz, a temperatura e os níveis de nutrientes no solo. Esses modelos, se devidamente parametrizados, calibrados e validados, fornecem estimativas confiáveis com capacidade preditiva espacial e temporal (BORGES, 2009).

O potencial de um modelo mecanístico foi descrito num estudo que se baseou nas relações entre as variáveis climáticas temperatura média, precipitação pluvial e radiação solar para prever a produtividade em cultura de cafeeiro (SANTOS, 2002). Outro bom exemplo de utilização deste tipo de modelo foi publicado por TURAK et al. (2011). Neste estudo, utilizaram-se modelos baseados em processos para estimar a condição da biodiversidade em um rio sob cenários alternativos de gestão na Região Hunter localizada no sudeste da Austrália. A modelagem relacionou dados de sobrevivência de macroinvertebrados a perturbações induzidas pelo homem em locais do rio. Esta

abordagem mostrou-se adequada para orientar a implementação integrada de três estratégias de proteção para as águas doces da região.

No setor florestal, BORGES (2009) simulou o crescimento de florestas de eucalipto na região do cerrado em Minas Gerais através do modelo 3-PG, o qual apresentou desempenho superior ao de outras parametrizações do modelo disponíveis. Este mesmo modelo também foi testado por GUIMARÃES et al. (2007) em florestas de eucalipto e novamente mostrou-se uma ferramenta adequada para estimar o potencial produtivo desta espécie em escala regional.

3.3. Modelos Empíricos

Os modelos empíricos não têm como objetivo o entendimento da relação de causa-efeito entre as variáveis de um sistema, mas sim nas relações quantitativas entre as mesmas e nos fatores considerados (PORTÉ e BARTELINK, 2002). São representados por equações cujos coeficientes são válidos apenas para as condições específicas em que foram estabelecidos; por isto são extremamente dependentes (ou influenciados) pelo banco de dados usado para o ajuste (CAMPOS e LEITE, 2013).

É inadequado utilizar o mesmo modelo empírico em diferentes regiões, devido às inúmeras combinações de espécies florestais e dos diferentes regimes de manejo. Porém, a adição de relações mecanísticas proporcionará melhora na qualidade desses modelos. Isto porque sua confiabilidade depende também do conhecimento dos processos fisiológicos e suas respostas nas condições de crescimento das espécies, do que somente no ajuste estatístico de alguns dados empíricos (PORTÉ e BARTELINK, 2002). Os modelos empíricos aplicados no setor florestal são subdivididos em três categorias: povoamento total (tabelas de produção e densidade variável), distribuição por classes de diâmetro e em nível de árvore individuais (DAVIS e JOHNSON, 1987).

3.3.1. Povoamento Total (Tabelas de Produção e Densidade Variável)

Os modelos de crescimento e produção em nível de povoamento, também denominados modelos do tipo povoamento total, estimam o crescimento e a produção a partir de atributos como idade, área basal e índice de local (CAMPOS e LEITE, 2013). Inicialmente, desenvolveram-se as tabelas de crescimento e produção com intuito de auxiliar a compreensão do comportamento das florestas e fornecer informações do povoamento sob condições específicas de manejo (VANCLAY, 1995). Em seguida,

surgiram os modelos de densidade variável, os quais incluem a projeção da densidade, pois esta variável é considerada como uma parte dinâmica do sistema de equações (CAMPOS e LEITE, 2013).

No modelo proposto por Clutter em 1963, a produção é derivada pela integração matemática do modelo de crescimento, o que lhe confere uma característica essencial: compatibilidade. As principais relações funcionais podem ser expressas em duas formas do modelo: completa ou original (1) e usual (2). Apesar desse sistema de equações não ser aplicado a todas as espécies e tipos de manejo, o modelo usual é o mais difundido nas empresas florestais.

$$\begin{aligned} \ln B_2 &= \ln B_1 \left(\frac{I_1}{I_2} \right) + \alpha_0 \left(1 - \frac{I_1}{I_2} \right) + \alpha_1 \left(1 - \frac{I_1}{I_2} \right) S_1 + \ln \varepsilon \\ \ln V_2 &= \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 \ln B_2 + \beta_4 S_1 \ln B_2 + \beta_5 S_1 I_2^{-1} + \beta_6 \ln B_2 I_2^{-1} + \ln \varepsilon \quad (1) \\ \ln V_2 &= \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 \ln B_2 + \ln \varepsilon \quad (2) \end{aligned}$$

em que: I_1 e I_2 = idade atual e futura, respectivamente; V_1 e V_2 = volumes nas idades I_1 e I_2 , respectivamente; S_1 = índice de local na idade atual I_1 ; e B_1 = área basal atual.

Contudo, alguns trabalhos reportam a necessidade do modelo completo ser aplicado. SALLES (2010) utilizou dados de clones de eucalipto implantados em um sistema agroflorestal para testar a contribuição de todas as variáveis do modelo de Clutter. O autor definiu sistemas de equações específicos para cada clone e concluiu sobre a relevância em se utilizar este modelo na sua forma original (completa), porém analisando-se a contribuição estatística de cada variável. ALCÂNTARA (2012) comparou diferentes abordagens para modelagem em nível de povoamento com o modelo de Clutter e também concluiu sobre a necessidade de avaliar a contribuição estatística das diferentes variáveis.

O modelo de Clutter é considerado como do tipo povoamento total porque, independentemente da classe de tamanho da árvore, a variável estimada é o volume por unidade de área. É também um modelo explícito porque a produção em volume é calculada diretamente; e consistente, pois as estimativas podem ser obtidas projetando-se a área basal ano a ano, ou diretamente de um para qualquer outro ano, com intervalos irregulares. Além disso, ao estimar a produção para uma determinada idade, o resultado terá valores semelhantes àqueles observados em campo (CAMPOS e LEITE, 2013).

Diversos resultados deste modelo já foram publicados nacionalmente e atestaram sua eficiência para, por exemplo, prever a produção atual e futura de povoamentos

desbastados (DIAS et al., 2005) e não desbastados (SANTANA et al., 2005) de eucalipto; e também na determinação de sua rotação técnica e econômica (RESENDE et al., 2004). NAKASHIZUKA e KOHYAMA (1995) estudaram a dinâmica a longo prazo de diferentes espécies em um povoamento utilizando um modelo de distribuição de densidade variável.

3.3.2. Modelo de Distribuição de Diâmetros

A distribuição em classe de diâmetros possibilita descrever a estrutura do povoamento (MALTAMO et al., 2000); ademais a análise desta distribuição permite estimar o número de árvores por hectare de cada classe de diâmetro, assim como determinar a altura média nessas classes (RETSLAFF et al., 2012). Esses dados, quando usados em conjunto com uma equação de volume, estimam a produção por classe de diâmetro e por unidade de área (CLUTTER et al., 1983).

A principal característica desta técnica está na separação das árvores com mesmas características de fuste em classes diamétricas, diminuindo assim, a amplitude volumétrica entre essas classes, o que permite obter coeficientes de melhor precisão para cada classe e mais confiabilidade ao modelo (SANTOS et al., 2006). Para realizar tal função é empregada uma função de densidade probabilidade (*f_{dp}*), sendo a mais comumente utilizada a função Weibull (CAMPOS e LEITE, 2013).

A projeção da distribuição diamétrica e estimativas de volume de madeira em diferentes classes fornecem uma ótima base quantitativa para a avaliação econômica e o gerenciamento de diferentes opções de manejo (TRINCADO et al., 2003). Os modelos de distribuição diamétrica são também importantes para estudos sobre multiprodutos de madeira (RENNOLLS et al., 1985; SOUZA et al., 2008; CAMPUS et al., 2013).

No Brasil, estudos sobre o crescimento utilizando este tipo de modelo abrangem povoamentos equiâneos (CUNHA NETO et al., 1994; NOGUEIRA, 2003; LEITE, et al., 2005; RETSLAFF et al., 2012) e inequiâneos (UMAÑA e ALENCAR, 1998; BARTOSZECK et al., 2004; SANTOS et al., 2006; MACHADO et al., 2009) implantados em diversas regiões.

3.3.3. Modelo de Árvores Individuais

A abordagem mais pormenorizada dos modelos é a que utiliza a árvore como unidade básica de modelagem. A entrada mínima necessária é uma lista especificando o

tamanho de cada árvore no povoamento. As informações por árvores são atualizadas para uma idade definida e, então, agregadas por área (CAMPOS e LEITE, 2013). Nas tabelas de produção usuais, a média das dimensões e totais por unidade de área será o somatório dos incrementos das árvores singulares.

Os modelos de árvores individuais são mais complexos do que os modelos em nível de povoamento e de distribuição diamétrica, pois são constituídos por três sub-modelos: mortalidade, regeneração ou ingresso e crescimento dimensional (altura e diâmetro) (HASENAUER, 2000). Alguns também necessitam de informações sobre a posição espacial da árvore, a altura e a classe de copa (VANCLAY, 1995).

O primeiro sub-modelo é o da mortalidade, o qual é simulado pela aplicação de uma função de probabilidade no período da projeção. Assume-se que esta variável possa ser causada por fatores não catastróficos, também denominados de regular, e que podem ser mensurados; ou ainda, pela consequência do ataque de pragas, incêndios, danos mecânicos, entre outros fatores difíceis de prever e mensurar. Já o recrutamento ou ingresso corresponde ao número de indivíduos vivos computados numa medição e não medidos em idades anteriores por não terem alcançado um tamanho (valor) mínimo (YOUNG e GIESE, 2003).

O último sub-modelo (crescimento dimensional) pode ser calculado por uma função modificadora do crescimento potencial, a qual utiliza índices de competição e variáveis do povoamento para deduzir o crescimento das árvores submetidas à competição (VANCLAY, 1994). Esses índices subdividem os modelos de árvore individual em duas categorias, *distância dependente* e *distância independente*, de acordo com a distância e o porte das árvores.

Os modelos de árvore individual podem ser utilizados tanto em povoamento equiâneos quanto em inequiâneos. Alguns têm demonstrado a eficiência deste tipo de modelo em manejo de eucalipto (FERRARI et al., 2005; MENDES et al., 2006; CASTRO, 2011; MARTINS, 2011) e em povoamentos naturais de araucária (NUTTO, 2001; CHASSOT et al., 2011). Contudo, uma desvantagem ao empregar o modelo de árvore individual é o seu alto custo computacional, especialmente quando se deseja simular processos fisiológicos (PORTÉ e BARTELINK, 2002).

4. Dados e Técnicas para Modelagem do Crescimento e Produção

A eficiência das previsões nos modelos de crescimento e produção depende de muitos fatores, sendo o principal deles a qualidade dos dados utilizados na calibração.

Considerando a confiabilidade destes dados um fator crucial, e com o propósito de reduzir os efeitos dos erros de amostragem nas estimativas, é essencial o controle das informações que servirão de entrada para os modelos. Por conseguinte, os inventários contínuos realizados em parcelas permanentes são as melhores fontes de dados para a modelagem do crescimento e da produção florestal (YOUNG e GIESE, 2003). Os dados disponíveis pelo inventário influenciarão principalmente na escolha do modelo ajustado. Além disso, a decisão deve ser baseada nos objetivos do manejo e nos resultados da avaliação estatística do modelo (CAMPOS e LEITE, 2013). Do ponto de vista da mensuração, por exemplo, sempre será escolhido o modelo que resultar em maior exatidão e ausência de *bias*, ou seja, tendência (SOARES et al., 2006).

Observa-se que diferentes alternativas de modelagem podem ser empregadas para entender o comportamento do crescimento e sua influência na produção florestal. Independente da categoria, a maioria dos modelos supracitados baseiam-se no ajuste de regressão (linear e não-linear) dos dados disponíveis (CAMPOS e LEITE, 2013). Modelos de regressão são modelos matemáticos que estudam o comportamento, ou melhor, a relação entre duas variáveis:

$$Y = \beta_1 + \beta_2 X + \epsilon$$

Neste caso, a variável X é considerada como independente na equação; enquanto que Y é a variável dependente das variações de X . β_1 e β_2 são os coeficientes da regressão, ou seja, parâmetros. ϵ é conhecido como termo de erro estocástico; em outras palavras, representa um componente aleatório que gera diferença entre os valores estimados (\hat{y}) e os observados de Y (y) que geralmente é proveniente de uma imprecisão, ou melhor, erros de medida, indisponibilidade dos dados e casualidades (GUJARATI, 2006).

A compreensão da relação entre as duas variáveis serve para realizar previsões sobre o comportamento futuro de algum fenômeno. Neste caso, extrapola-se para o futuro as relações de causa-efeito já observadas no passado entre as variáveis X e Y . Outra utilização do modelo de regressão é simular os efeitos sobre uma variável Y em decorrência de alterações introduzidas nos valores de uma variável X .

Independente da finalidade, a construção de qualquer modelo envolve duas etapas essenciais: a estrutura e os parâmetros. A estrutura é a representação da relação existente entre as variáveis numa forma funcional, podendo ser exponencial, quadrática, entre outras. Já os parâmetros consistem nos números que tornam possível a previsão

quantitativa das relações representadas pela estrutura funcional do modelo (SCHWAAB e PINTO, 2007). No exemplo supracitado, pode-se dizer que as duas variáveis X e Y estão relacionadas linearmente devido a estrutura do modelo. No entanto, para que o modelo possa ser utilizado é necessário definir quais valores representarão os parâmetros β_1 e β_2 .

A estimativa dos parâmetros ocorre geralmente pelo método dos mínimos quadrados, o qual consiste em minimizar a soma do quadrado dos resíduos (ε):

$$\sum \varepsilon_i^2 = \sum (y_i - \hat{y}_i)^2$$

O emprego do modelo de regressão na estimativa dos parâmetros por mínimos quadrados requer algumas condições específicas, tais como: a linearidade dos parâmetros, homocedasticidade da variância dos resíduos, resíduo com média zero, ausência de autocorrelação residual e número de observações maior que o número de parâmetros (GUJARATI, 2006).

Entretanto, os dados oriundos dos atuais povoamentos florestais apresentam características que violam algumas pressuposições estatísticas da distribuição dos resíduos (GARCÍA, 1988). Como exemplo, podem-se citar a ocorrência de heterocedasticidade da variância e autocorrelação residual, que estão associadas às medições repetidas das variáveis ao longo do tempo e à idade dos povoamentos (CALEGARIO et al., 2005; OLIVEIRA, 2012). Ademais, para estimar os parâmetros pelo modelo de regressão é necessário no mínimo duas medições nas parcelas. Porém, dependendo desses dados, alguns termos (variáveis) no sistema de equações podem ser modificados, ou até mesmo eliminados, devido a inconsistências dos parâmetros. Segundo VANCLAY (1995), dados limitados, mas de confiança e que cobrem os extremos, são mais úteis do que dados copiosos agrupados em torno da média.

Outro grande problema na modelagem de dados florestais é que no processo de ajuste de modelos há a necessidade de estratificação do povoamento a qual é geralmente realizada utilizando-se a capacidade produtiva do lugar, além de outras características edáfica ou climática. Existem duas opções: estratificar muito e usar modelos do tipo $V = f(I)$, ou estratificar pouco e optar por modelos mais abrangentes, como o de densidade variável. Contudo, ambas as opções também podem proporcionar uma ineficiência ao modelo de regressão, se houver limitação na representatividade dos dados (CAMPOS e LEITE, 2013).

5. Outras Técnicas para Modelagem do Crescimento e Produção Florestal

Uma alternativa à modelagem do crescimento e da produção florestal é a aplicação de Inteligência Artificial (IA), como por exemplo as técnicas de Aprendizado de Máquina: Máquinas de Vetores Suporte (MVS) e Redes Neurais Artificiais (RNA) (McQUEEN et al., 1995; KING et al., 2000; RECKNAGEL, 2001; GORGENS, 2006; SILVA et al., 2009; BINOTI, 2010; ÖZÇELIK et al, 2010; CASTRO, 2011; BINOTI, 2012; CRISCI et al., 2012; ZHENG e LU, 2012; ÖZÇELIK et al, 2013). Estudos na área de mensuração e manejo florestal com aplicação dessas técnicas são incipientes e muito promissores; haja visto poder incluir dados categóricos aos modelos, não ser necessária a estratificação, assim como a superioridade de precisão das estimativas BINOTI (2010).

5.1. Redes Neurais Artificiais

Rede Neural Artificial (RNA) é um sistema de processamento de informações desenvolvido a partir de modelos matemáticos simplificados. A expressão “rede neural” é denominada pela tentativa destes modelos imitarem a capacidade que o cérebro humano possui de reconhecer, associar e generalizar padrões. Uma RNA se assemelha ao cérebro em dois aspectos: o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem e as forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido (HAYKIN, 2001).

Um neurônio artificial, assim como o biológico, compreende a unidade básica de processamento de uma RNA, a qual apresenta duas facetas elementares: a arquitetura e o algoritmo de aprendizagem. Essa divisão surge naturalmente pelo paradigma de como a rede é treinada, por exemplos. Os dados de entrada, também denominados de sinal de entrada ou estímulo, são captados de outro neurônio ou do ambiente, em seguida são processados e assim geram um resultado (HAYKIN, 2001).

Efetivamente, o algoritmo de aprendizagem generaliza esses dados e memoriza o conhecimento dentro dos “parâmetros” adaptáveis da rede, os quais são denominados de pesos. O aprendizado da rede ocorre pelo ajuste desses pesos, pois são neles que fica guardado o conhecimento acerca do problema (BRAGA et al., 2000). Sendo assim, o construtor de um sistema baseado em RNA deve definir o tipo de rede que melhor resolverá o problema em consideração e o algoritmo para treinar a rede, ou seja, para

adaptar os pesos da rede (RAUBER, 2013). Normalmente, o tipo de processamento de um único neurônio é a combinação linear das entradas com os pesos, seguida pela passagem da combinação linear por uma função de ativação (HAYKIN, 2001) (Figura 1).

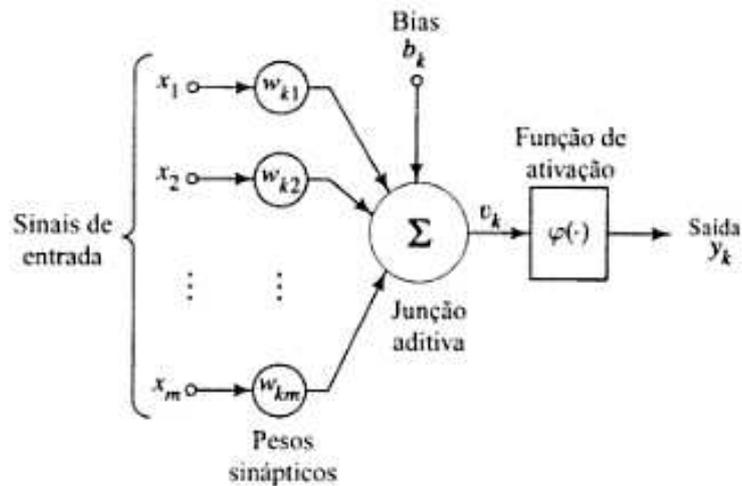


Figura 1 – Modelo de um neurônio artificial (Fonte: HAYKIN, 2001).

Do ponto de vista da modelagem, uma grande vantagem das RNAs sobre os modelos tradicionais está no fato dela poder aproximar uma função sem ser preciso se preocupar com a especificação do modelo. Ademais, são também capazes de aprender relações complexas, mesmo com poucos dados (BRAGA et al., 2000). Por conta disso, RNA é conhecida como uma importante técnica estatística não-paramétrica, capaz de resolver uma gama de problemas complexos, principalmente quando o ambiente dos dados muda muito (VELLASCO, 2013).

As principais áreas de atuação enquadram-se nas tarefas de aprendizagem, podendo-se citar: aproximação de função, classificação de padrões, agrupamento de dados, previsão (séries temporais), otimização, recuperação de conteúdo e controle (HAYKIN, 2001).

5.2. Máquina de Vetores Suporte

As Máquinas de Vetores Suporte – MVS (*Support Vector Machine* – SVM), desenvolvidas inicialmente por Vladimir Vapnik e colaboradores, constituem uma técnica de Aprendizado de Máquina baseada nos princípios da Teoria de Aprendizado Estatístico (TAE). Esta teoria estabelece as condições matemáticas essenciais que

permitem a escolha de um classificador com boa generalização, ou seja, com bom desempenho para os conjuntos de treinamento e teste (SMOLA e SCHOLKOPF, 2004). Os resultados da aplicação desta técnica são comparáveis aos obtidos por outros algoritmos de aprendizado, como as Redes Neurais Artificiais (CORTES e VAPNIK, 1995).

Em essência, MVS é uma entidade matemática, um algoritmo ou receita, que otimiza uma determinada função a partir de uma coleção de dados (NOBLE, 2006). São capazes de resolver problemas de classificação e regressão, adquirindo com o aprendizado a capacidade de generalização (CRISCI et al., 2012). Por exemplo, uma SVM pode aprender a reconhecer um cartão de crédito fraudulento examinando centenas de cartões de crédito fraudulentos e não fraudulentos.

A técnica de Aprendizado de Máquina é um campo de pesquisa da Inteligência Computacional que utiliza o princípio da indução, no qual o conhecimento é obtido a partir de um conjunto de exemplos (Figura 2). Os algoritmos de aprendizado de máquina têm o papel de descobrir o relacionamento entre as variáveis de entrada e de saída de um sistema (CHERKASSKY e MULIER, 2007). Este aprendizado pode ser de dois tipos: supervisionado e não supervisionado.

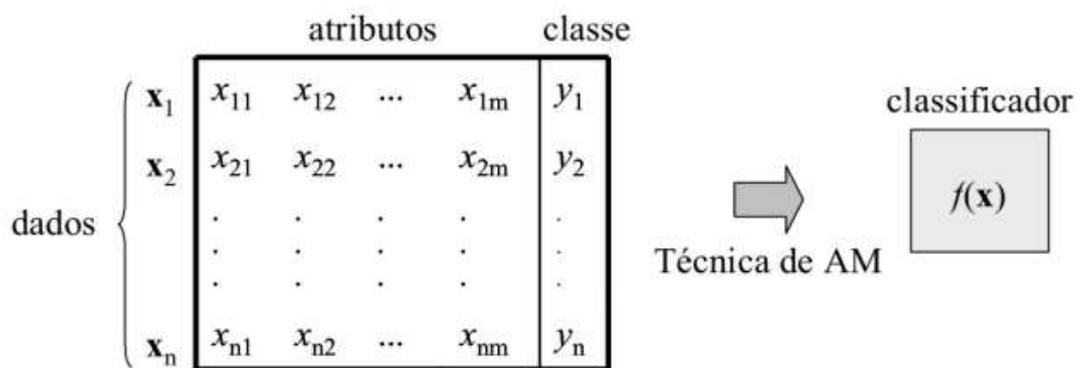


Figura 2 – Ilustração de uma Máquina de Vetores Suporte, evidenciando a técnica de Aprendizado de Máquina através da indução de classificação em um exemplo de aprendizado supervisionado (Fonte: LORENA, 2006).

No aprendizado supervisionado um conjunto de exemplos, denominado modelos, é fornecido para que a máquina possa aprender. Estes exemplos do conjunto de treinamento são compostos por vetores, ou seja, pares de entradas com suas respectivas saídas, de onde o algoritmo extrai o conhecimento necessário para assim poder

generalizar. A generalização consiste em produzir saídas corretas para novas entradas nunca antes apresentadas para a máquina (HAYKIN, 2001).

Já no aprendizado não supervisionado, o algoritmo aprende com a ajuda de uma medida de qualidade, pois não há dados prévios que informem sobre o domínio do problema. Esta técnica geralmente é usada em um conjunto de dados complexos quando o objetivo é encontrar padrões e/ou tendências que auxiliem no entendimento desses dados (SOUTO et al., 2003).

Os pares de entrada, ou exemplos, são também denominados de vetores de características, e podem se assumir como um atributo nominal (categórico) ou contínuo (numérico) (LORENA, 2006). Os atributos categóricos apresentam uma quantidade finita de valores; enquanto que os atributos contínuos assumem uma ordem linear entre seus diferentes valores. No setor florestal, o tipo de solo, clone e projeto podem ser citados como dados categóricos. As medidas mensuráveis, como diâmetro a altura do peito (DAP), volume, altura, entre outras, são consideradas como atributos contínuos.

Eventualmente os atributos contínuos podem ser transformados em atributos categóricos; por exemplo, a declividade é um atributo contínuo se seus valores pertencerem ao conjunto dos números reais, porém esta variável pode se assumir como um atributo categórico se seus valores forem classificados como alto, médio e baixo.

As ideias básicas que fundamentam o algoritmo de MVS podem ser explicadas pela compreensão de quatro conceitos básicos: (i) função *kernel*, (ii) hiperplano de separação ótima, (iii) margem e (iv) vetores suporte (NOBLE, 2006). Na Figura 3, supõe-se que as bolas pertencem ao conjunto dos dados de treinamento: $\{(vermelha_1, azul_1), \dots, (vermelha_n, azul_n)\} \subset X \times R$, onde X indica o espaço dos padrões de entrada.

As funções de *kernel* representam o produto dos dados de entrada mapeados no espaço de característica de dimensão mais elevada. Elas têm a finalidade de projetar os vetores do espaço de entrada, os quais se encontram em um espaço não linearmente separável, para um espaço de características de alta dimensão Z , através de um mapeamento não-linear escolhido a priori para classificação do problema (NOBLE, 2006). Isto é realizado dado que, segundo HAYKIN (2001), a probabilidade desse problema se tornar linearmente separável em um espaço de alta dimensão é maior se comparada a um espaço de baixa dimensão.

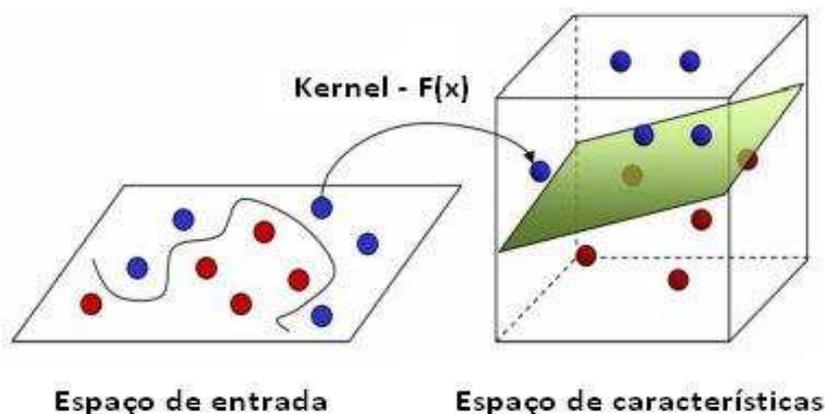


Figura 3 – Ilustração do funcionamento de uma Máquina de Vetores Suporte, evidenciando a função *kernel* e a transformação de um problema não linearmente separável em um problema linearmente separável.
 (Fonte: <http://www.imtech.res.in/raghava/rbpred/svm.jpg>)

Nesse espaço de características com alta dimensão é construída uma superfície de decisão linear, a qual é também definida como hiperplano de separação ótima (máxima). Este hiperplano é constituído por uma função de decisão linear com margem de separação máxima entre os vetores das duas classes, sendo este construído com propriedades especiais que garantem alta capacidade de generalização para a máquina (CORTES e VAPNIK, 1995). Ressalta-se que somente uma pequena quantidade dos dados de treinamento é considerada para a construção do hiperplano, os chamados vetores suporte (FIGURA 4).

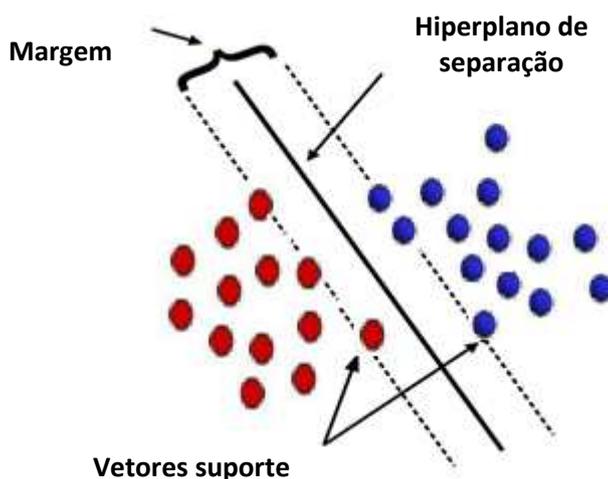


Figura 4 – Ilustração de um hiperplano de separação ótima, evidenciando a margem máxima entre os vetores suporte das duas classes. Adaptado de MEYER (2013).

Uma característica importante para uma boa Máquina de Vetores Suporte é a capacidade de lidar com dados imperfeitos, também chamados de ruídos; e de

minimizar os *outliers*, ou seja, exemplos isolados que destoam dos demais e que podem ocorrer devido a presença de ruídos (LORENA, 2006). A máquina deve idealmente ser robusta aos ruídos e *outliers* presentes nos dados, não obtendo os classificadores sobre estes tipos de informações (CORTES e VAPNIK, 1995).

5.2.1. Teoria da MVS na Aproximação por Regressão

As máquinas de vetor de suporte por mínimos quadrados, ou *Least Square Support Vector Machine* (LS-SVM), baseiam-se nos mesmos fundamentos da SVM, porém utilizam um sistema de equações lineares e uma função custo por mínimos quadrados ao invés da programação quadrática. Isto proporciona uma menor exigência computacional e elevada capacidade de generalização, sendo esta o principal benefício desta metodologia (CRISTIANINI e SHAW-TAYLOR, 2000). O objetivo da máquina na aproximação por regressão é aprender o mapeamento $x_i \rightarrow y_i$, encontrando uma função f que relacione a entrada com a saída ($x \rightarrow f(x)$), sendo que essa aprendizagem é feita com base em dados de treinamento (SMOLA e SCHOLKOPF, 2004).

Considerando um conjunto de n amostras $\{(x_i, y_i), \dots, (x_n, y_n)\} \subset X \times \mathbb{R}$, onde $x_i \in X$, são os dados de entrada e $y_i \in \mathbb{R}$, as classes de saída. Os valores alvo para i variam de 1 até n . Assume-se que os dados são independentes, identicamente distribuídos e que existe alguma distribuição de probabilidade desconhecida $P(x, y)$ de onde os dados são obtidos (DIAS, 2007). Determina-se também um critério para avaliar a qualidade dessas estimativas, como a diferença entre o valor dado por $f(x_i)$ e o valor y_i , a qual é chamado de resíduo. Ressalta-se a importância desta precisão; haja vista a possibilidade de se aceitar resíduos pequenos, mas de rejeitar resíduos grandes (CORTES e VAPNIK, 1995).

A fim de penalizar esses erros de estimação, ou seja, falhas para prever precisamente uma dada variável, estabelece-se uma função de custo, também denominada de função de perda. Na maioria dos casos a mesma será do tipo $c(x, y, f(x)) = c(f(x) - y)$. Dessa maneira, podem-se usar os dados de treinamento para estimar uma função f que reduza o valor esperado da perda fornecido pelo risco funcional (CRISTIANINI e SHAW-TAYLOR, 2000; DIAS, 2007):

$$R[f] = \int c(x, y, f(x)) dP(x, y);$$

onde $c(x, y, f(x))$ é a função de custo.

Entretanto, como se desconhece a distribuição $P(x, y)$, pode-se estimar uma função f que de alguma maneira esteja “próxima” daquela que minimiza o risco funcional $R[f]$. A aproximação possível para o risco funcional consiste em substituir a integral por uma estimativa empírica; obtendo-se assim, o risco funcional empírico:

$$R_{emp} [f] := \frac{1}{l} \sum_{i=1}^l c(x_i, y_i, f(x_i))$$

Este risco pode ser compreendido como uma medida do erro médio no conjunto de treinamento, para um número de observações finitas e fixas. A máquina garante que assintoticamente quando $l \rightarrow \infty$, o risco empírico irá convergir para o risco esperado, desde que sejam fornecidas algumas condições (KECMAN, 2001).

Todavia, quando se tem poucos dados em espaços de grande dimensionalidade, podem ocorrer alguns problemas ao se utilizar uma função f para minimizar o risco empírico de uma classe de função F , como por exemplo: grandes desvios, *overfitting* (excesso de ajuste) e generalizações ruins. Para contornar tal situação, deve-se restringir a classe de soluções admissíveis, ou seja, reduzir a classe F (CRISTIANINI e SHAWE-TAYLOR, 2000; KECMAN, 2001; SMOLA e SCHOLKOPF, 2004).

No contexto da teoria da MVS na aproximação por funções de regressão (LS-SVM), alguns estudos foram realizados para consolidar a aplicação desta técnica em diversos campos da ciência. DIAS (2007) conclui que a LS-SVM é uma técnica promissora para previsão da Estrutura a termo da Taxa de Juros Brasileira (ETTJ brasileira) ao comparar os resultados deste método com os tradicionais modelos VAR (Vetor Auto-Regressivo) e ECM (Modelos de Correção de Erros). CAO e TAY (2003) também verificou a aplicação de SVM em séries temporais financeiras e ressaltou o elevado desempenho na etapa de generalização, com resultados comparáveis aos obtidos com Redes Neurais do tipo RBF.

Reporta-se a utilização de LS-SVM em outras áreas, como por exemplo: Zootecnia, pela pesquisa de ALONSO et al. (2013) ter comprovado a eficiência da LS-SVM para prever o peso da carcaça de bovinos de corte; Fruticultura, com o estudo de classificador de mínimos quadrados para detectar o grau de escurecimento em função da dimensão fractal nos frutos de manga (ZHENG e LU, 2012); Química, com a LS-SVM resultando em menor erro quando comparada a regressão para determinar parâmetros de qualidade em matrizes lácteas (BORIN, 2007), e também com a maior capacidade preditiva da técnica dos constituintes químicos lignina e celulose em madeira de

eucalipto testados por NUNES et al. (2011). Contudo, desconhecem-se estudos testando o potencial da LS-SVM para problemas no âmbito do manejo florestal, principalmente no tocante a prognose da produção.

6. Regulação Florestal

Regular a produção de uma floresta significa determinar estratégias ou prescrições de manejo para converter uma estrutura já existente de classes de idades em uma progressão aritmética. Isto feito, ao final de um horizonte de planejamento (HP), essas idades poderão variar de I até a idade de rotação (R anos), resultando assim em uma produção sustentável (CAMPOS e LEITE, 2013). Dentre os vários benefícios gerados por uma floresta regulada, cita-se a uniformidade da produção no tocante ao volume produzido, tamanho, valor e qualidade; promovendo assim, uma auto-suficiência do setor e prevenindo contra a escassez de madeira ou exaustão da floresta (DAVIS, 1966). Contudo, essa produção sustentável deve atender às especificações de mercado, oscilando dentro de um limite determinado, mas atendendo à demanda, a qual pode variar ao longo do tempo. Isto possibilita também uma base regular de emprego e o equilíbrio entre despesas anuais e receitas (RODRIGUES, 1997).

A ideia da estrutura regulada se insere no campo do planejamento estratégico, haja vista esta ser uma atividade a longo prazo. Dentre as questões básicas que devem ser do conhecimento do manejador, cita-se: qual a melhor composição de espécies (ou clones), espaçamento inicial, intensidade e ciclo de desbaste, além de como definir a rotação florestal de colheita. Percebe-se então que os modelos de regulação são construídos visando atender os objetivos estipulados no plano da empresa florestal, com a inclusão de restrições de ordem econômica, social e ambiental (CAMPOS e LEITE, 2013). Além de demandar habilidade e experiência do gestor florestal, um modelo de regulação da produção é extremamente complexo devido ao excesso de variáveis envolvidas.

As diversas prescrições silviculturais geradas para cada unidade de manejo são avaliadas por diferentes métodos matemáticos, com intuito de se selecionar o melhor plano de manejo. Por conseguinte, diversas ferramentas de pesquisa operacional foram introduzidas na planejamento florestal, permitindo assim uma redução de riscos e incertezas na tomada de decisão dos planos empresariais (BASKENT e KELES, 2005). Atualmente, utilizam-se métodos numéricos para processamento dos modelos de

regulação da produção. Dentre aqueles relacionados com a pesquisa operacional, exemplifica-se a programação linear (PL), programação multiobjetivo (PMO), programação inteira (PI), programação inteira mista (PIM), programação dinâmica (PD), metas-heurísticas e simulação (BINOTI, 2010). Ressalta-se que a escolha da melhor ferramenta para alocação ótima dos recursos florestais dependerá da formulação do modelo e das características da floresta (BETTINGER et al., 2009).

Diversos autores aplicaram os modelos de regulação em problemas do setor florestal: SANTOS (2008) desenvolveu uma metodologia para definir um regime ótimo de desbastes e a produção de multiprodutos de madeira de clones de eucalipto, visando o aproveitamento de diferentes sortimentos e posterior definição de uma estratégia de regulação. RODRIGUES (2006) avaliou estratégias para gerar alternativas de manejo na formulação e solução de problemas de planejamento florestal com restrições de recobrimento. NALLI et al. (1996) e NASSET (1997) são exemplos onde objetivos espaciais foram incluídos no planejamento de modelos, definindo restrições de área para cada unidade de gestão.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALCÂNTARA, A. E. M. de. **Alternativas de modelagem para projeção do crescimento de eucalipto em nível de povoamento**. Viçosa: UFV, 2012. 66p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.
- ALONSO, J.; CASTANON, A. R.; BAHAMONDE, A. Support Vector Regression to predict carcass weight in beef cattle in advance of the slaughter. **Computers and Electronics in Agriculture**, n.91, p. 116-120, 2013.
- ARCE, J. E.; GOMES, F. dos S.; SANQUETTA, C. R.; CRUZ, E. C. da. Utilização dos processos de difusão na avaliação e simulação precoces do crescimento de povoamentos de *Pinus taeda* L. **Revista Cerne**, v.4, n.1, p.154-170, 1998.
- BARTOSZECK, A. C. de P. e S.; MACHADO, S. do A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E. B. Distribuição diamétrica para bracingais em diferentes idades, sítios e densidades na região metropolitana de Curitiba. **Revista Floresta**, Curitiba-PR, v.3, n.34, p.305-323, set/dez. 2004.
- BASKENT, E. Z.; KELES, S. Spatial forest planning: a review. **Ecological Modelling**, n. 188, p. 145-173, 2005.
- BETTINGER, P.; BOSTON, K.; SIRY, J. P.; GRENNER, D. L. **Forest Management and Planning**. Elsevier, 2009. 342p.
- BINOTI, M. L. M. S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. Viçosa: UFV, 2010. 54p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.
- BINOTI, D. H. B. **Sistemas computacionais aplicados ao manejo florestal**. Viçosa: UFV, 2012. 127p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.
- BORGES, J. S. **Parametrização, calibração e validação do modelo 3-PG para eucalipto na região do cerrado de Minas Gerais**. Viçosa: UFV, 2009. 77p. Dissertação (Mestrado) – Programa de Pós-Graduação em Solos, Universidade Federal de Viçosa, Viçosa, 2009.
- BORIN, A. **Aplicações de Máquinas de vetores de suporte por mínimos quadrados (LS-SVM) na quantificação de parâmetros de qualidade de matrizes lácteas**. Campinas: UNICAMP, 2007. 128 p. Tese (Doutorado) – Programa de Pós-Graduação em Química Analítica, Universidade Estadual de Campinas, Campinas, 2007.
- BRAGA, A. de P.; CARVALHO, A. P. de L. F. de; LUDEMIR, T. B. **Redes Neurais Artificiais: teorias e aplicações**. Rio de Janeiro: Editora LTC, 2000. 262p.
- BUGMANN, H. A review of forest gap models. **Climatic Change**, n.51, p.259–305, 2001.

BUONGIORNO, J.; GILLESS, J. K. **Decision Methods for Forest Resource Management**. Elsevier, 2003. 458p.

CALEGARIO, N.; CALEGARIO, C. L. L.; MAESTRI, R.; DANIELS, R. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Florestalis**, n. 69, p. 38-50, dez. 2005.

CAMPOS, B. P. F.; BINOTI, D. H. B.; SILVA, M. L. da; LEITE, H. G.; BINOTI, M. L. M. da S. Conversão de árvore em multiprodutos da madeira utilizando programação inteira. **Revista Árvore**, Viçosa-MG, v.37, n.5, p.881-887, 2013.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. Viçosa: Editora UFV, 2013. 605 p.

CAO, L. J.; TAY, F. E. H. Support Vector Machine With Adaptive Parameters in Financial Time Series Forecasting. **IEEE Transactions on Neural Networks**, v. 14, n. 6, p.1506-1518, dez. 2003.

CASTRO, R. V. O. **Modelagem do crescimento em nível de árvores individuais utilizando redes neurais e autômatos celulares**. Viçosa: UFV, 2011. 90p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2011.

CHASSOT, T.; FLEIG, F. D.; FINGER, C. A. G.; LONGHI, S. J. Modelos de crescimento em diâmetro de árvores individuais de *Araucaria angustifolia* (Bertol.) Kuntze em Floresta Ombrófila Mista. **Revista Ciência Florestal**, Santa Maria, v. 21, n. 2, p. 303-313, abr.-jun. 2011.

CHERKASSKY, V.; MULIER, F. **Learning from Data: Concepts, Theory, and Methods**. Wiley: IEEE Press, 2007.

CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, G. H.; BAILEY, R. L. **Timber Management: a quantitative approach**. New York: J. Wiley, 1983. 333p.

CORTES, C.; VAPNIK, V. Support-Vector Networks. **Machine Learning**, n. 20, p.273-297, 1995.

CRISCI, C.; GHATTASB, B.; PERERAC, G. A review of supervised machine learning algorithms and their applications to ecological data. **Ecological Modelling**. n.240, p.113– 122, 2012.

CRISTIANINI, N.; SHAW-TAYLOR, J. **An introduction to Support Vector Machines and other kernel-based learning methods**. Cambridge University, 2000. 195p.

CUNHA NETO, F. R. da; SCOLFORO, J. R. S.; CALEGARIO, N.; DE OLIVEIRA, A. D.; KANEGAE JÚNIOR, H. Modelo para predição da produção por classe de diâmetro para *Eucalyptus grandis*. **Revista Cerne**, v.1, n.1, p. 108-122, 1994.

DAVIS, L. S.; JOHNSON, K. N. **Forest Management**. New York: McGraw-Hill, 1987. 790p.

DIAS, A. N.; LEITE, H. G.; CAMPOS, J. C. C.; COUTO, L.; DE CARVALHO, A. F. Emprego de um modelo de crescimento e produção em povoamentos desbastados de eucalipto. **Revista Árvore**, v. 29, n. 5, p. 731-739, set./out. 2005.

DIAS, M. S. **O uso de Máquina de Suporte Vetorial para Regresso (SVR) na Estimação da Estrutura a Termo da Taxa de Juros do Brasil**. Rio de Janeiro: PUC, 2007. 77p. Dissertação (Mestrado) – Programa de Pós-Graduação em Matemática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.

FERRARI, M. P.; FERREIRA, C. A.; DA SILVA, H. D.; TREVISAN, R. Prognose do crescimento volumétrico individual de árvores de *Eucalyptus* em povoamentos na Região Centro Sul. **Boletim Pesquisa Florestal**, Colombo, n. 51, p. 05-16, jul./dez. 2005.

GARCÍA, O. Growth modeling: a (re)view. **New Zealand Forestry**, v.33, n.3, p.14-17, 1988.

GORGENS, E. B. **Estimação do volume de árvores utilizando Redes Neurais Artificiais**. Viçosa: UFV, 2006. 94p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2006.

GUIMARÃES, D. P.; SILVA, G. G. C.; SANS, L. M. A.; LEITE, F. P. Uso do modelo de crescimento 3-PG para o zoneamento do potencial produtivo do eucalipto no estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Piracicaba, v.15, n.2, p. 192-197, 2007.

GUJARATI, D. N. **Econometria Básica**. São Paulo: Makron Books, 2006. 846p.

HASENAUER, H. Princípios para a modelagem de ecossistemas florestais. **Ciência e Ambiente**, Santa Maria, v. 20, n. 1, p. 53-69, 2000.

HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre, 2001. 900 p.

KECMAN, V. **Learning and soft computing: support vector machine, neural networks and fuzzy logistic**. MIT Press, 2001. 540 p.

KING, S. L.; BENNETT, K. P.; LIST, S. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. **Computers and Electronics in Agriculture**, n. 27, p.401-406, 2000.

LEITE, H. G.; NOGUEIRA, G. S.; CAMPOS, J. C. C.; DE SOUZA, A. L.; CARVALHO, A. Avaliação de um modelo de distribuição diamétrica ajustados para povoamentos de *Eucalyptus* sp. submetidos a desbaste. **Revista Árvore**, Viçosa-MG, v.29, n.2, p.271-280, 2005.

LORENA, A. C. **Investigação de estratégias para a geração de máquinas de vetores suporte multiclassés**. São Paulo: USP, 2006. 229p. Tese (Doutorado) – Programa de

Pós-Graduação em Ciências de Computação e Matemática Computacional, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2006.

MACHADO, S. do A.; AUGUSTYNCZIK, A. L. D.; NASCIMENTO, R. G. M.; TÊO, S. J. ; MIGUEL, E. P.; FIGURA, M. A.; SILVA, L. C. R. da. Funções de distribuição diamétrica em um fragmento de Floresta Ombrófila Mista. **Revista Ciência Rural**, Santa Maria, v.39, n.8, p.2428-2434, 2009.

MALTAMO, M.; KANGAS, A.; UTERRA, J.; TORNIAINEN, T. SARAMÄKI, J. Comparison of percentile based prediction na the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. **Forest Ecology and Management**, v. 133, p. 263-274, 2000.

MARTINS, F. B. **Modelagem de crescimento em nível de árvore individual para plantios comerciais de eucaliptos**. Viçosa: UFV, 2011. 159p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2011.

McQUEEN, R. J.; GAMER, S. R.; NEVILL-MANNING, C. G.; WITTEN, I. H. Applying machine learning to agricultural data. **Computers and Electronics in Agriculture**, n. 12, p. 275-293, 1995.

MENDES, B. R.; CALEGARIO, N.; VOLPATO, C. E. S.; DE MELLO, A. A. Desenvolvimento de modelos de crescimento de árvores individuais fundamentados em equações diferenciais. **Revista Cerne**, Lavras, v. 12, n. 3, p. 254-263, jul./set. 2006.

MEYER, D. **Support Vector Machine**. Disponível em: <http://www.researchgate.net/publication/242323440_Support_Vector_Machines_The_Interface_to_libsvm_in_package_e1071>. Acesso em: 18 jul. 2013.

NAESSET, E. A spatial decision support system for long-term forest management planning by means of linear programming and GIS. **Scandinavian Journal of Forest Research**, n. 12, p. 77-88, 1997.

NAKASHIZUKA, T.; KOHYAMA, T. The significance of the asymmetric effect of crowding for coexistence in a mixed temperate forest. **Vegetal Science**, v. 6, n.4, p. 509-516, 1995.

NALLI, A.; NUUTINEN, T.; PAIVINEN, R. Site specific constraints in integrated forest planning. **Scandinavian Journal of Forest Research**, n. 11, p. 85-96, 1996.

NOBLE, W. S. What is a support vector machine? **Nature Biotechnology**, v. 24, n. 12, p. 1565-1567, 2006.

NOGUEIRA, G. S. **Modelagem do crescimento e da produção de povoamentos de *Eucalyptus* sp. e de *Tectona grandis* submetidos a desbaste**. Viçosa: UFV, 2003. 126p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2003.

NUNES, C. A.; LIMA, C. F.; BARBOSA, L. C. de A.; COLODETTE, J. L.; FIDÊNCIO, P. H. Determinação de constituintes químicos em madeira de eucalipto por PI-CG/EM e calibração multivariada: comparação entre Redes Neurais Artificiais e Máquinas de Vetor Suporte. **Química Nova**, v. 34, n. 2, p. 279-283, 2011.

NUTTO, L. Manejo do crescimento diamétrico de *Araucaria angustifolia* (Bert.) O. Ktze. baseado na árvore individual. **Revista Ciência Florestal**, Santa Maria, v. 11, n. 2, p. 9-25, 2001.

OLIVEIRA, E. B. de. **Um sistema computadorizado de prognose do crescimento e produção de *Pinus taeda* L., com critérios quantitativos para a avaliação técnica e econômica de regimes de manejo.** Curitiba: UFPR, 1995. 126p. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia Florestal, Universidade Federal do Paraná, Curitiba, 1995.

OLIVEIRA, J. de C. **Mínimos quadrados generalizados: uma proposta à modelagem do crescimento e produção em povoamentos não desbastados de eucalipto.** Viçosa: UFV, 2012. 80p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; BROOKS, J. R.; WIANT JUNIOR, H.V. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of Environmental Management**, n. 91, p. 742-753, 2010.

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; CRESCENTE-CAMPO, F. ELER, U. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology and Management**, n. 306, p. 52-60, 2013.

PORTÉ, A.; BARTELINK, H. H. Modelling mixed Forest growth: a review of models for forest mangement. **Ecological Modelling**, n.150, p. 141-188, 2002.

RAUBER, T. W. **Redes Neurais Artificiais.** Disponível em:
<<http://www.inf.ufes.br/~thomas/pubs/eri98.pdf>> Acesso em: 17 jul. 2013.

RECKNAGEL, F. Applications of machine learning to ecological modeling. **Ecological Modelling**, n. 146, p. 303-310, 2001.

RENNOLLS, K.; GEARY, D. N.; ROLLINSON, T. J. D. Characterizing diameter distributions by the use of the Weibull distribution. **Forestry**, v.58, n.1, p.57-66, 1985.

RESENDE, R. R.; DO VALE, A. B.; SOARES, T. S.; DA SILVA, M. L.; COUTO, L.; DO VALE, R. S. Emprego de um modelo de crescimento e produção para determinação de rotação em povoamentos de eucalipto. **Revista Árvore**, Viçosa-MG, v.28, n.2, p.219-225, 2004.

RETSLAFF, F. A. de S.; FIGUEIREDO FILHO, A.; DIAS, A. N.; BERNETT, L. G.; FIGURA, M. A. Prognose do crescimento e da produção em classe de diâmetro para povoamentos desbastados de *Eucalyptus grandis* no sul do Brasil. **Revista Árvore**, Viçosa-MG, v.36, n.4, p.719-732, 2012.

RODRIGUES, F. L. **Regulação de florestas equiâneas utilizando programação linear**. Viçosa: UFV, 1997. 109p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 1997.

RODRIGUES, F. L.; SILVA, G. F. da; LEITE, H. G.; XAVIER, A. C.; PEZZOPANE, J. E. M. Um modelo de regulação florestal e suas implicações na formulação e solução de problemas com restrições de recobrimento. **Revista Árvore**, Viçosa-MG, v.30, n.5, p.769-778, 2006.

SALLES, T. T. **Modelagem do crescimento e produção de clones de eucalipto em um sistema agroflorestal**. Viçosa: UFV, 2010. 87p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.

SANTANA, C.; MELLO, A. A. de; EISFELD, R. de L.; SANQUETTA, C. R. Sistemas de equações para simulação do crescimento e da produção em povoamentos de *Eucalyptus grandis* Hill ex Maides. sem desbaste baseado no modelo de Clutter. **Revista Ambiência**, v.1, n. 2, p. 239-256, jun./dez. 2005.

SANTOS, A. A. **Desenvolvimento e avaliação de um modelo mecanístico para estimativa da produtividade da cultura do cafeeiro**. Viçosa: UFV, 2002. 54p. Dissertação (Mestrado) – Programa de Pós-Graduação em Meteorologia Agrícola, Universidade Federal de Viçosa, Viçosa, 2002.

SANTOS, K. dos; SANQUETTA, C. R.; EISFIELD, R. de L.; WATZLAWICK, L. F.; ZILIOOTTO, M. A. B. Equações volumétricas por classe diamétrica para algumas espécies folhosas da Floresta Ombrófila Mista no Paraná, Brasil. **Revista Ciências Exatas e Naturais**, v. 8, n.1, p.99-112, jan./jun. 2006.

SANTOS, F. L. de. **Regulação da produção de floresta de eucalipto submetida a desbaste e destinada a multiprodutos**. Viçosa: UFV, 2008. 80p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2008.

SCHWAAB, M.; PINTO, J.C. **Análise de Dados Experimentais I: fundamentos de estatística e estimação de parâmetros**. Rio de Janeiro: E-papers, 2007. 462p.

SILVA, A. A. L. **Emprego de modelos de crescimento e produção em análise econômica de decisões de manejo florestal**. Viçosa: UFV, 2001. 69 p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2001.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, v.33, n.6, p.1133-1139, 2009.

SMOLA, A. J.; SCHOLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, n.14, p.199–222, 2004.

SOARES, C. P. B.; PAULA NETO, F. de; SOUZA, A. L. de. **Dendrometria e Inventário Florestal**. Viçosa: Editora UFV, 2006. 276p.

SOUTO, M. C. P.; LORENA, A. C.; DELBEM, A. C. B.; CARVALHO, A. C. P. L. F. Técnicas de Aprendizado de Máquina para problemas de Biologia Molecular. In: XXIII Congresso da Sociedade Brasileira de Computação, 2003. **Anais...** 2003. p. 103-152.

SOUZA, C. A. M. de; CHASSOT, T.; FINGER, C. A. G.; SCHNEIDER, P. R.; FLEIG, F. D. Modelos de afilamento para o sortimento do fuste de *Pinus taeda* L. **Ciência Rural**, Santa Maria, v.38, n.9, p.2506-2511, dez. 2008.

SULLIVAN, A. D.; CLUTTER, J. L. A simultaneous growth and yield model for Loblolly Pine. **Forest Science**, v. 18, n. 1, p. 76-86, 1972.

TRINCADO, G. V.; QUEZADA, R. P.; KLAUS, V. G. A comparison of two stand table projection methods for young *Eucalyptus nitens* (Maiden) plantations in Chile. **Forest Ecology and Management**, v180, p.443-451, 2003.

TURAK, E.; FERRIER, S.; BARRETT, T.; MESLEY, E.; DRIELSMA, M.; MANION, G.; DOYLE, G.; STEIN, J.; GORDON, G. Planning for the persistence of river biodiversity: exploring alternative futures using process-based models. **Freshwater Biology**, v. 56, p. 39-56, 2011.

UMAÑA, C. L. A.; ALENCAR, J. da C. Distribuições diamétricas da floresta tropical úmida em uma área no município de Itacoatiara-AM. **Acta Amazonica**, n. 28, p. 167-190, 1998.

VANCLAY, J. K. **Modelling forest growth and yield: applications to mixed tropical forests**. Copenhagen: CAB International, 1994. 312 p.

VANCLAY, J.K. Growth models for tropical forests: a synthesis of models and methods. **Forest Science**, v.41, p. 7-42, 1995.

VELLASCO, M. M. B. R. **Redes Neurais Artificiais: Introdução**. Disponível em: <<http://www.ica.ele.puc-rio.br/Downloads%5C33/ICA-introdu%C3%A7%C3%A3o%20RNs.pdf>> Acesso em: 29 mar. 2013.

YOUNG, R. A.; GIESE, R. L. **Introduction to Forest Ecosystem Science and Management**. Elsevier, 2003. 589 p.

ZEIDE, B. Analysis of Growth Equations. **Forest Science**, v. 39, n. 3, p. 594-616, 1993.

ZHENG, H.; LU, H. A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). **Computers and Electronics in Agriculture**, n.83, p. 47-51, 2012.

CAPÍTULO 2: TÉCNICAS PARA PROJEÇÃO DA PRODUÇÃO FLORESTAL

RESUMO

Entender o comportamento do crescimento das árvores e prever sua produção sempre foi interesse dos profissionais do setor florestal. A importância disto se reflete no ordenamento da produção, assegurando-a como socialmente justa, ecologicamente sustentável e economicamente viável. Diante disso, este trabalho objetivou contribuir para o estudo da modelagem do crescimento e da produção florestal ao comparar o desempenho de três técnicas (Análise de Regressão pelo Modelo de Crescimento, Máquina de Vetores de Suporte – MVS e Redes Neurais Artificiais – RNA) na projeção da produção de eucalipto. Como modelo de crescimento testemunha empregou-se o Modelo de Clutter na sua forma usual e completa. Para as técnicas de MVS e RNA, a produção foi estimado através de dois métodos: um direto, utilizando os valores de área basal auferidos em campo; e outro, indireto, no qual primeiramente estimou-se a área basal e estes valores foram utilizados como input para estimar a produção. Em ambos os casos, o volume por hectare foi considerado como variável dependente, sendo utilizado pelas MVS e RNA como variáveis independentes contínuas: idade (I), altura dominante (Hd), área basal (Ab) e índice de local (S); e como variáveis categóricas: rotação, tipo de solo, bacia hidrográfica, altitude, mecanização, material genético, nome do projeto e espaçamento. As estimativas foram analisadas calculando-se as estatísticas: coeficiente de correlação ($r_{y\hat{y}}$), coeficiente de determinação ($R^2\%$), coeficiente de determinação ajustado ($\bar{R}^2\%$) e erro padrão da estimativa (S_{yx}). Além disso, efetuou-se a análise gráfica de resíduos. O modelo de Clutter apresentou coeficiente de determinação de 94,43. Todavia, observou-se que as técnicas de inteligência artificial promoveram um ganho em precisão nas estimativas da produção. Dentre todas as MVS testadas, a que utilizou o método indireto com função *kernel* RBF apresentou o melhor resultado devido aos seus resíduos, aproximadamente 60%, encontrarem-se entre as amplitudes de classes de mais ou menos 10%. Com relação a RNA, os resultados obtidos pelo método indireto com a rede MLP 143-14-1 foram mais precisos, pois aproximadamente 60%, encontrarem-se entre as amplitudes de classes de mais ou menos 5%. Diante disso, é seguro afirmar que as técnicas de inteligência artificial, MVS e RNA, apresentam-se como uma ferramenta alternativa para estudos de projeção da produção florestal, gerando um aumento na precisão das estimativas, uma redução no tempo em relação ao método usual de modelagem e a vantagem de se poderem incluir dados categóricos aos modelos.

Palavras-chave: Modelagem florestal, Aprendizagem de máquina, Regressão Linear.

1. INTRODUÇÃO

A necessidade de deter o conhecimento sobre a situação dos povoamentos, bem como de sua dinâmica, conduz a um contínuo aprimoramento das técnicas de modelagem do crescimento e da produção florestal (RESENDE et al., 2004). Para desempenhar tal função, o gestor florestal deve conhecer as técnicas para modelagem dos dados, os valores numéricos a serem incorporados nessas equações e a lógica necessária para interligá-las de uma forma significativa (VANCLAY, 1995).

Ademais, na perspectiva de eficiência produtiva, as intervenções realizadas pelo gestor ao longo da vida do povoamento são tanto importantes quanto a instalação e a colheita do mesmo. Torna-se assim, essencial antever com precisão os resultados das diferentes estratégias de ação, bem como detectar e corrigir possíveis falhas durante suas execuções (PORTÉ e BARTELINK, 2002; YOUNG e GIESE, 2003).

Diante disso, os modelos de crescimento e produção apresentam-se como uma maneira eficiente para compreender a dinâmica de um povoamento florestal e colaborar na tomada de decisão; haja vista poderem englobar o crescimento através de diversas variáveis, incluindo a mortalidade, e quaisquer outras alterações observadas na composição e estrutura do povoamento (BUONGIORNO e GILLESS, 2003).

Geralmente, esses modelos são constituídos por equações que consideram o volume como variável dependente, e o relacionam com diversas variáveis independentes de fácil mensuração na floresta, como o diâmetro à altura do peito e a altura (CAMPOS e LEITE, 2013). Uma técnica amplamente utilizada na estimativa dos parâmetros nessas equações volumétricas é a regressão (GARCÍA, 1988).

Apesar de diversos trabalhos científicos atestarem sua eficiência, a modelagem do crescimento utilizando a técnica de regressão requer que o povoamento florestal seja estratificado, o que pode proporcionar uma ineficiência ao modelo. Além disso, algumas características dos dados oriundos de plantios clonais utilizados atualmente podem causar violação de determinadas pressuposições estatísticas da distribuição dos erros (GARCÍA, 1988). Dentre elas, cita-se: a ocorrência de heterocedasticidade da variância, decorrente do aumento da idade dos povoamentos, e autocorrelação residual, a qual está associada à remedição das variáveis ao longo do tempo (OLIVEIRA, 2012).

Contudo, uma alternativa a modelagem do crescimento e da produção florestal pelo método de regressão linear é a aplicação de técnicas de inteligência artificial (McQUEEN et al., 1995; RECKNAGEL, 2001; GORGENS, 2006; SILVA et al., 2009;

BINOTI, 2010; KING et al., 2000; CASTRO, 2011; BINOTI, 2012; CRISCI et al., 2012; ZHENG e LU, 2012); como as Redes Neurais Artificiais (RNA), Máquinas de Vetores Suporte (MVS) e Autômatos Celulares. Estudos na área de mensuração e manejo florestal com aplicação destas técnicas são incipientes e muito promissores; haja vista a superioridade de precisão das estimativas fornecidas pelas mesmas (BINOTI, 2010).

Sendo assim, com a finalidade de contribuir para o estudo da modelagem de dados florestais e aumentar o número de ferramentas disponíveis ao manejador, sobretudo ferramentas que melhorem a precisão das estimativas, testou-se o desempenho de duas técnicas de inteligência artificial não-paramétrica, Redes Neurais Artificiais (RNA) e Máquina de Vetor Suporte (MSV), para a projeção da produção de povoamentos de eucalipto.

2. HIPÓTESES

H_0 : A projeção da produção de povoamentos de eucalipto utilizando as técnicas de inteligência artificial difere daquela projeção obtida pelo método de regressão linear.

H_1 : Não H_0 .

3. METODOLOGIA

3.1. Descrição dos dados

O estudo foi conduzido em povoamentos de eucalipto, em uma fazenda modelo localizada na região Centro Oeste do Estado de Minas Gerais, entre as coordenadas geográficas 19^o13'28'' de latitude sul e 43^o4'0'' de longitude oeste. Segundo a classificação climática de Köppen, o clima predominante na região é Aw – Clima Tropical chuvoso de savana, o qual determina que o inverno seja seco e que ocorram chuvas máximas no verão (SOUZA et al., 2003). A precipitação anual média é de 1212 mm, com temperatura variando em média de 15°C a 30°C (INMET, 2013). A área apresenta relevo fortemente ondulado, marcada por solos, em sua maioria, classificados como latossolo vermelho amarelo, podendo ocorrer também cambissolo háplico e neossolo flúvico (FEAM, 2010).

Utilizaram-se dados oriundos de inventário florestal contínuo conduzido em 1452 parcelas permanentes de área fixa de cerca de 400 m² e medidas de 2005 até 2012. Encontram-se nessa área 83 clones de eucalipto e alguns talhões com plantios seminais implantados sob diferentes arranjos espaciais (3.00 x 2.00; 3.00 x 2.20; 3.00 x 2.50; 3.00 x 2.70; 3.00 x 2.80; 3.00 x 3.00; 3.00 x 3.30; 3.33 x 2.00; 4.00 x 2.50). Em todas as parcelas mediram-se anualmente a área basal (*Ab*), a altura das árvores dominantes (*Hd*) e o volume até 4 cm com casca (*V*) (Tabela 1). Esses dados foram submetidos a consistência, de acordo com recomendações de CAMPOS e LEITE (2013).

Tabela 1 – Descrição dos dados de uma fazenda modelo utilizados para projeção da produção pelas técnicas de Modelo de Crescimento e Máquina de Vetores Suporte.

Casse de Idade (meses)	Número de árvores por hectare – <i>N</i>	Altura dominante (m) – <i>Hd</i>			Área basal (m ² ha ⁻¹) – <i>Ab</i>			Volume até 4 cm de diâmetro com casca (m ³ ha ⁻¹) – <i>V</i>			Índice de local para uma idade-índice de 72 meses (m) – <i>S</i>		
		Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max
17-35	1515	7,57	15,53	26,03	2,82	10,35	20,30	8,94	67,78	198,75	18,74	30,38	46,30
36-54	1220	11,67	22,54	31,06	8,25	17,90	27,43	45,68	166,51	347,41	16,36	28,66	38,18
55-73	1255	17,00	27,09	37,30	8,63	22,17	34,57	61,29	243,51	429,40	18,43	28,47	37,80
74-92	975	19,77	31,03	41,43	0,83	25,75	37,36	8,30	322,45	559,31	18,44	29,43	39,02
93-111	778	16,68	33,29	45,03	10,91	27,30	40,79	63,54	357,89	616,43	14,52	29,67	41,07
112-130	156	19,77	35,20	45,00	9,64	28,27	43,65	70,17	391,31	669,37	16,79	30,13	38,26
131-149	48	23,03	34,88	46,10	16,47	28,13	46,43	164,87	373,67	728,29	18,83	28,76	38,35
150-168	25	23,43	34,00	41,10	7,12	26,53	45,72	90,12	354,23	711,40	18,88	27,41	33,13
169-187	6	32,42	40,96	48,88	24,93	35,72	43,59	289,46	544,77	708,42	25,55	32,28	38,52
188-207	12	32,40	42,29	51,00	24,88	36,30	45,28	284,20	588,73	808,45	25,28	32,86	39,80

3.2. Classificação da capacidade produtiva

A capacidade produtiva dos povoamentos foi determinada pelo método da curva-guia (CLUTTER *et al.*, 1983), estipulando-se 72 meses como idade índice, através do modelo exponencial desenvolvido por Schumacher (CAMPOS e LEITE, 2013):

$$Hd = e^{\alpha + \beta I^{-1}} + \varepsilon$$

em que: Hd = altura dominante, em metros; I = idade, em meses; α , β e γ = parâmetros; e ε = erro aleatório, $\varepsilon \sim \text{NID}(0, \sigma^2)$.

3.3. Modelagem do crescimento e da produção

O estudo do crescimento e da produção foi conduzido em nível de povoamento, sendo considerado como unidade de manejo a parcela. Foram empregadas duas alternativas: A) Modelo de crescimento proposto por Clutter (sistemas de equações), B) Máquina de Vetores Suporte (MVS) e C) Redes Neurais Artificiais (RNA).

3.3.1. Modelo de Crescimento

O modelo de crescimento considerado foi o proposto por Clutter em 1963. Nele a produção é expressa em função das idades atual e futura (I_1 e I_2), do índice de local na idade atual (S_1) e das densidades atual e futura (B_1 e B_2) (CAMPOS e LEITE, 2013). Empregou-se o modelo de Clutter na sua forma usual (1) e na forma completa (2). Os sistemas de equações testados foram:

$$1 \left\{ \begin{array}{l} \text{Ln}B_2 = \text{Ln}B_1(I_1 I_2^{-1}) + \alpha_0(1 - I_1 I_2^{-1}) + \alpha_1(1 - I_1 I_2^{-1})S_1 + \varepsilon \\ \text{Ln}V_2 = \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 \text{Ln}B_2 + \varepsilon \end{array} \right.$$

$$2 \left\{ \begin{array}{l} \text{Ln}B_2 = \text{Ln}B_1(I_1 I_2^{-1}) + \alpha_0(1 - I_1 I_2^{-1}) + \alpha_1(1 - I_1 I_2^{-1})S_1 + \varepsilon \\ \text{Ln}V_2 = \beta_0 + \beta_1 I_2^{-1} + \beta_2 S_1 + \beta_3 \text{Ln}B_2 + \beta_4 S_1 \text{Ln}B_2 + \beta_5 S_1 I_2^{-1} + \beta_6 \text{Ln}B_2 I_2^{-1} + \varepsilon \end{array} \right.$$

em que: V_2 = volume na idade de projeção I_2 , em $\text{m}^3 \text{ha}^{-1}$; I_1 = idade atual, em meses; I_2 = idade de projeção, em meses; S_1 = índice de local na idade atual, em m; B_1 = área basal na idade atual I_1 , em $\text{m}^2 \text{ha}^{-1}$; B_2 = área basal na idade de projeção I_2 , em $\text{m}^2 \text{ha}^{-1}$; α_i e β_i = parâmetros; Ln = logaritmo neperiano; e ε = erro aleatório, $\varepsilon \sim \text{NID}(0, \sigma^2)$.

Com intuito de obter tendências consistentes para a modelagem, alguns dados das parcelas foram eliminados, haja vista serem identificados como *outliers* (DRAPPER e SMITH, 1968; GUJARATI, 2006).

Os ajustes destes sistemas de equações foram realizados pelo método dos mínimos quadrados em dois estágios. A avaliação estatística do modelo baseou-se na análise dos resíduos das variáveis área basal e volume. Além disso, foi observada a capacidade dos sistemas em descrever o fenômeno biológico implícito no crescimento do povoamento.

3.3.2. Máquina de Vetores Suporte (MVS)

Para testar esta técnica, a produção foi estimada através de dois métodos: um direto, utilizando os valores de área basal auferidos em campo; e outro indireto, no qual primeiramente estimou-se a área basal e estes valores serviram como input para estimar a produção. Os valores que extrapolaram a tendência geral de cada projeto, em área basal ou volume, não foram eliminados a fim de verificar a capacidade das máquinas em lidar com *outliers* e ruídos.

Em ambos os casos, a produção foi considerada como variável dependente, sendo utilizado como variáveis independentes contínuas: idade (I), altura dominante (Hd), área basal (Ab) e índice de local (S); e como variáveis categóricas: rotação, tipo de solo, bacia hidrográfica, altitude, mecanização, material genético (nome do clone), projeto (nome do projeto) e espaçamento. Ressalta-se que as variáveis categóricas altitude e mecanização foram divididas em classes; enquanto que para a variável categórica rotação, atribuíram-se valor 1 para alto fuste e 2 para talhadia.

Avaliaram-se duas funções de erro, tipo I e II, também conhecidas como *epsilon-SVM regression* e *nu-SVM regression*, respectivamente.

A função de erro tipo I: $\frac{1}{2}w^T + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^*$, foi sujeita as seguintes restrições:

$$\begin{aligned} w^T \phi(x_i) + b - y_i &\leq \varepsilon + \xi_i^* \\ y_i - w^T \phi(x_i) + b_i &\leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* &\geq 0, i = 1, \dots, N \end{aligned}$$

Enquanto que a função de erro tipo II: $\frac{1}{2}w^T w - C \left(v\varepsilon + \frac{1}{N} \sum_{i=1}^N (\xi_i + \xi_i^*) \right)$, ficou sujeita as seguintes restrições:

$$\begin{aligned} (w^T \phi(x_i) + b) - y_i &\leq \varepsilon + \xi_i \\ y_i - (w^T \phi(x_i) + b_i) &\leq \varepsilon + \xi_i^* \\ \xi_i^*, \xi_i &\geq 0, i = 1, \dots, N, \varepsilon \geq 0 \end{aligned}$$

em que: C é a constante de capacidade; w é o vetor de coeficientes; b é uma constante;

ξ_i são os parâmetros; i é o índice que representa os casos de treinamentos; N é o número total de casos; e ϕ representa o *kernel* utilizado.

Testaram-se também os quatro tipos de função *kernel*: Linear, Polynomial, RBF (Função de Base Radial) e Sigmoid (Tabela 2). A construção de todas MVS citadas foi realizada no Statistic 12[®].

Tabela 2 – Tipos de funções de *kernel* testadas.

Tipo de Kernel	Função $K(X_i, X_j)$
Linear	$X_i \cdot X_j$
Polinomial	$(\gamma X_i \cdot X_j + C)^d$
RBF	$\exp(-\gamma X_i - X_j ^2)$
Sigmoid	$\tanh(\gamma X_i \cdot X_j + C)$

Em que: $K(X_i, X_j) = \phi(X_i) * \phi(X_j)$.

3.3.3. Redes Neurais Artificiais (RNA)

Treinaram-se as redes utilizando a metodologia desenvolvida por BINOTI (2010). Os dados, englobando todas as medições disponíveis, foram divididos aleatoriamente em dois conjuntos, um para o treinamento das redes e outro para avaliar a generalização das redes treinadas com uma proporção de 70% dos dados para treinamento e 30% para generalização.

Assim como na metodologia aplicada para as MVS, os valores que extrapolaram a tendência geral de cada projeto, em área basal ou volume, não foram eliminados a fim de verificar a capacidade das redes em lidar com *outliers* e ruídos.

Foram treinadas redes para estimar a produção e também para estimar a área basal, pois além deste valor ser considerado como input para projetar a produção, este resultado foi comparado com o resultado do método indireto obtido com a MVS. As variáveis de input seguiram a mesma descrição do modelo utilizado pela MVS. Treinaram-se 20 redes do tipo Perceptrons de Múltiplas Camadas (MLP); das quais foram selecionadas as 5 melhores, com base no coeficiente de correlação entre os valores observados e estimados e na distribuição gráfica dos resíduos.

A construção de todas as redes citadas, bem como a aplicação das redes treinadas aos dados de generalização, ou seja, aqueles que não foram utilizados na etapa do treinamento, foi realizada no Statistic 12[®].

3.4. Avaliação e aplicação das alternativas de modelagem

A seleção da melhor metodologia (Regressão, MSV ou RNA), ou seja, aquela que fornecesse as estimativas mais precisas de produção, baseou-se na magnitude e na distribuição gráfica dos resíduos (CAMPOS e LEITE, 2013). Para determinar a exatidão das estimativas, foram calculadas as estatísticas: coeficiente de determinação ($R^2\%$), coeficiente de determinação ajustado ($\bar{R}^2\%$), coeficiente de correlação ($r_{y\hat{y}}$) e erro padrão residual ($S_{y\hat{y}}$) (CAMPOS e LEITE, 2013).

O coeficiente de determinação (R^2) representa uma medida do grau de ajustamento dos dados. Deve ser interpretado como a proporção de variação total da variável dependente Y que é explicada pela reta de regressão, sendo válido apenas para equações que apresentem β_0 . Enquanto que o \bar{R}^2 deve ser compreendido como o valor de R^2 corrigido com relação ao número de parâmetros (GUJARATI, 2006).

$$R^2 = \frac{\text{variação explicada}}{\text{variação total}} \implies R^2\% = \left[1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2} \right] * 100$$
$$\bar{R}^2\% = \left[1 - \left(\frac{n-1}{n-p} \right) * \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2} \right] * 100$$

O coeficiente de correlação entre os valores observados e estimados é uma medida do grau de relação linear entre essas duas variáveis quantitativas (GUJARATI, 2006). Apesar de não se poder inferir sobre a igualdade entre os valores estimados e observados, juntamente com a análise de resíduos, esta análise permite inferir sobre a qualidade da prognose ou estimativa (CAMPOS e LEITE, 2013).

$$r_{y\hat{y}} = \frac{\text{cov}(y, \hat{y})}{\sqrt{s^2(Y) s^2(\hat{y})}}$$

em que, s^2 é a variância e cov a covariância. Os limites de r são -1 e $+1$; isto é r pertence ao intervalo $[-1, +1]$. Se $r_{y\hat{y}} = +1$, a correlação linear entre as variáveis é considerada como perfeita positiva; caso $r_{y\hat{y}} = -1$, a correlação é classificada como perfeita negativa; porém se $r_{y\hat{y}} = 0$, não existe correlação linear entre as variáveis.

O erro padrão da estimativa é uma medida das distâncias, ou seja, das diferenças entre os valores amostrais y observados e os valores preditos \hat{y} obtidos através da reta de regressão, da MVS ou da RNA.

$$S_{y\hat{y}} = \sqrt{\frac{\sum(y - \hat{y})^2}{n - p - 1}}$$

em que: Y é o valor observado em campo nos inventários; \bar{Y} é a média dos valores observados; \hat{Y} é o valor estimado n é o número de observações; e p é o número de variáveis independentes.

A análise gráfica consistiu na inspeção estatística da dispersão dos erros (resíduos) percentuais em relação aos valores observados, sendo:

$$Erro_{\%} = \frac{(\hat{y} - y)}{y} 100$$

4. RESULTADOS E DISCUSSÃO

Os modelos de crescimento e produção devem ser suficientemente confiáveis para descreverem irrefutavelmente a dinâmica do complicado sistema não-linear que é o crescimento de um povoamento florestal (PORTÉ e BARTELINK, 2002). Isto porque na prática, a precisão desses modelos é um fator que pode negativamente influenciar nas estimativas de produção, e conseqüentemente, no planejamento florestal (BETTINGER et al., 2009).

O tipo proposto por Clutter em 1963 é o modelo mais difundido nas empresas florestais brasileiras (CAMPOS e LEITE, 2013). Nele a produção é derivada pela integração matemática do modelo de crescimento, o que lhe confere sua principal característica: a compatibilidade. Neste trabalho, a produção foi estimada pelo Modelo de Clutter na sua forma completa (2). Conforme recomendação de SALLES (2010) e ALCÂNTARA (2012), avaliou-se a contribuição estatística das variáveis nesse sistema de equações; e diante da inconsistência de seus parâmetros, as variáveis $S_1 I_2^{-1}$ e $LnB_2 I_2^{-1}$ foram eliminadas. Os demais parâmetros apresentaram valor de $p < 0,001$.

$$LnB_2 = LnB_1(I_1 I_2^{-1}) + 4,323606(1 - I_1 I_2^{-1}) + 0,021744(1 - I_1 I_2^{-1})S_1 \quad R^2 = 94,43$$

$$LnV_2 = 3,904683 + 17,998810I_2^{-1} - 0,075513S_1 + 0,361053LnB_2 + 0,033308S_1 LnB_2 \quad R^2 = 93,38$$

Observa-se que a projeção da produção pelo método de Clutter resultou em estimativas não tendenciosas. Foi estimado um coeficiente de correlação de 0,97 e um coeficiente de determinação de 94,43; com 80% dos resíduos concentrando-se entre a amplitude das classes de $\pm 10\%$ de erro (Figuras 1 e 2).

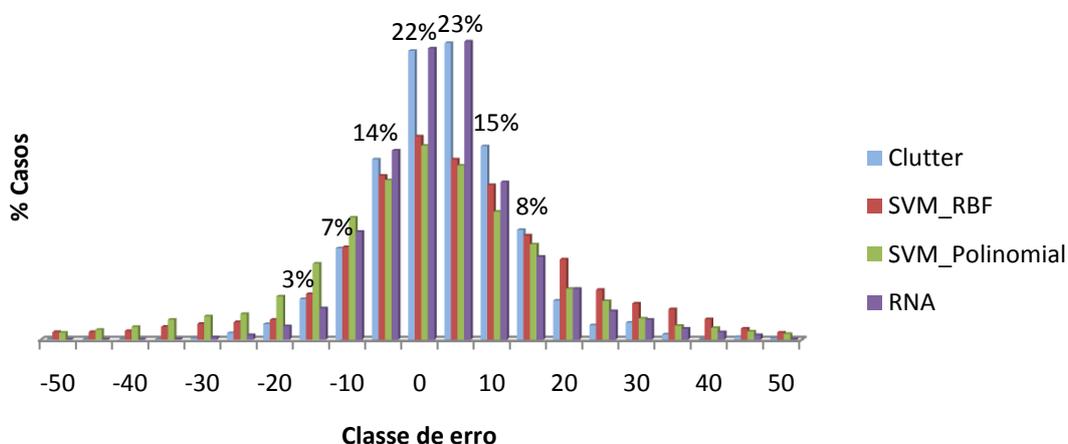


Figura 1 – Distribuição por classe de erro dos resíduos (percentagem de casos) gerados pelas três metodologias para estimativa da produção.

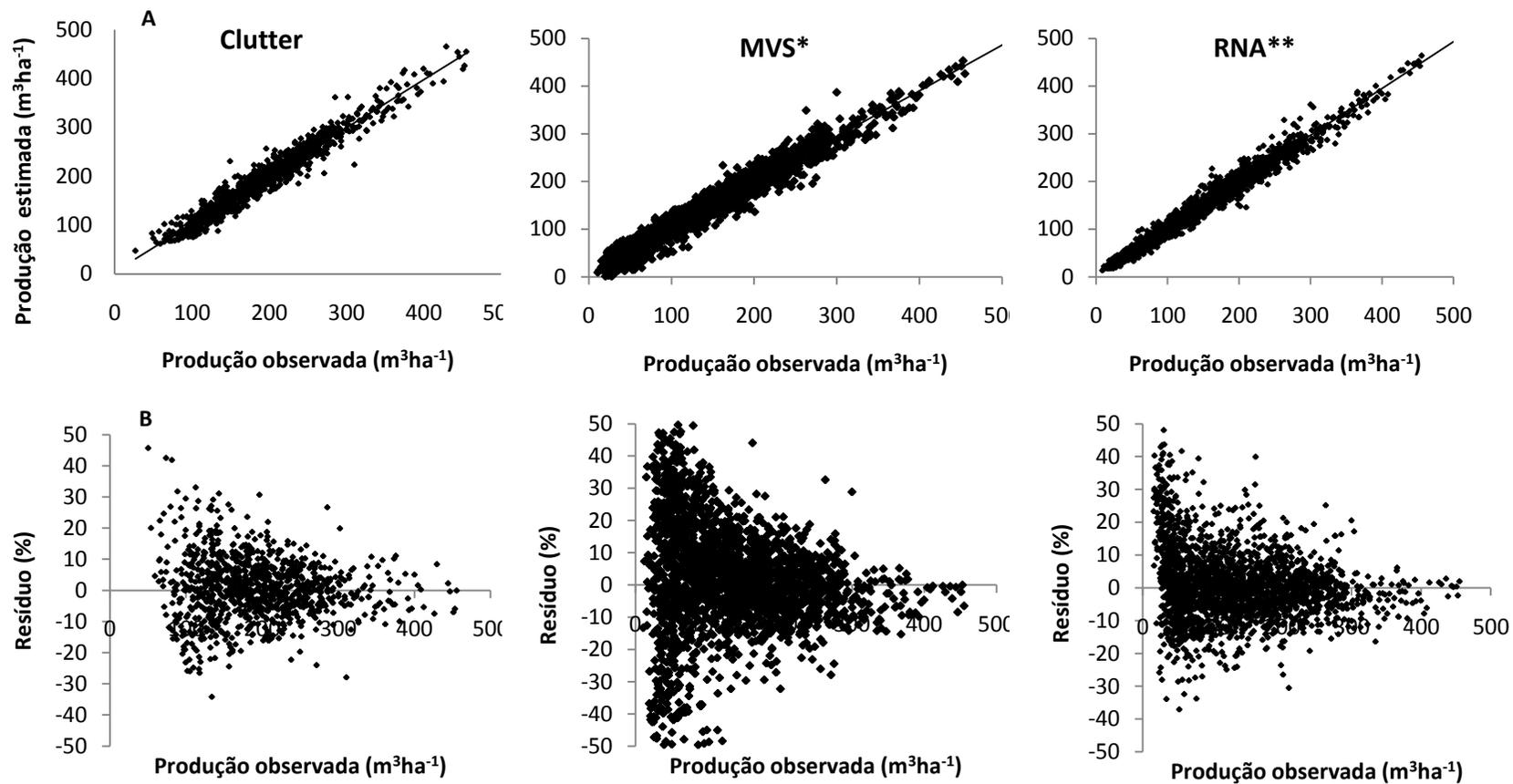


Figura 2 – A) Distribuição gráfica da produção observada em campo e a estimada por diferentes metodologias; e B) Distribuição gráfica dos resíduos obtidos por diferentes metodologias de estimativa da produção.

* Máquina de Vetores Suporte com *kernel* RBF e função de erro tipo 2.

** Rede Neural Artificial do tipo Perceptron da Múltiplas Camadas (MLP).

Uma desvantagem em se utilizar o modelo de regressão para a modelagem florestal é que as informações provenientes de povoamentos florestais clonais apresentam características que violam algumas pressuposições estatísticas (GARCÍA, 1988). Como exemplo, cita-se a ocorrência de heterocedasticidade da variância e autocorrelação residual, que estão associadas às medições repetidas das variáveis ao longo do tempo e à idade dos povoamentos (CALEGARIO et al., 2005; OLIVEIRA, 2012).

Ademais, a necessidade de estratificação do povoamento pode também proporcionar uma ineficiência ao modelo de regressão. Outra particularidade é a amplitude entre as medições de idade (I_2-I_1), se constante em todo o banco de dados, o ajuste do modelo pelo método de mínimos quadrados em dois ou três estágios torna-se impossível (CAMPOS e LEITE, 2013).

Uma alternativa para esses casos é o emprego de técnicas de inteligência computacional (RNA e MVS). Neste estudo, a aplicação dos algoritmos de aprendizado de máquina demonstrou-se como uma eficiente ferramenta para estimar tanto a produção em volume e quanto em área basal. Conclusão similar foi obtida por ÖZÇELİK et al. (2010) ao analisar o desempenho de redes neurais artificiais para estimar o volume do fuste de quatro diferentes espécies na Turquia.

Semelhante ao modelo de crescimento, estas metodologias também apresentaram resultados sem tendenciosidade e livre de bias (Figura 2). Diante dos dois métodos testados para ajustar as MVS, observou-se que a produção estimada a partir dos dados de área basal estimadas, ou seja, o método indireto apresentou melhores resultados. Comparando-se as duas funções de erro, a função tipo II proporcionou maiores valores de coeficiente de correlação. Com relação aos *kernels*, observa-se que a função Sigmoid foi a única que não resultou em boas estimativas (Tabela 2).

Tabela 2 – Estatísticas relacionadas aos resultados das produções estimadas por diferentes Máquinas de Vetores de Suporte (MVS).

Funções <i>kernel</i>	Função de Erro Tipo 1					Função de Erro Tipo 2				
	R^2 (%)	\bar{R}^2 (%)	S_{yx}	$r_{y\hat{y}}$	Nº Vetores suporte	R^2 (%)	\bar{R}^2 (%)	S_{yx}	$r_{y\hat{y}}$	Nº Vetores suporte
Método Direto										
Linear	98,63	98,63	19,61	0,98	60	98,00	97,99	22,63	0,97	85
Polinomial	99,27	99,27	14,05	0,99	89	99,82	99,82	6,74	0,99	234
RBF	99,41	99,41	12,67	0,99	77	99,76	99,76	7,78	0,99	221
Método Indireto										
Linear	92,80	92,76	22,65	0,94	37	89,40	89,35	27,48	0,89	222
Polinomial	96,02	96,00	16,84	0,96	112	95,66	95,64	17,58	0,95	260
RBF	96,29	96,27	16,26	0,96	67	95,91	95,89	17,07	0,96	210

Dentre todas as MVS construídas, a que utilizou o método indireto com a função *kernel* RBF explicitou melhor resultado, haja vista o elevado valor de coeficiente de correlação (0,96) e, aproximadamente, 60% dos resíduos estarem distribuídos entre a amplitude de classes de ± 10 . Esta solução seguiu a mesma tendência daquela com o modelo de Clutter na sua forma completa (Figura 8).

Assim como para MVS, o método indireto de estimativa da produção com a RNA apresentou mesmo desempenho, pois também resultou em estimativas mais precisas. Foram testadas 20 diferentes arquiteturas da rede Perceptrons de Múltiplas Camadas (MLP), selecionando-se a MLP 143-14-1 devido a superioridade de seus resultados. Diferentemente da distribuição dos resíduos obtidos com a MVS, aproximadamente 60% dos resíduos obtidos pela RNA encontraram-se distribuídas entre as amplitudes de classes de ± 5 .

Tabela 3 – Estatísticas relacionadas com os resultados das produções estimadas por diferentes Redes Neurais Artificiais (RNA).

Arquitetura	R^2 (%)	\bar{R}^2 (%)	S_{yx}	$r_{y\hat{y}}$
Método Direto				
1) MLP 145-10-1	99,95	99,95	3,69	0,99
2) MLP 145-14-1	99,94	99,94	3,78	0,99
3) MLP 145-13-1	99,95	99,95	3,57	0,99
4) MLP 145-9-1	99,94	99,94	3,87	0,99
5) MLP 145-4-1	99,94	99,94	3,86	0,99
Método Indireto				
1) MLP 143-7-1	99,36	99,35	12,80	0,99
2) MLP 143-9-1	99,40	99,39	12,41	0,99
3) MLP 143-14-1	99,43	99,43	12,05	0,99
4) MLP 143-12-1	99,37	99,37	12,65	0,99
5) MLP 143-5-1	99,37	99,37	12,66	0,99

Uma característica importante das técnicas de aprendizado de máquina é a capacidade de lidar com dados imperfeitos, também chamados de ruídos; e de minimizar os chamados *outliers*, ou seja, exemplos isolados que destoam dos demais de uma determinada classe e que podem ocorrer devido a presença de ruídos (CORTE e VAPNIK, 1995; HAYKIN, 2001; LORENA, 2006). Esses atributos foram comprovados neste estudo, uma vez que os dados imperfeitos não foram eliminados no processo de modelagem pelas técnicas MVS e RNA.

Apesar de alguns dos resultados obtidos com a RNA apresentarem maior precisão, é importante ressaltar a influência do procedimento de separação do banco de dados na etapa de treinamento das redes. Esta metodologia permitiu que um número maior de redes fosse construído, podendo-se assim, selecionar aquelas que permitissem melhor resultado. No entanto, apenas uma MVS foi construída para cada tipo de *kernel* e função de erro.

A técnica de MVS se destaca também pela elevada robustez em exemplos com grandes dimensões, devido a convexidade de sua função objetivo (CORTE e VAPNIK, 1995). Isto porque a aplicação da MVS implica na otimização de uma função quadrática a qual possui apenas um mínimo global (SMOLA e SCHOLKOPF, 2004). Esta é uma vantagem sobre, por exemplo, as Redes Neurais Artificiais, na qual há a presença de mínimos locais na função objetivo a ser minimizada (HAYKIN, 2001).

No tocante ao setor florestal, a imensa vantagem na utilização das técnicas de inteligência artificial para projetar a produção é a possibilidade de inserir dados categóricos aos modelos, como informações sobre o tipo de solo, bacia hidrográfica, tipo de mecanização, material genético, rotação, entre outras. A incorporação dessas variáveis ambientais permite maior flexibilidade ao modelo de projeção, influenciando assim, no aumento de precisão das estimativas.

Uma consequência positiva da plasticidade permitida pelos algoritmos de aprendizado de máquina está no planejamento do setor florestal. A simulação de diferentes cenários assegura a tomada de decisão dos gestores florestais, além de permitir uma melhor compreensão dos impactos das alternativas de manejo sobre a estrutura do povoamento. Por exemplo, no método de regressão, as variáveis de entradas são apenas aqueles dados contínuos auferidas em campo, como: diâmetro a altura do peito (DAP), altura e idade. Caso aconteça alguma alteração no ambiente

onde foram coletadas, não será possível inseri-la ao modelo; situação esta que pode ser contornada pelos algoritmos computacionais.

Outro fator relevante é a duração para a modelagem do crescimento e da produção do povoamento. Comparando-se as três técnicas, a MVS proporcionou uma redução no tempo de processamento desses dados. Este resultado pode impulsionar o surgimento de novas ferramentas computacionais para o setor de manejo florestal, as quais além de agilizar o trabalho do gestor, concedem também maior precisão as estimativas. Porquanto, mais estudos devem ser realizados, principalmente no tocante a parametrização desta técnica, bem como sua aplicação em outros problemas da área florestal.

5. CONCLUSÕES

A comparação entre a precisão das estimativas obtidas por diferentes técnicas de modelagem do crescimento e da produção (Modelo de Crescimento, Redes Neurais Artificiais e Máquinas de Vetores Suporte) permite inferir que a projeção da produção do povoamento de eucalipto utilizando os algoritmos computacionais não só diferem daquela projeção obtida pelo método de regressão linear, como também são eficientes ferramentas para estudos de projeção da produção, com a vantagem de permitirem agregar dados categóricos aos modelos. Neste contexto, destaca-se a Máquina de Vetores Suporte pela alta precisão das estimativas e menor tempo de processamento dos dados da fazenda modelo.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALCÂNTARA, A. E. M. de. **Alternativas de modelagem para projeção do crescimento de eucalipto em nível de povoamento**. Viçosa: UFV, 2012. 66p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

BETTINGER, P.; BOSTON, K.; SIRY, J. P.; GREBNER, D. L. **Forest Management and Planning**. Elsevier, 2009. 342p.

BINOTI, M. L. M. S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**. Viçosa: UFV, 2010. 54 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.

BINOTI, D. H. B. **Sistemas computacionais aplicados ao manejo florestal**. Viçosa: UFV, 2012. 127p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

BINOTI, M. L. M. S. **Emprego de redes neurais artificiais em mensuração e manejo florestal**. Viçosa: UFV, 2012. 136p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

BUONGIORNO, J.; GILLESS, J. K. **Decision Methods for Forest Resource Management**. Elsevier, 2003. 458p.

CALEGARIO, N.; CALEGARIO, C. L. L.; MAESTRI, R.; DANIELS, R. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não lineares generalizados. **Scientia Florestalis**, n. 69, p. 38-50, dez. 2005.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. Viçosa: Editora UFV, 2013. 605 p.

CASTRO, R. V. O. **Modelagem do crescimento em nível de árvores individuais utilizando redes neurais e autômatos celulares**. Viçosa: UFV, 2011. 90p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2011.

CLUTTER, J. L.; FORTSON, J. C.; PIENAAR, L. V.; BRISTER, G. H.; BAILEY, R. L. **Timber Management: a quantitative approach**. New York: J. Wiley, 1983. 333p.

CORTE, C.; VAPNIK, V. Support-Vector Networks. **Machine Learning**, n. 20, p.273-297, 1995.

CRISCI, C.; GHATTASB, B.; PERERAC, G. A review of supervised machine learning algorithms and their applications to ecological data. **Ecological Modelling**, n.240, p.113-122, 2012.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. New York: John Wiley & Sons, 1966. 407p.

MINAS GERAIS. Secretaria de Meio Ambiente e Desenvolvimento Sustentável. Fundação Estadual do Meio Ambiente. **Mapa de solos do Estado de Minas Gerais: legenda expandida**. Belo Horizonte, 2010. Disponível em: <<http://www.feam.br/noticias/1/949-mapas-de-solo-do-estado-de-minas-gerais>> Acesso em: 13 out. 2013.

GARCÍA, O. Growth modeling: a (re)view. **New Zealand Forestry**, v.33, n.3, p.14-17, 1988.

GORGENS, E. B. **Estimação do volume de árvores utilizando Redes Neurais Artificiais**. Viçosa: UFV, 2006. 94p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2006.

GUJARATI, D. N. **Econometria Básica**. São Paulo: Makron Books, 2006. 846p.

HAYKIN, S. **Redes Neurais: princípios e prática**. Porto Alegre, 2001. 900 p.

MINISTÉRIO DA AGRICULTURA, PECUÁRIA e ABASTECIMENTO. Instituto Nacional de Meteorologia. **Base de dados climatológica**. 2013. Disponível em: <<http://www.inmet.gov.br/portal/index.php?r=clima/mesTempo>> Acesso em: 15 out. 2013.

KING, S. L.; BENNETT, K. P.; LIST, S. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. **Computers and Electronics in Agriculture**, n. 27, p.401-406, 2000.

LORENA, A. C. **Investigação de estratégias para a geração de máquinas de vetores suporte multiclases**. São Paulo: USP, 2006. 229p. Tese (Doutorado) – Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2006.

McQUEEN, R. J.; GAMER, S. R.; NEVILL-MANNING, C. G.; WITTEN, I. H. Applying machine learning to agricultural data. **Computers and Electronics in Agriculture**, n. 12, p. 275-293, 1995.

OLIVEIRA, J. de C. **Mínimos quadrados generalizados: uma proposta à modelagem do crescimento e produção em povoamentos não desbastados de eucalipto** Viçosa: UFV, 2012. 80p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

ÖZÇELİK, R.; DIAMANTOPOULOU, M. J.; BROOKS, J. R.; WIANZ JUNIOR, H.V. Estimating tree bole volume using artificial neural network models for four species in Turkey. **Journal of Environmental Management**, n. 91, 2010, p. 742-753.

PORTÉ, A.; BARTELINK, H. H. Modelling mixed Forest growth: a review of models for forest mangement. **Ecological Modelling**, v.150, p. 141-188, 2002.

RECKNAGEL, F. Applications of machine learning to ecological modeling. **Ecological Modelling**, n. 146, p. 303-310, 2001.

RESENDE, R. R.; DO VALE, A. B.; SOARES, T. S.; DA SILVA, M. L.; COUTO, L.; DO VALE, R. S. Emprego de um modelo de crescimento e produção para determinação de rotação em povoamentos de eucalipto. **Revista Árvore**, Viçosa-MG, v.28, n.2, p.219-225, 2004.

SALLES, T. T. **Modelagem do crescimento e produção de clones de eucalipto em um sistema agroflorestal**. Viçosa: UFV, 2010. 87p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Ajuste do modelo de Schumacher e Hall e aplicação de redes neurais artificiais para estimar volume de árvores de eucalipto. **Revista Árvore**, v.33, n.6, p.1133-1139, 2009.

SMOLA, A. J.; SCHOLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, n.14, p.199–222, 2004.

SOUZA, M. J. H.; RIBEIRO, A.; LEITE, F. P. Balanço hídrico e caracterização climática de Guanhões, Nova Era e Rio Doce. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 13., 2003, Santa Maria. **Anais...** Santa Maria: UNIFRA, SBA, UFSM, 2003, v 2., p.131-132.

VANCLAY, J. K. Growth models for tropical forests: a synthesis of models and methods. **Forest Science**, v.41, p. 7-42, 1995.

YOUNG, R. A.; GIESE, R. L. **Introduction to Forest Ecosystem Science and Management**. Elsevier, 2003. 589 p.

ZHENG, H.; LU, H. A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). **Computers and Electronics in Agriculture**, n.83, p. 47–51, 2012.

CAPÍTULO 3: EFEITOS DE DIFERENTES ALTERNATIVAS DE MODELAGEM DO CRESCIMENTO E DA PRODUÇÃO SOBRE A REGULAÇÃO FLORESTAL

RESUMO

Este trabalho teve como objetivo a aplicação do software RPF para o estudo de caso de um problema de regulação florestal com tabelas de produção obtidas por três diferentes metodologias para estimar a produção: Modelo de Clutter, Máquina de Vetores Suporte (MVS) e Redes Neurais Artificiais (RNA). Utilizaram-se dados de uma fazenda modelo, cuja área total era de 16057.5 ha, distribuídos em 71 unidades de manejo. O horizonte de planejamento foi de 18 anos e a taxa de descontos de 8.5% ao ano, com rotação regulatória permitida aos 6 anos. Testaram-se como função objetivo: minimizar o custo médio de produção, minimizar o custo total, maximizar a produção e maximizar a receita líquida. A produção total foi limitada entre 700.000 e 1.000.000.000m³ por ano. Os resultados deste estudo demonstram a eficiência do sistema RPF 2.0[®], o qual se apresenta bastante útil aos gestores no processo de construção e interpretação de modelos de regulação da produção florestal. Observou-se que as diferentes tabelas de produção interferiram nas prescrições geradas, principalmente a obtido pelo Modelo de Clutter, no tocante a produção total anual e rendimento acumulado ao longo do horizonte de planejamento.

Palavras-chaves: Regulação florestal, Manejo florestal, Produção florestal

1. INTRODUÇÃO

Entende-se como gestão o ato de gerenciar, administrar uma instituição (LACOMBE, 2004). Em outras palavras, conceituar-se-ia como a otimização do funcionamento e desenvolvimento das instituições, baseando-se em decisões racionais, fundamentadas em informações relevantes. Neste contexto, definiu-se como gestão dos recursos florestais a arte ou a ciência de tomar decisões no que se refere à organização, ao uso e à conservação das florestas e de seus recursos (BUONGIORNO e GILLESS, 2003). Para desempenhar as funções administrativas, os gestores florestais recorrem a outras áreas científicas como economia, matemática e estatística. Tais disciplinas direcionam a tomada de decisão a longo prazo, ou seja, o planejamento estratégico dos empreendimentos florestais.

A ideia da estrutura regulada de uma floresta se insere no planejamento estratégico, pois é uma atividade a longo prazo. Regular uma floresta significa determinar onde, como e quando produzir bens e serviços da floresta de modo sustentável para melhor alcançar os objetivos do empreendimento (DAVIS e JHONSON, 1987). A regulação florestal permite melhor definição das práticas silviculturais a serem aplicadas em cada compartimento da floresta, constituindo-se assim um plano de manejo. A principal vantagem gerada por uma floresta regulada são produções periódicas de igual volume, tamanho e qualidade; conseqüentemente, proporcionando a sustentabilidade da floresta (LEUSCHNER, 1984).

O processamento dos modelos de regulação da produção é realizado utilizando-se métodos numéricos, como: programação linear (PL), programação multiobjetivo (PMO), programação inteira (PI), programação inteira mista (PIM), programação dinâmica (PD), meta-heurísticas e simulação (BINOTI, 2010). Atualmente, diversos softwares relacionados a manejo podem ser encontrados para auxiliar na solução dos modelos de regulação florestal (Forplan – Forest Planning, Planflor, *SysFlor*, SifPlan, Remsoft, OpTimber-LP e RPF – Regulação da Produção Florestal).

Ressalta-se o programa RPF[®], haja vista ser este gratuito, versátil e fácil de usar. Tais características foram testadas por CARVALHO (2012) para analisar a influência de diferentes variáveis econômicas em modelos de regulação florestal, utilizando dados de 10 florestas-modelo, cada uma com 100 unidades de manejo. Este autor relatou a eficiência para a resolução dos modelos de regulação formulados de acordo com o modelo I da programação linear. Dentre as diferentes funções

objetivo estudadas, o software mostrou que as de minimização tiveram a produção volumétrica estabilizada a partir do horizonte de 16 anos.

Diante desse contexto, pretende-se com este estudo analisar os efeitos de diferentes alternativas de modelagem do crescimento e da produção sobre a regulação florestal de uma fazenda modelo; testando também o desempenho do software RPF para processamento dos modelos de regulação.

2. HIPÓTESES

H'_0 : As diferentes tabelas de produção interferem no valor de rendimento acumulado numa floresta regulada após horizonte de planejamento.

H'_1 : Não H'_0 .

3. METODOLOGIA

3.1. Estudo de caso

A fazenda modelo citada no capítulo 2 apresenta uma área total de 16057.5 ha, distribuídos em 71 unidades de manejo com diferentes dimensões. Para a formulação dos modelos de regulação, foram construídas tabelas de produção obtidas com as três alternativas utilizadas de prognose da produção: Máquina de Vetores Suporte, Redes Neurais Artificiais e Modelo de Clutter, conforme capítulo 2 (Figura 1). As tabelas de produção foram geradas por unidade de manejo (talhão). Para cada UM foram geradas todas as possíveis prescrições; ou seja, uma sequência de ações (corte ou corte com reforma imediata ou corte e venda de madeira ou corte e condução ou reforma ou condução ou desbaste ou fertilização ou desrama) estabelecidas para um horizonte de planejamento.

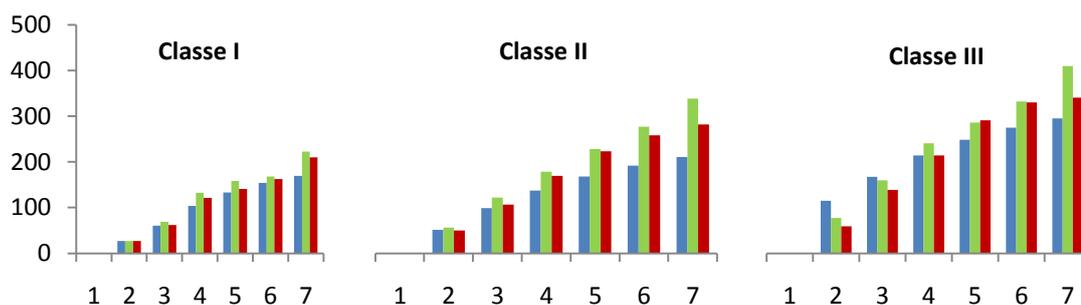


Figura 1 – Produção estimada ($m^3 ha^{-1}$) através das três metodologias* proposta para prognose da produção por idade (anos), distribuídos por classe.

* ■ Modelo de Clutter, ■ Rede Neural Artificial e ■ Máquina de Vetores

3.2. Formulação e Resolução do Modelo de Regulação

Considerou-se na formulação que uma vez assinalada a uma unidade de manejo, a prescrição se manteria até o fim do horizonte de planejamento. Para as análises econômicas, adotou-se uma matriz de custos (Tabela 1), juntamente com os dados de preço de venda da madeira (R\$ 52,00/m³) e o valor da venda da terra (R\$ 4.500,00/ha). Como parâmetros do modelo, considerou-se uma taxa de juros de 8,5% ao ano (i_{aa}) em um horizonte de planejamento (H) de 18 períodos (*anos*), uma rotação regulatória aos 6 anos (R) e uma variação na área de corte de 5%.

Tabela 1 – Descrição dos custos florestais por ano de ocorrência e regime de manejo.

Descrição	Ano de ocorrência	Alto fuste Valor (R\$/ha)	Talhadia Valor (R\$/ha)
Insumos	0	1.939,54	1.550,69
Assistência técnica	0	130,11	130,11
Implantação/condução	0	1050,00	434,00
Manutenção 1	1	240,00	240,00
Manutenção 2	2	210,00	210,00
Custo anual da terra	0-n	120,00	120,00
Colheita (R\$/m ³)*	n	15,42	15,42
Transporte (R\$/m ³ /km)*	n	0,23	0,23

De posse das tabelas de produção geradas por talhão, da planilha de custos, do preço da madeira, da definição da rotação regulatória, da definição do horizonte de planejamento, gerou-se e processou-se o modelo de regulação florestal em formato de programação linear inteira utilizando o software RPF 2.0[®] (BINOTI, 2012).

3.2.1. Função Objetivo

Além de minimizar o custo médio de produção (A), conforme proposto por CARVALHO (2012), considerou-se como objetivo do empreendedor: minimizar o custo total (B), maximizar a produção (C) e maximizar a receita líquida (D). Cabe ressaltar que quando se maximiza a produção, pretende-se o valor ótimo de produção possível para a floresta, porém sem desrespeitar as restrições do modelo.

O Custo Médio de Produção (CMP) é o resultado da relação entre o custo total corrigido e a produção equivalente, ambos considerados no valor presente (SILVA et al., 2002). A função cujo objetivo é o de minimizar o somatório do CMP da floresta assume a seguinte configuração:

$$Min. Z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij};$$

em que Z é o somatório do CMP (R\$/ha); x_{ij} são variáveis de decisão, representando a j -ésima alternativa de manejo adotada na i -ésima unidade de manejo; c_{ij} é CMP de cada unidade de manejo i , manejada segundo a alternativa de manejo j ; m é o número total de unidades de manejo; n é número total de alternativas de manejo para a i -ésima unidade de manejo.

As demais funções objetivo seguiram o mesmo padrão da de minimizar o somatório do CMP, alterando o valor de Z para o somatório do custo total (R\$/ha), da produção (m^3ha^{-1}) e da receita líquida (R\$/ha).

3.2.2. Restrições

Impuseram-se restrições de área, de produção ou demanda e regulação. A restrição de área deve estar presente em todos os modelos de regulação da produção florestal, pois a mesma impede que a soma das áreas da cada prescrição exceda a área da unidade de manejo (BINOTI, 2012). É importante que áreas aproximadamente iguais sejam colhidas todos os anos para garantir a sustentabilidade do projeto. A definição do limite máximo e mínimo das áreas de corte é dada por:

$$\sum_{i=1}^m \sum_{j=1}^n a_k x_{ij} \geq R_{min_k}, \text{ com } k = 0, 1, \dots, H - 1;$$

$$\sum_{i=1}^m \sum_{j=1}^n a_k x_{ij} \leq R_{max_k}, \text{ com } k = 0, 1, \dots, H - 1;$$

em que X_{ij} são variáveis de decisão, representando a j -ésima alternativa de manejo adotada na i -ésima unidade de manejo; m é o número total de unidades de manejo; n é número total de alternativas de manejo para a i -ésima unidade de manejo; a_k valor 1 se a variável de decisão apresentar colheita no período k , valor 0 caso a variável de decisão não apresentar colheita no período k ; R_{min_k} e R_{max_k} representam o somatório das áreas disponíveis para colheita mínima e máxima respectivamente, em cada período do horizonte de planejamento.

A restrição de produção volumétrica, também conhecida como de restrição de demanda, visa limitar a produção volumétrica entre limites desejados para atendimento da demanda de consumo (BINOTI, 2012). Esta restrição é dada pela soma da produtividade das unidades de manejo multiplicada pela área assinalada. Considerou-se também que a produção de cada projeto não mudaria de um ciclo de corte para outro.

$$\sum_{i=1}^m \sum_{j=i}^n v_{ijk} x_{ij} \geq D_{min_k}, \text{ com } k = 0, 1, \dots, H - 1;$$

$$\sum_{i=1}^m \sum_{j=i}^n v_{ijk} x_{ij} \geq D_{max_k} \text{ com } k = 0, 1, \dots, H - 1;$$

em que x_{ij} são variáveis de decisão, representando a j -ésima alternativa de manejo adotada na i -ésima unidade de manejo; m é o número total de unidades de manejo; n é número total de alternativas de manejo para a i -ésima unidade de manejo; v_{ijk} é o volume (m^3) produzido por unidade de área, no período k , pela i -ésima unidade de manejo adotando-se a j -ésima alternativa de manejo; D_{min_k} e D_{max_k} são as demandas volumétricas (m^3) mínima e máxima respectivamente, em cada período do horizonte de planejamento.

A produção total foi limitada entre 700.000,00 e 1.000.000.000,00 m^3 por ano, com um manejo contemplando o corte seguido de reforma imediata da área ou condução da talhadia.

4. RESULTADOS E DISCUSSÃO

O sistema denominado RPF 2.0[®] (Regulação da Produção Florestal) foi desenvolvido utilizando-se a linguagem de programação Java, em ambiente de desenvolvimento IDE (Integrated Development Environment) Netbeans 7.0.1 e JDK 7 (Java Development Kit), com testes realizados em ambiente Windows (Figura) (BINOTI, 2012). Durante a tabulação e análise dos dados, observou-se que programa apresentava uma interface bastante amigável para os usuários. Atribuí-se esta vantagem à ampla visualização do processo matemático necessário para a geração dos modelos de regulação da produção, além da flexibilidade na simulação dos cenários.

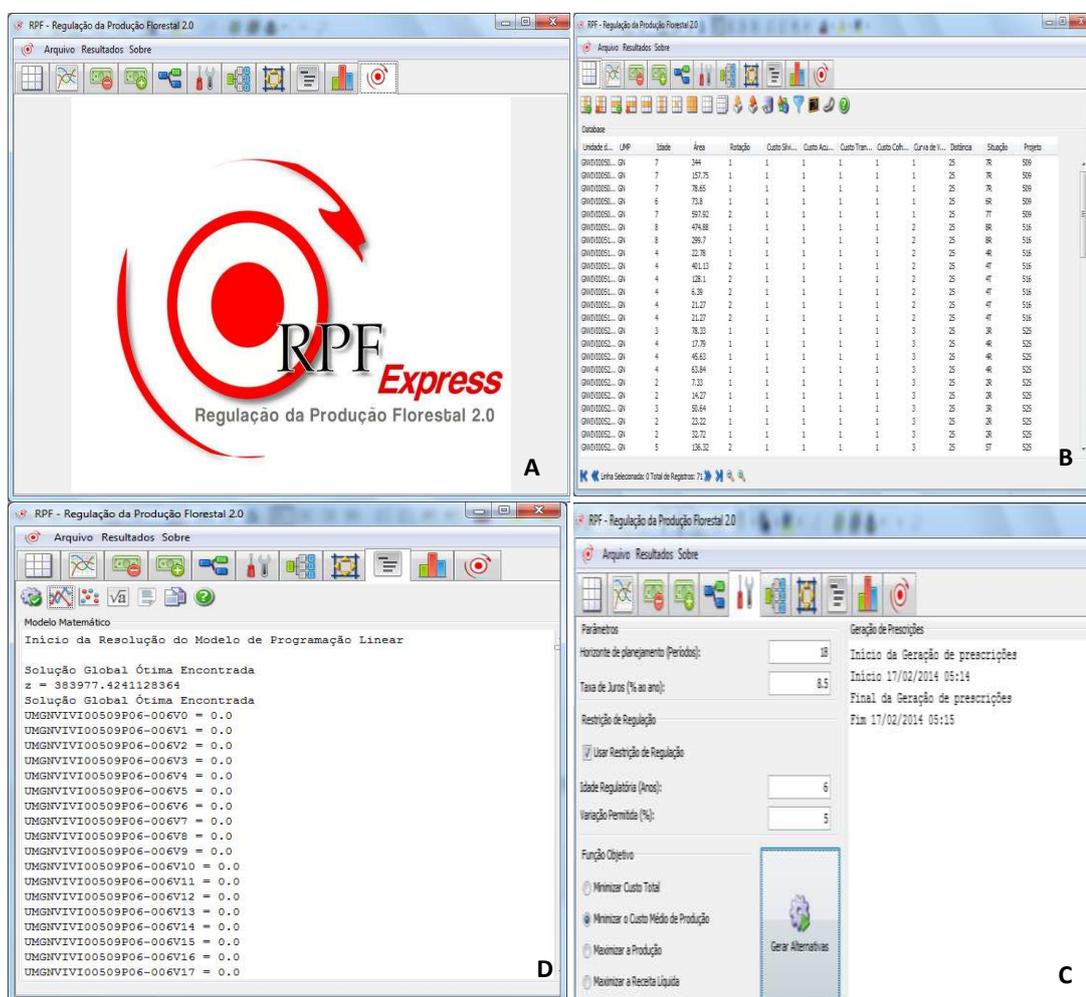


Figura 2 – A) Interface do programa Regulação da Produção Floresta (RPF 2.0[®]), evidenciando as etapas de: B) tabulação dos dados, C) formulação e D) processamento do modelo de regulação da produção florestal em uma fazenda modelo.

Após formulação do modelo de programação linear, as tabelas de produção obtidas com as três técnicas de projeção da produção Modelo de Clutter, Máquina de Vetores Suporte (MVS) e Redes Neurais Artificiais (RNA), geraram-se três planos de manejo, os quais foram ilustrados pelas tabelas de Prescrição, Idade, Rotação, Volume, Custos, Receitas e Resumo preenchidas pelo RPF. O programa permitiu visualizar as alternativas de manejo elaboradas pelo sistema (Figura 2B), bem como seus cálculos (Figura 2D). Os três planos de regulação formulados no software RPF apresentaram soluções viáveis para todos os cenários avaliados. A Figura 3 ilustra os resultados das quatro funções objetivo testadas para as tabelas de produção obtidas com os resultados das três alternativas de prognose da produção.

Primeiramente, observa-se que as tabelas de produção geradas a partir das diferentes técnicas de prognose da produção (Clutter, MVS e RNA) interferiram não só no valor de rendimento acumulado, como também na produção total anual numa floresta regulada após horizonte de planejamento. Com relação às questões de ordem da regulação da floresta, observa-se para o cenário de minimizar o custo médio de produção, uma pequena variação entre a produção total resultantes dos três planos de regulação (Figura 3A). Do segundo ao quarto ano do horizonte de planejamento, as prescrições obtidas com as curvas de produção geradas pelas técnicas de aprendizado de máquina (SVM e RNA) permitem um corte próximo ao limite superior da restrição de demanda (1.000.000,00 m³). Isto pode ser explicado devido ao povoamento na fazenda modelo apresentar classes de sítio em período de rotação, conforme Tabela 1 do capítulo 2. Contudo, este valor é muito superior ao indicado pela prescrição obtida com o Modelo de Clutter, aproximadamente 700.000 m³. Além disso, no ano de 2025, indicou-se uma produção muito abaixo do limite inferior da restrição de demanda (700.000,00 m³), o que permite inferir sobre um erro desta metodologia.

Situação semelhante se repete nos cenários para maximizar a produção e maximizar a receita líquida. A variação entre essas prescrições proporciona uma tendência dos planos gerados a partir da prognose com o Modelo de Clutter de subestimar a produção da floresta em estudo. Cabe exemplificar uma ocasião na qual o empreendedor poderia deixar de investir na compra da fazenda, caso se baseasse nas informações geradas por esta metodologia. Todavia, o fato de as técnicas de inteligência computacional, SVM e RNA, explicitarem maior precisão das

estimativas de produção, beneficiaria a tomada de decisão do gestor florestal; uma vez que indicaria a melhor opção de investimento e asseguraria a tomada de decisão.

Focando nas questões de ordem econômica, os resultados do plano gerado com as informações obtidas com Modelo de Clutter demonstraram valores subestimados da receita a ser gerada na fazenda modelo (Figura 4). Ademais, o valor acumulado final não diferiu para os quatro diferentes cenários propostos (R\$ 3.120.497,58), circunstância que não condiz com a realidade financeira de um empreendimento. Desta maneira, enfatiza-se o desempenho inferior do Modelo de Clutter aos obtidos com as técnicas de inteligência computacional, RNA e MVS.

A vantagem de inserir variáveis categóricas aos modelos de crescimento e produção, proporcionadas pela MVS e RNA, sem que houvesse perda da precisão nas estimativas, se refletiu nos resultados deste estudo, representando melhor o potencial da fazenda modelo. Consequentemente, isto influenciaria positivamente a tomada de decisão, tornando-a mais segura e confiável.

É interessante notar a superioridade dos resultados de receita acumulada prescritos pela RNA (MLP 143-14-1). Ressalta-se que as estimativas de produção obtidas por esta técnica demonstraram elevada precisão; haja vista os altos valores de coeficiente de determinação, de correlação e baixo erro padrão residual: 99.43, 0.99 e 12.05, respectivamente, calculados no capítulo 2. Todavia, enfatiza-se o fato de ter-se construído um número maior de redes que de máquinas, o que influenciou para um maior tempo computacional despendido com o processamento dos dados pela RNA.

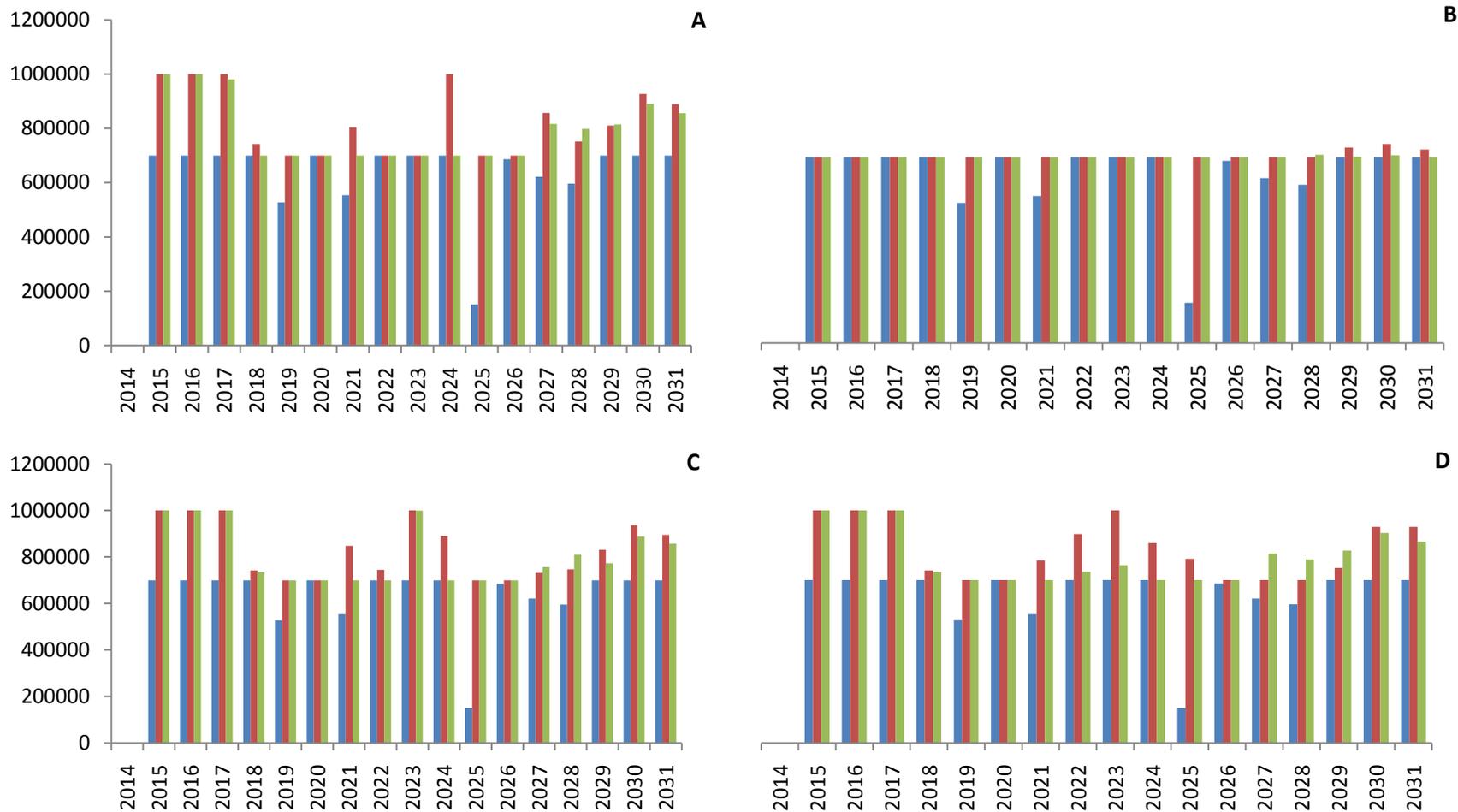


Figura 3 – Produção total prescrita pelos modelos de regulação gerados com os resultados de três metodologias* para prognose da produção, considerando-se como função objetivo: A) minimizar o custo médio de produção. B) minimizar o custo total. C) maximizar a produção e D) maximizar a receita líquida.* ■ Modelo de Clutter, ■ Rede Neural Artificial e ■ Máquina de Vetores

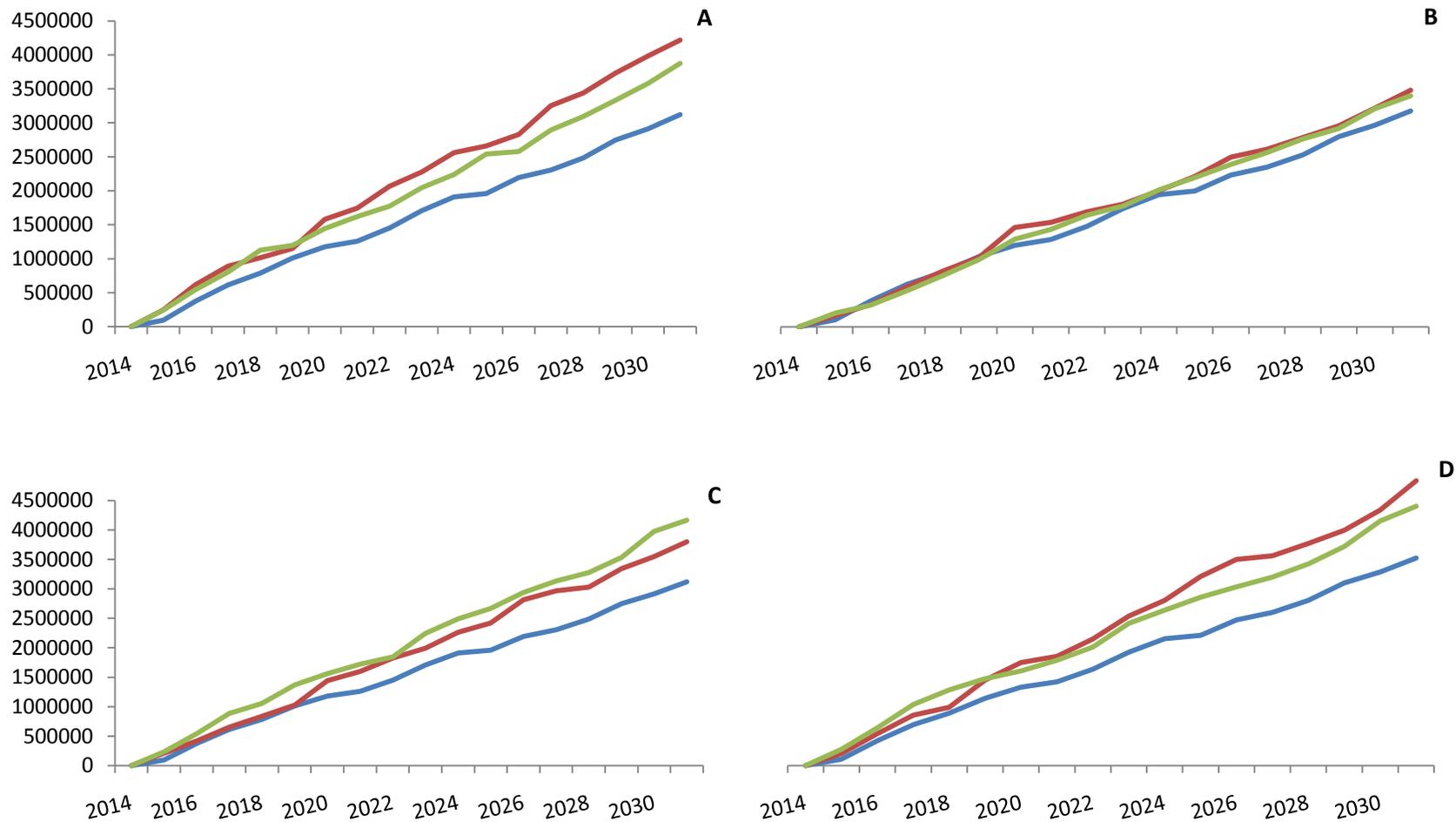


Figura 4 – Receita acumulada prescrita pelos modelos de regulação gerados com os resultados de três metodologias* para prognose da produção, considerando-se como função objetivo: A) minimizar o custo médio de produção, B) minimizar o custo total, C) maximizando a produção, D) maximizar a receita líquida. * ■ Modelo de Clutter, ■ Rede Neural Artificial e ■ Máquina de Vetores

5. CONCLUSÕES

Diante dos resultados, pode-se concluir que o software RPF[®] é eficiente e flexível para a construção de modelos e simulação de cenários de regulação da produção florestal. As tabelas de produção geradas a partir das diferentes técnicas de prognose da produção (Modelo de crescimento proposto por Clutter, Máquina de Vetores Suporte e Redes Neurais Artificiais) interferem nos resultados da floresta regulada após horizonte de planejamento. Apesar dos três planos de regulação formulados pelo RPF apresentarem soluções viáveis para todos os cenários avaliados, o modelo de Clutter não foi uma alternativa confiável, pois suas prescrições falharam ao tentar demonstrar a realidade da fazenda modelo. Idealizações. Contudo, as técnicas de inteligência artificial foram mais representativas; além de demonstrarem elevada precisão às estimativas de produção, proporcionam maior segurança a tomada de decisão do empreendedor florestal.

6. REFERÊNCIAS BIBLIOGRÁFICAS

BINOTI, M. L. M. S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto.** Viçosa: UFV, 2010. 54 p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2010.

BINOTI, D. H. B. **Sistemas computacionais aplicados ao manejo florestal.** Viçosa: UFV, 2012. 127p. Tese (Doutorado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

BUONGIORNO, J.; GILLESS, J. K. **Decision Methods for Forest Resource Management.** Elsevier, 2003. 458p.

CARVALHO, K. H. A. de. **Influência de variáveis econômicas em modelos de regulação florestal. Dissertação.** Viçosa: UFV, 2012. 114p. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência Florestal, Universidade Federal de Viçosa, Viçosa, 2012.

DAVIS, L. S.; JOHNSON, K. N. **Forest management.** New York: Mc Graw-Hill Book Company, 1987. 790p.

LACOMBE, F. **Dicionário de Administração.** Rio de Janeiro: Saraiva, 2004. 368p.

LEUSCHNER, W. A. **Introduction to forest resource management.** New York: John Willey & Sons, 1984. 298p.

4. CONCLUSÕES GERAIS

O modelo proposto através da Máquina de Vetores Suporte (MVS) é uma excelente alternativa para estudos de modelagem do crescimento e da produção florestal, cuja eficiência e vantagens foram demonstradas neste estudo; e que pode ser empregado para prognose e regulação da produção de povoamentos de eucalipto, destacando-se pela redução no tempo computacional para processamento dos dados.