

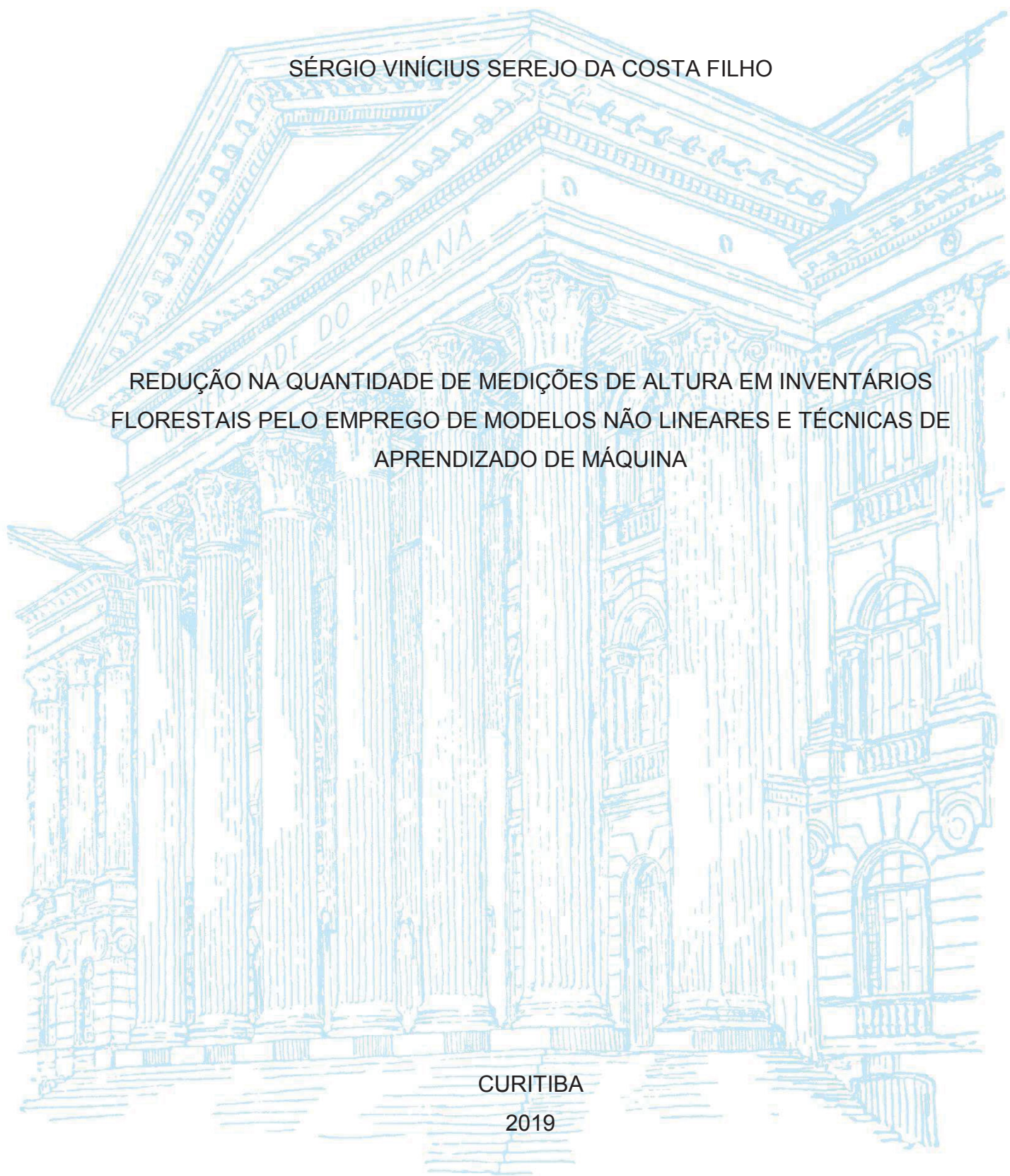
UNIVERSIDADE FEDERAL DO PARANÁ

SÉRGIO VINÍCIUS SEREJO DA COSTA FILHO

REDUÇÃO NA QUANTIDADE DE MEDIÇÕES DE ALTURA EM INVENTÁRIOS
FLORESTAIS PELO EMPREGO DE MODELOS NÃO LINEARES E TÉCNICAS DE
APRENDIZADO DE MÁQUINA

CURITIBA

2019



SÉRGIO VINÍCIUS SEREJO DA COSTA FILHO

REDUÇÃO NA QUANTIDADE DE MEDIÇÕES DE ALTURA EM INVENTÁRIOS
FLORESTAIS PELO EMPREGO DE MODELOS NÃO LINEARES E TÉCNICAS DE
APRENDIZADO DE MÁQUINA

Dissertação apresentada ao curso de Pós-Graduação em Engenharia Florestal, Setor de Ciências Agrárias, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Engenharia Florestal.

Orientador: Prof. Dr. Julio Eduardo Arce

Coorientador: Prof. Dr. Alexandre Behling

CURITIBA

2019

Ficha catalográfica elaborada pela
Biblioteca de Ciências Florestais e da Madeira - UFPR

Costa Filho, Sérgio Vinícius Serejo da

Redução na quantidade de medições de altura em inventários florestais pelo emprego de modelos não lineares e técnicas de aprendizado de máquina / Sérgio Vinícius Serejo da Costa Filho. – Curitiba, 2019.

121 f. : il.

Orientador: Prof. Dr. Julio Eduardo Arce

Coorientador: Prof. Dr. Alexandre Behling

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Ciências Agrárias, Programa de Pós-Graduação em Engenharia Florestal. Defesa: Curitiba, 25/02/2019.

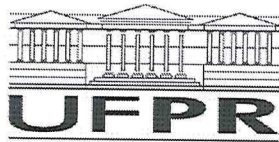
Área de concentração: Manejo Florestal.

1. Levantamentos florestais - Programas de computador. 2. Árvores - Medição. 3. Inteligência artificial. 4. Redes neurais (Computador). 5. Teses. I. Arce, Julio Eduardo. II. Behling, Alexandre. III. Universidade Federal do Paraná, Setor de Ciências Agrárias. IV. Título.

CDD – 634.9

CDU – 634.0.524.6

Bibliotecária: Berenice Rodrigues Ferreira – CRB 9/1160



MINISTÉRIO DA EDUCAÇÃO
SETOR SETOR DE CIÊNCIAS AGRARIAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA
FLORESTAL - 40001016015P0

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA FLORESTAL da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **SÉRGIO VINÍCIUS SEREJO DA COSTA FILHO** intitulada: **REDUÇÃO NA QUANTIDADE DE MEDIÇÕES DE ALTURA EM INVENTÁRIOS FLORESTAIS PELO EMPREGO DE MODELOS NÃO LINEARES E TÉCNICAS DE APRENDIZADO DE MÁQUINA**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 25 de Fevereiro de 2019.

JULIO EDUARDO ARCE

Presidente da Banca Examinadora (UFPR)

ANA PAULA DALLA CORTE

Avaliador Interno (UFPR)

RAZER ANTHOM NIZER ROJAS MONTAÑO

Avaliador Externo (UFPR)



À minha esposa Amanda e aos meus filhos Ian Caetano e Ravi,
dedico.

AGRADECIMENTOS

Agradeço aos meus avós, que em minha vida tiveram papel de pais, pela dedicação em minha criação e por tudo o que me proporcionaram, inclusive esta conquista.

À minha esposa Amanda pela compreensão, atenção e apoio incondicional nos momentos de dedicação aos estudos e a este trabalho, mas acima de tudo, pelo amor diário.

Aos meus filhos, pois admirar os filhos aquece e alimenta a alma, trazendo motivação para tudo o que se faz.

Ao meu orientador professor Julio Eduardo Arce por ter aceitado a incumbência de me orientar, mas principalmente pela sempre pronta atenção, pelas revisões, sugestões e pertinentes contribuições a este trabalho.

Ao meu coorientador professor Alexandre Behling pelas enriquecedoras orientações, pela prestatividade e pelo tempo despendido em longas conversas que foram de grande auxílio para o desenvolvimento desta pesquisa.

Aos professores, funcionários e acadêmicos do curso de Pós-Graduação em Engenharia Florestal da Universidade Federal do Paraná, pela excelência no trabalho prestado, que faz deste o melhor curso de Engenharia Florestal do país.

À Remasa Reflorestadora S.A., na pessoa de Thomas Edson Ramos, coordenador de planejamento, e Augusto Borges, analista florestal, pelo tempo disponibilizado para conversa e cessão dos dados utilizados no segundo capítulo desta dissertação.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pela bolsa de estudos que oportunizou a dedicação a este projeto de pesquisa.

*“Quando uma criatura humana desperta para um grande sonho e sobre ele lança
toda a força de sua alma, todo o universo conspira a seu favor”*

Johann Goethe

RESUMO GERAL

A redução no número de medições de alturas em parcelas de inventário florestal tem sido tema de interesse prático nas empresas florestais, uma vez que está relacionado à redução de custos. O desafio consiste na diminuição do trabalho em campo, sem que isso comprometa a acuracidade dos modelos hipsométricos que representarão a altura das árvores não mensuradas. A partir disso, o objetivo do trabalho foi modelar a relação hipsométrica de povoamentos de *Pinus taeda* L. por meio da aplicação de algoritmos de aprendizado de máquina e modelos não lineares com inclusão de covariantes, visando reduzir o número de medições de altura em campo. A apresentação dos resultados foi realizada em dois capítulos. O primeiro teve como objetivo geral a otimização de hiperparâmetros de algoritmos de aprendizado de máquina aplicados à modelagem da relação hipsométrica. Para maximização do desempenho preditivo dos algoritmos k-Vizinhos Mais Próximos (k-VMP), Floresta Aleatória (FA), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA) foi empregado o método de busca em grade com validação cruzada *k-fold*. O melhor modelo gerado para cada técnica teve o desempenho comparado com modelos não lineares modificados pela inserção de covariantes e modelos lineares genéricos. As técnicas de melhor desempenho preditivo foram RNA e MVS. Os modelos não lineares modificados também resultaram em indicadores favoráveis ao seu uso. No segundo capítulo, o objetivo foi testar a redução do número de medições de altura em inventários de *Pinus taeda* por meio de técnicas de aprendizado de máquina e modelos não lineares modificados, de modo que o uso desses não reduzisse a acuracidade obtida pelo emprego do método tradicional de modelagem da relação hipsométrica. O método tradicional (considerado testemunha) refere-se ao utilizado pela empresa que disponibilizou a base de dados e consiste na mensuração de 20 árvores por parcela em média, e no ajuste de modelos por estrato (idade, fase de manejo, sítio e região, totalizando 146 estratos). Dois subconjuntos da base de dados original foram gerados para simular a medição de alturas de três e de cinco árvores por parcela. Essas foram utilizadas para aplicação das técnicas RNA, MVS e modelo não linear com inclusão de covariantes. Para avaliação do desempenho dos métodos propostos, foram utilizados os indicadores estatísticos: erro absoluto médio, erro quadrático médio e análise gráfica de resíduos. Adicionalmente, foi realizada a análise de variância e o erro médio absoluto resultante da aplicação das técnicas (tratamentos) foram comparadas com o método tradicional (controle) pelo teste de Dunnett a 5% de probabilidade. A mensuração de 5 árvores por parcela foi suficiente para obter estimativas de altura estatisticamente equivalentes ao método tradicional, independente da técnica de ajuste dos modelos, o que não foi constatado para os tratamentos que consideraram a mensuração da altura de 3 árvores. Os estudos demonstraram que o procedimento de otimização de hiperparâmetros confere melhorias ao desempenho preditivo de técnicas de aprendizado de máquina, e que a medição da altura de 5 árvores por parcela é suficiente para modelar a relação hipsométrica de povoamentos de *Pinus taeda* por meio de MVS, RNA e modelos não lineares modificados pela inclusão de covariantes.

Palavras-chave: Redes Neurais Artificiais. Máquina de Vetores de Suporte. k-Vizinhos Mais Próximos. Inteligência artificial. Floresta Aleatória. Inclusão de covariantes.

GENERAL ABSTRACT

The reduction in the number of height measurements in forest inventory plots has been a matter of practical interest in forestry companies since it contributes to reduce costs. The challenge is to reduce the need for field operational work without compromising the accuracy of height-diameter models that will represent the height of unmeasured trees. The objective of this research was to model the height-diameter relationship in *Pinus taeda* L. stands through the application of machine learning algorithms and nonlinear models with inclusion of covariates, aiming to the reduction of the number of height measurements in forest inventories in *P. taeda* stands. The main objective of the first chapter was to optimize the hyperparameters of machine learning algorithms applied to modeling the height-diameter relationship. To maximize the predictive performance of the k-Nearest Neighbors (k-VMP), Random Forest (FA), Support Vector Machines (MVS) and Artificial Neural Networks (RNA) algorithms, the grid search method with *k*-fold cross validation was used. The best model generated for each technique had the prediction performance compared to linear and nonlinear models with covariate inclusion. The best results were obtained by the RNA and MVS models. The modified nonlinear models also presented favorable indicators to their use. In the second chapter, the objective was to reduce the number of height measurements in forest inventories using machine learning and nonlinear models, in order to maintain the accuracy obtained by the use of the traditional method. The traditional method (adopted as the control) refers to what is used by the company that provides the data set and consists of the fit of models by stratum (age, management phase, site and region). Two subsets of the original database were generated to simulate the measurement of the heights of three and five trees per plot, respectively. These were used to apply the RNA, MVS and non-linear model with inclusion of covariates. To evaluate the performance of the proposed methods, statistical indicators were used (mean absolute error and mean squared error) and scatter plots of residues. In addition, the analysis of variance and the techniques (treatments) mean absolute errors were compared with the traditional method (control) by the traditional method, by the Dunnett test at 5% probability. The measurement of five trees per plot was sufficient to obtain height estimates statistically equivalent to the traditional method, independent of the regression technique used, what was not verified for the treatments that considered the height measurement of three trees per plot. The studies showed that optimization procedures of hyperparameters improve the predictive performance of machine learning techniques, and that the height measurement of five trees per plot is sufficient to model the height-diameter relationship in *Pinus taeda* stands by MVS, RNA and nonlinear models modified by the inclusion of covariates.

Keywords: Artificial Neural Networks. Support Vector Machines. k-Nearest Neighbors.

Artificial intelligence. Random Forest. Covariate inclusion.

LISTA DE FIGURAS

FIGURA 1 - QUANTIDADE DE PESQUISAS QUE ENVOLVEM A TEMÁTICA RELAÇÃO HIPSOMÉTRICA PUBLICADAS EM REVISTAS CIENTÍFICAS DE 2010 A 2017, CLASSIFICADAS POR TÉCNICA DE MODELAGEM.....	18
FIGURA 2 - HIERARQUIA DO APRENDIZADO.	20
FIGURA 3 - ILUSTRAÇÃO DE UM PROBLEMA DE CLASSIFICAÇÃO BINÁRIA LINEARMENTE SEPARÁVEL (A); SOLUÇÕES POSSÍVEIS PARA O PROBLEMA DA SEPARAÇÃO DOS CONJUNTOS ATRAVÉS DE UM HIPERPLANO (B); SEPARAÇÃO DOS DADOS PELO HIPERPLANO DE SEPARAÇÃO ÓTIMA (C).....	25
FIGURA 4 - ERROS DE CLASSIFICAÇÃO EM UM CONJUNTO DE OBSERVAÇÕES PERTENCENTES A DUAS CLASSES.....	26
FIGURA 5 - EXEMPLO DE UM PROBLEMA DE CLASSIFICAÇÃO BINÁRIA NÃO SEPARÁVEL LINEARMENTE NO ESPAÇO ORIGINAL (A). MAPEAMENTO DAS VARIÁVEIS ORIGINAIS NO ESPAÇO DE CARACTERÍSTICAS DE ALTA DIMENSÃO POR MEIO DE UMA FUNÇÃO KERNEL (B).....	27
FIGURA 6 - CONJUNTO DE DADOS PARA REGRESSÃO (A); CRIAÇÃO DOS CONJUNTOS NEGATIVO E POSITIVO A PARTIR DA SUBTRAÇÃO E DA ADIÇÃO DA CONSTANTE D (B); HIPERPLANO DE MÁXIMA SEPARAÇÃO (C).....	29
FIGURA 7 - LIMITE DE TOLERÂNCIA DETERMINADO POR ξ EM MÁQUINAS DE VETORES DE SUPORTE.....	30
FIGURA 8 - EXEMPLO DE ÁRVORE DE REGRESSÃO PARA PREDIÇÃO DE ALTURA TOTAL, EM QUE AS SETAS REPRESENTAM OS RAMOS, OS RETÂNGULOS BRANCOS REPRESENTAM OS NÓS DE DECISÃO E OS CINZAS AS FOLHAS TERMINAIS.	31
FIGURA 9 - ILUSTRAÇÃO DO MÉTODO DE CRIAÇÃO DE ÁRVORES DE REGRESSÃO OU DECISÃO NO ALGORITMO FLORESTA ALEATÓRIA.....	32
FIGURA 10 - ARQUITETURA HIPOTÉTICA DE UMA REDE NEURAL ARTIFICIAL DE MÚLTIPLAS CAMADAS.....	34

FIGURA 2.1 - ESTATÍSTICA DESCRITIVA DA BASE DE DADOS OBTIDOS EM INVENTÁRIOS EM POVOAMENTOS DE <i>PINUS TAEDA</i> NO ESTADO DO PARANÁ.	80
FIGURA 2.2 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DAS ALTURAS ESTIMADAS PARA POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ.	95
FIGURA 2.3 - FREQUÊNCIA DE RESÍDUOS DAS PREDIÇÕES REALIZADAS PELOS MÉTODOS ESTUDADOS NO CONJUNTO DE TESTE DE POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ.	95
FIGURA 2.4 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL EM FUNÇÃO DOS DIÂMETROS À ALTURA DO PEITO (DAP) REALIZADAS PELOS MÉTODOS ESTUDADOS, PARA POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ.	96
FIGURA 2.5 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL EM FUNÇÃO DA IDADE, REALIZADA PELOS MÉTODOS ESTUDADOS, PARA POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ.	96
FIGURA 2.6 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL PARA POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DA VARIÁVEL H_{dn} , EM QUE $n = 6$ NO MÉTODO TRADICIONAL, $n = 3$ NO CONJUNTO A3 E $n = 5$ NO CONJUNTO A5.	97
FIGURA 2.7 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL PARA POVOAMENTOS DE <i>PINUS TAEDA</i> L. NO PARANÁ REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DO DIÂMETRO MÉDIO QUADRÁTICO.	97
FIGURA 2.8 - ALTURAS OBSERVADAS E ESTIMADAS POR DIFERENTES MÉTODOS PARA <i>PINUS TAEDA</i> L., PARA 5 PARCELAS SELECIONADAS ALEATORIAMENTE NO CONJUNTO DE TESTE.	98

FIGURA 2.9 - NUVENS DE OBSERVAÇÕES DOS CONJUNTOS DE AJUSTE/TREINAMENTO UTILIZADOS NO PRESENTE ESTUDO.	101
FIGURA 2.10 - SEQUÊNCIA DE PROCEDIMENTOS PARA MEDIÇÃO DE ALTURAS DE CINCO ÁRVORES POR UNIDADE AMOSTRAL PELO MÉTODO PROPOSTO NESTE ESTUDO.....	102

LISTA DE QUADROS

QUADRO 1 - PARADIGMAS DO APRENDIZADO	21
QUADRO 2 - FUNÇÕES KERNEL MAIS UTILIZADAS EM MÁQUINAS DE VETORES DE SUPORTE.....	28
QUADRO 3 - FUNÇÕES DE ATIVAÇÃO COMUMENTE UTILIZADAS NO TREINAMENTO DE REDES NEURAIAS ARTIFICIAIS.....	33
QUADRO 1.1 - MODELOS LINEARES SELECIONADOS PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA EM POVOAMENTOS DE <i>Pinus taeda</i>	51
QUADRO 1.2 - MODELOS NÃO LINEARES SELECIONADOS PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE <i>Pinus taeda</i>	51
QUADRO 1.3 - MÉTRICAS DE DISTÂNCIA TESTADAS PARA O ALGORITMO K- VMP.....	52
QUADRO 1.4 - FUNÇÕES <i>KERNEL</i> TESTADAS PARA K-VIZINHOS MAIS PRÓXIMOS.....	52
QUADRO 1.5 - FUNÇÕES DE ATIVAÇÃO TESTADAS NO TREINAMENTO DE REDES NEURAIAS ARTIFICIAIS.	54

LISTA DE TABELAS

TABELA 1.1 - COEFICIENTES DE REGRESSÃO E INDICADORES ESTATÍSTICOS DOS MODELOS HIPSOMÉTRICOS AJUSTADOS.....	57
TABELA 1.2 - INDICADORES DE DESEMPENHO DOS MODELOS ESTATÍSTICOS E DE APRENDIZADO DE MÁQUINA NA ESTIMATIVA DA ALTURA TOTAL DE ÁRVORES DE <i>PINUS TAEDA</i> COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.....	64
TABELA 2.1 - SORTIMENTOS UTILIZADOS PARA ESTIMATIVA DOS VOLUMES COMERCIAIS.	86
TABELA 2.2 - ESTIMATIVAS DOS PARÂMETROS PARA OS MODELOS NÃO LINEARES AJUSTADOS PARA A RELAÇÃO HIPSOMÉTRICA DE ÁRVORES DE <i>PINUS TAEDA</i> NO PARANÁ.	88
TABELA 2.3 - ESTATÍSTICAS DE TESTE DOS MÉTODOS EMPREGADOS NA ESTIMATIVA DA ALTURA TOTAL E DOS VOLUMES TOTAL E POR SORTIMENTO PARA POVOAMENTOS DE <i>PINUS TAEDA</i> EM DIVERSAS IDADES NO ESTADO DO PARANÁ.	93
TABELA 2.4 - DIFERENÇAS PERCENTUAIS OBSERVADAS ENTRE OS INDICADORES ESTATÍSTICOS DOS MÉTODOS PROPOSTOS E DO MÉTODO TRADICIONAL PARA POVOAMENTOS DE <i>PINUS TAEDA</i> EM DIVERSAS IDADES NO ESTADO DO PARANÁ.	93
TABELA 2.5 - ANÁLISE DE VARIÂNCIA PARA A DIFERENÇA ABSOLUTA ENTRE AS ALTURAS OBSERVADAS E ESTIMADAS.....	99
TABELA 2.6 - TESTE DE DUNNETT PARA COMPARAÇÃO DAS MÉDIAS DOS TRATAMENTOS COM O CONTROLE.	99

SUMÁRIO

1 INTRODUÇÃO GERAL	16
2 OBJETIVOS	19
2.1 OBJETIVO GERAL	19
2.2 OBJETIVOS ESPECÍFICOS	19
3 REFERENCIAL TEÓRICO	19
3.1 APRENDIZADO DE MÁQUINA.....	19
3.2 MÉTODOS DE APRENDIZAGEM SUPERVISIONADA	21
3.2.1 k-Vizinhos mais próximos.....	21
3.2.2 Máquinas de Vetores de Suporte	24
3.2.3 Floresta Aleatória	30
3.2.4 Redes Neurais Artificiais	32
3.3 MODELOS NÃO LINEARES	35
3.4 RELAÇÃO HIPSOMÉTRICA.....	37
4 REFERÊNCIAS	39
CAPÍTULO 1 – TÉCNICAS DE APRENDIZADO DE MÁQUINA E MODELOS NÃO LINEARES COM INCLUSÃO DE COVARIANTES NA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE <i>Pinus taeda</i> L.....	46
1.1 INTRODUÇÃO	48
1.2 MATERIAIS E MÉTODOS	49
1.2.1 Base de dados	49
1.2.2 Modelos de regressão.....	50
1.2.2.1 Modelos lineares	50
1.2.2.2 Modelos não lineares com adição de covariantes	51
1.2.3 Algoritmos de aprendizado de máquina	51
1.2.3.1 k-Vizinhos Mais Próximos (k-VMP).....	52
1.2.3.2 Floresta Aleatória (FA)	53
1.2.3.3 Máquinas de Vetores de Suporte (MVS)	53
1.2.3.4 Redes Neurais Artificiais (RNA).....	53
1.2.4 Otimização dos hiperparâmetros dos algoritmos de aprendizado de máquina	54
1.2.5 Avaliação do desempenho dos modelos	55
1.3 RESULTADOS E DISCUSSÃO.....	56
1.3.1 Ajuste dos modelos de regressão	56

1.3.2 Busca em grade	58
1.3.2.1 K-vizinhos Mais Próximos.....	58
1.3.2.2 Floresta Aleatória	60
1.3.2.3 Máquinas de Vetores de Suporte	61
1.3.2.4 Redes Neurais Artificiais.....	62
1.3.3 Avaliação dos modelos	63
1.3.4 Considerações sobre a otimização de hiperparâmetros de algoritmos de aprendizado de máquina.....	70
1.4 CONCLUSÕES	72
1.5 REFERÊNCIAS.....	73
CAPÍTULO 2 – REDUÇÃO NA QUANTIDADE DE MEDIÇÕES DE ALTURA EM INVENTÁRIOS DE <i>Pinus taeda</i> L. PELO EMPREGO DE MODELOS NÃO LINEARES E TÉCNICAS DE APRENDIZADO DE MÁQUINA.....	76
2.1 INTRODUÇÃO	78
2.2 MATERIAIS E MÉTODOS	79
2.2.1 Base de dados	79
2.2.2 Redução do número de medições de alturas	80
2.2.3 Método de ajuste tradicional.....	81
2.2.4 Treinamento dos algoritmos de aprendizado de máquina	82
2.2.4.1 Máquinas de Vetores de Suporte	82
2.2.4.2 Redes Neurais Artificiais.....	83
2.2.4.3 Ajuste do modelo não linear	84
2.2.5 Avaliação do desempenho dos métodos	84
2.2.5.1 Desempenho na estimativa da altura total	85
2.2.5.2 Impacto nos volumes total e por sortimento.....	85
2.2.6 Análise de variância e teste de comparação de médias	87
2.3 RESULTADOS E DISCUSSÃO.....	87
2.3.1 Ajustes de modelos de regressão estatística	87
2.3.2 Redes neurais artificiais	88
2.3.2.1 Conjunto de treinamento A3	88
2.3.2.2 Conjunto de treinamento A5	89
2.3.3 Máquinas de vetores de suporte	91
2.3.3.1 Conjunto de treinamento A3	91
2.3.3.2 Conjunto de treinamento A5	91

2.3.4 Análise de indicadores estatísticos.....	92
2.3.5 Análise de variância e teste de comparação de médias	99
2.3.6 Considerações sobre a redução do número de medições de altura em parcelas de inventário florestal	100
2.4 CONCLUSÕES	103
2.5 REFERÊNCIAS.....	104
CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES	105
REFERÊNCIAS GERAIS.....	107
APÊNDICE 1 – COEFICIENTES DOS MODELOS SELECIONADOS NO AJUSTE PELO MÉTODO TRADICIONAL (POR ESTRATO).	116
APÊNDICE 2 – CÓDIGOS DE PROGRAMAÇÃO EM R PARA TREINAMENTO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA E AJUSTE DE MODELOS NÃO LINARES	119

1 INTRODUÇÃO GERAL

A utilização da relação hipsométrica em povoamentos florestais surgiu a partir da necessidade de se reduzir os custos expressivos com a mensuração da altura de todas as árvores da unidade amostral. A altura é medida indiretamente por meio de instrumentos como o hipsômetro, sendo sujeita a erros especialmente na presença de fatores climáticos desfavoráveis como o vento, o que a torna mais onerosa.

Em inventários florestais, especialmente em florestas plantadas, são medidos todos os diâmetros e parte das alturas. Esses dados são utilizados para ajustar equações hipsométricas por meio de análise de regressão, conseqüentemente, obter as alturas das demais árvores nas parcelas.

Schneider (1986) afirma que para o ajuste da relação hipsométrica são necessárias entre 30 e 40 alturas distribuídas em toda a amplitude diamétrica, no entanto, segundo SILVA *et al.* (2007) é usual nas empresas florestais a medição de 15 ou 20 árvores por unidade amostral.

Barros *et al.* (2002) e Bartoszeck *et al.* (2004) recomendam o ajuste de uma equação hipsométrica para cada parcela, ou para um grupo homogêneo de parcelas de mesma idade, sítio e densidade, de modo reduzir a alta variabilidade de dados advindos de povoamentos com diferentes características.

Diversos fatores afetam a relação hipsométrica, tais como: idade do povoamento, sítio, densidade, posição sociológica e práticas silviculturais, o que direcionou o desenvolvimento de modelos lineares genéricos, que incluem uma ou mais das variáveis que afetam a relação hipsométrica, visando aumentar a capacidade de generalização das equações (BARROS *et al.*, 2002; SCOLFARO, 2005; CAMPOS e LEITE, 2006).

A utilização de modelos lineares genéricos em substituição aos modelos lineares tradicionais (equações por parcela) permitiu a redução do trabalho de ajuste e seleção de equações por parcela ou povoamento e conseqüentemente dos custos com processamento de dados (BARROS *et al.* 2002). No entanto, é escassa a abordagem referente à redução do trabalho de medição das alturas em campo.

O que foi exposto até aqui serviu de base para a formulação de duas perguntas: O número de árvores mensuradas em cada parcela pode ser reduzido sem aumento do trabalho de processamento? Outras técnicas de modelagem, mais

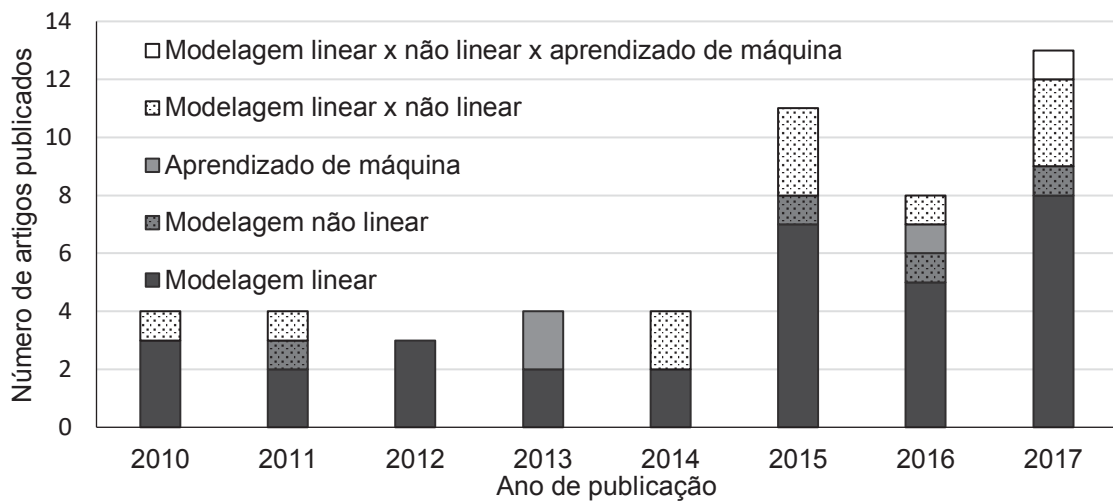
sofisticadas, podem contribuir para essa redução? Essas são as duas questões principais abordadas nesta pesquisa.

O avanço da ciência computacional permitiu o desenvolvimento e a popularização de técnicas mais elaboradas de modelagem, como os modelos não lineares. Devido à sua maior flexibilidade e fundamentação biológica, esses permitem a geração de previsões mais realistas, uma vez que a relação entre o DAP e a altura das árvores muitas vezes não é linear (CALEGARIO *et al.*, 2005). Ainda, modelos não lineares podem ter caráter genérico por meio da decomposição de seus parâmetros e inserção de variáveis como altura dominante e idade como covariantes (ALVES *et al.*, 2015; SENA *et al.*, 2015).

Recentemente, pesquisas com algoritmos de aprendizado de máquina trouxeram novas discussões para o campo mensuração florestal. Devido à sua elevada capacidade de generalização, também de permitir a inclusão de variáveis contínuas e categóricas de maneira facilitada, algoritmos como Redes Neurais Artificiais (RNA) e Máquinas de Vetores de Suporte (MVS) vêm ganhando destaque e têm sido objeto de análise em estudos de modelagem da relação hipsométrica (BINOTI, 2012; BINOTI *et al.*, 2013a; BINOTI *et al.*, 2013b; CAMPOS *et al.*, 2016; MONTAÑO, 2016; VENDRUSCOLO *et al.*, 2017; MENDONÇA *et al.*, 2018).

Houve um crescimento na quantidade de pesquisas envolvendo a modelagem da relação hipsométrica nos últimos anos (Figura 1). A maioria dos trabalhos envolve comparações entre os métodos tradicionais (normalmente caracterizados por análises de regressão) e algoritmos de aprendizado de máquina. Pode-se considerar que há uma popularização da utilização de algoritmos de aprendizado de máquina em pesquisas voltadas à mensuração florestal, mas a utilização de modelos lineares tradicionais ainda é predominante, e, provavelmente, também nas empresas florestais.

FIGURA 1 - QUANTIDADE DE PESQUISAS QUE ENVOLVEM A TEMÁTICA RELAÇÃO HIPSONÔMICA PUBLICADAS EM REVISTAS CIENTÍFICAS DE 2010 A 2017, CLASSIFICADAS POR TÉCNICA DE MODELAGEM.



FONTE: O autor (2019).

NOTA: Consulta realizada nos bancos de dados: ResearchGate, SciELO e RedALyC, utilizando a palavra-chave: "relação hipsométrica".

Para que os benefícios da utilização de técnicas de aprendizado de máquina e modelos estatísticos mais sofisticados, como os não lineares com inserção de covariantes, possam ser maximizados, é fundamental a busca pela otimalidade na geração de modelos de regressão. Para o primeiro método envolve a otimização de hiperparâmetros e para o segundo a decomposição de parâmetros e apropriada seleção de covariantes.

Para responder as questões produzidas nesta pesquisa, optou-se por apresentar 2 capítulos. O primeiro foi direcionado à melhoria da performance preditiva de algoritmos de aprendizado de máquina por meio da otimização de seus hiperparâmetros e avaliação do desempenho dessas técnicas em comparação a modelos de regressão não lineares modificados pela adição de covariantes, e lineares. No segundo, buscou-se desenvolver um método para redução do número de medições de altura em inventários florestais, por meio do uso de técnicas de aprendizado de máquina e modelos não lineares.

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral desta pesquisa foi testar se técnicas de aprendizado de máquina e modelos não lineares aplicados à modelagem da relação hipsométrica de *Pinus taeda* permite a redução do número de árvores a serem mensuradas em inventários florestais.

2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos do estudo foram:

- Gerar diretrizes para configuração e otimização de parâmetros de algoritmos de aprendizado de máquina aplicados à mensuração florestal;
- Comparar as técnicas de aprendizado de máquina, regressão linear e regressão não linear com inclusão de covariantes na modelagem da relação hipsométrica;
- Definir um número mínimo de medições de alturas em inventários florestais sem perda de confiabilidade nas predições.

3 REFERENCIAL TEÓRICO

3.1 APRENDIZADO DE MÁQUINA

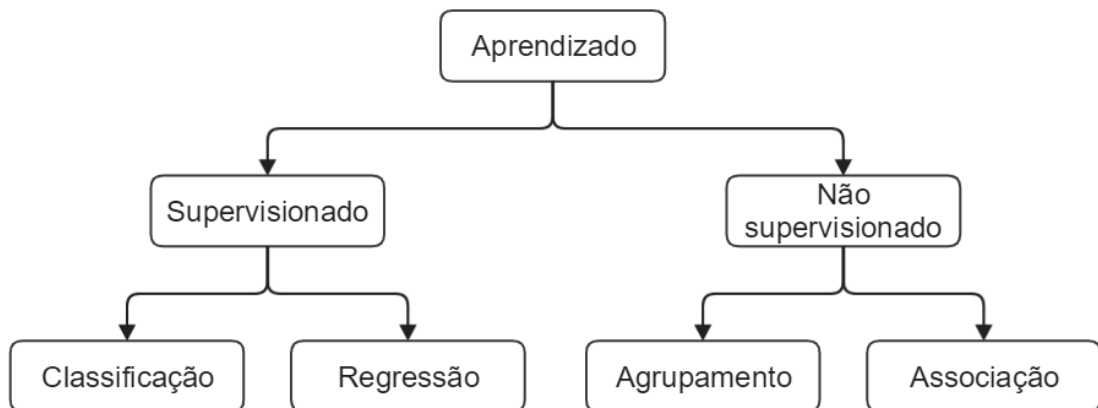
O aprendizado de máquina é uma área da inteligência artificial voltada à construção de sistemas capazes de adquirir conhecimento de forma automática. Um sistema de aprendizado é um programa de computador capaz de tomar decisões baseado em experiências acumuladas por meio da solução bem-sucedida de problemas anteriores (MONARD e BARANAUSKAS, 2003).

O aprendizado de máquina baseia-se no conceito de aprendizado indutivo, que é caracterizado como o raciocínio que se origina em um conceito específico e o generaliza, ou seja, da parte para o todo, do específico ao geral. Na indução, um conceito é aprendido por meio de uma inferência indutiva sobre exemplos apresentados. A indução é um recurso comumente utilizado pelo cérebro humano

para derivar novos conhecimentos, no entanto deve ser utilizada com cautela pois se os exemplos forem insuficientes em quantidade ou qualidade, as hipóteses obtidas podem ser de pouco valor (REZENDE, 2003).

O aprendizado indutivo pode ser dividido em supervisionado e não-supervisionado, conforme ilustrado na Figura 2.

FIGURA 2 - HIERARQUIA DO APRENDIZADO.



FONTE: O autor (2019).

No aprendizado supervisionado cada exemplo é apresentado ao algoritmo de aprendizado, ou indutor, com seu respectivo atributo de saída, ou variável dependente. Quando a variável dependente é discreta, o problema de aprendizado é caracterizado como de classificação, e quando a variável dependente é contínua, como de regressão. O objetivo dos sistemas de aprendizado supervisionado é construir um sistema capaz de determinar corretamente classes ou valores de novos exemplos cujas classes ou valores são desconhecidos (LARRANAGA *et al.*, 2006; MONTAÑO, 2016).

No aprendizado não-supervisionado os atributos de saída não são conhecidos e o objetivo do algoritmo de aprendizado é analisar os exemplos fornecidos e agrupá-los de alguma maneira formando agrupamentos ou *clusters*, ou associá-los a partir de regras de associação identificadas durante o treinamento (MONARD e BARANAUSKAS, 2003; MONTAÑO, 2016).

Segundo Rezende (2003), os algoritmos de aprendizado podem ser categorizados em cinco paradigmas, apresentados no Quadro 1.

QUADRO 1 - PARADIGMAS DO APRENDIZADO

Paradigma	Descrição
Simbólico	Os sistemas de aprendizado simbólico utilizam representações simbólicas por meio da análise de exemplos e contra-exemplos. Estas representações geralmente estão na forma de alguma expressão lógica, árvore de decisão, regras ou rede semântica.
Baseado em exemplos	Os sistemas baseados em exemplos armazenam o conjunto de dados de treinamento e utilizam medidas de similaridade para identificar os casos mais similares ao exemplo a ser analisado. Este tipo de sistema de aprendizado, denominado <i>lazy</i> , necessita manter os exemplos na memória para prever novos casos (exemplos). O algoritmo k-vizinho mais próximo (k-VMP) é um exemplo deste paradigma.
Estatístico	Os sistemas de aprendizado estatísticos utilizam modelos estatísticos para aproximação do conceito induzido. Dentre os métodos estatísticos destacam-se os de aprendizado probabilístico Bayesiano e as Máquinas de Vetores de Suporte (MVS).
Conexionista	Os sistemas conexionistas envolvem unidades de processamento altamente interconectadas. São exemplos as Redes Neurais Artificiais (RNA), construções matemáticas simplificadas inspiradas no modelo biológico do sistema nervoso.
Evolutivo	Os algoritmos de aprendizado evolutivo consistem na competição de elementos de classificação que competem para realizar a predição. Nestes sistemas os elementos que possuem baixa performance são descartados, enquanto os de alta performance se perpetuam. Este paradigma possui uma analogia direta com a teoria de Darwin, na qual os indivíduos mais adaptados sobrevivem. São exemplos os algoritmos genéticos.

FONTE: O autor (2019).

3.2 MÉTODOS DE APRENDIZAGEM SUPERVISIONADA

Nesta seção serão apresentadas as características dos algoritmos de aprendizado que serão estudados nesta dissertação.

3.2.1 K-VIZINHOS MAIS PRÓXIMOS

O método do Vizinho mais próximo (*Nearest neighbor*) (FIX e HODGES, 1951) é uma das mais simples e intuitivas técnicas da área de aprendizado de máquina. Trata-se de um método não paramétrico onde uma nova observação recebe o valor ou a classe do caso mais próximo contido no conjunto de exemplos de treinamento.

Formalmente, o método pode ser descrito como $L = \{(y_i, x_i)\}$ sendo o conjunto de dados de treinamento, onde y_i denota a variável dependente, e x_i as variáveis preditoras. A determinação dos vizinhos mais próximos é baseada em uma função de distância $d(x_i, x_j)$. O valor de uma nova observação (y, x) será o mesmo

de seu vizinho mais próximo $(y_{(1)}, x_{(1)})$, em que a proximidade é medida pela função $d(x, x_1) = \min_i(d(x, x_i))$, ou seja, o vizinho mais próximo é aquele que possui a menor distância calculada da observação, dentre todas as observações empregadas no treinamento do algoritmo (HECHENBICHLER e SCHLIEP, 2004).

A distância é calculada a partir de métricas definidas, como a distância euclidiana ou a distância de Manhattan (BHATIA, 2010). A métrica de distância é um hiperparâmetro do algoritmo e deve ser definido para realização do treinamento. A título de exemplo, a distância euclidiana entre duas observações pode ser calculada por meio da seguinte expressão:

$$d_{euclidiana}(x_i, x_j) = \sqrt{\sum_{s=1}^p d(x_{is} - x_{js})^2}$$

Uma extensão difundida deste método é conhecida como k-Vizinho mais próximo (*k-Nearest neighbor* ou kNN). Nesse algoritmo não somente o exemplo mais similar é utilizado para determinação do rótulo de um novo caso, mas sim os k exemplos mais similares (HECHENBICHLER e SCHLIEP, 2004). No kNN o parâmetro k é um hiperparâmetro e deve ser determinado pelo próprio usuário.

Quando o hiperparâmetro $k = 1$, o algoritmo corresponde ao método do vizinho mais próximo, e quando $k > 1$, o rótulo do novo exemplo é determinado a partir da média entre os k vizinhos mais próximos para os problemas de regressão. Quando o problema é de classificação, a classe de maior frequência dentre os k vizinhos mais próximos é atribuída ao novo caso (HECHENBICHLER e SCHLIEP, 2004).

Um refinamento do método do k-Vizinhos mais próximos, conhecido como k-Vizinhos mais próximos ponderado (*Weighted k-Nearest Neighbors* ou wkNN), descrito por Hechenbichler e Schliep (2004). Em wkNN parte-se do pressuposto de que os casos mais próximos possuem maior peso na determinação do rótulo de um novo caso, em comparação aos vizinhos que se encontram mais distantes. Desse modo, o rótulo de um novo caso é obtido pelo cálculo da média ponderada pela distância do caso até os vizinhos mais próximos.

A ponderação dos k-vizinhos mais próximos é realizada por meio de uma função *kernel* $K(D)$ que possui valor máximo quando a distância é zero, e valores

decrecentes a medida em que a distância se torna maior. Existem diversas funções *kernel* que podem ser aplicadas ao kNN ponderado. Algumas comumente utilizadas são apresentadas no Quadro 2.

QUADRO 2 - FUNÇÕES *KERNEL* COMUMENTE APLICADAS EM K-VIZINHOS MAIS PRÓXIMOS.

Função <i>kernel</i>	Expressão matemática
Retangular	$K = \frac{1}{2}$
Triangular	$K(d) = (1 - d)$
Quadrática	$K(d) = \frac{15}{16}(1 - d^2)^2$
Cúbica	$K(d) = \frac{35}{32}(1 - d^2)^3$
Inversa	$K(d) = \frac{1}{ d }$

FONTE: O autor (2019).

No wkNN, $D(x, x_1)$ é a distância $d(x, x_1)$ padronizada em função da distância $d(x, x_{k+1})$, conforme a expressão a seguir:

$$D(x, x_i) = \frac{d(x, x_i)}{d(x, x_{(k+1)})}$$

a qual garante que as funções *kernel* resultem em zero quando $D = 1$. Ainda, para que todo $D > 1$ seja descartado do cálculo, garantindo que somente os k vizinhos componham a média, define-se o limite da janela da função *kernel* como $D = 1$. Desse modo, a predição (\hat{y}) de uma nova observação é a média ponderada das observações que se encontram dentro da janela local definida. A função de predição da nova observação é descrita matematicamente pela seguinte expressão.

$$\hat{y} = \frac{\sum_{i=1}^k K\left(\frac{d(x, x_i)}{d(x, x_{(k+1)})}\right) y_i}{\sum_{i=1}^k K\left(\frac{d(x, x_i)}{d(x, x_{(k+1)})}\right)} = \frac{\sum_{i=1}^k w_{(i)} y_i}{\sum_{i=1}^k w_{(i)}}$$

Particularmente, no que concerne às aplicações na área de manejo de recursos florestais, nota-se a popularização do método dos K-Vizinhos mais próximos para a produção de estimativas espacialmente contíguas de atributos de um ecossistema florestal (CHIRICI *et al.*, 2016). A título de exemplo, por meio da associação entre dados obtidos por sensoriamento remoto e dados provenientes de

inventário florestal, podem-se mapear atributos florestais, tais como estoques volumétricos, idade, altura e densidade de populações, gerando mapas contínuos que fornecem suporte a políticas públicas e planos de manejo (REESE *et al.*, 2005; BEAUDOIN *et al.*, 2014; WOJCIECHOWSKI, 2015; SILVA *et al.*, 2018).

3.2.2 MÁQUINAS DE VETORES DE SUPORTE

As Máquinas de Vetores de Suporte (MVS ou do inglês *support vector machines* - SVM) são algoritmos de aprendizado de máquina baseados na teoria de aprendizado estatístico, desenvolvida por Vapnik (1995). As MVS foram originalmente desenvolvidas como algoritmos de classificação binária. No entanto, posteriormente foram estendidas para as tarefas de regressão por meio de duas abordagens: Regressão de Vetores de Suporte (RVS) ou *support vector regression* (SVR) e Máquinas de Vetores de Suporte de mínimos quadrados (MVS-MQ) ou *least-squares support vector machines* (LS-SVM) (ÁLVAREZ MEZA, 2012).

Para compreensão dos princípios fundamentais das MVS, podemos assumir um conjunto de dados de entrada (x_i, y_i) , em que $x_i \in R$ é um vetor que pertence a uma de duas classes $y_i \in \{-1; +1\}$ (Figura 3A). O objetivo da MVS é separar as classes por meio de um hiperplano, expresso por $f(x) = w \cdot x + b$, em que w é o conjunto de parâmetros e b o desvio aleatório.

Considerando que existem inúmeras soluções possíveis (mínimos locais) que separem o conjunto de dados em dois (Figura 3B), o algoritmo MVS emprega o princípio da minimização do risco estrutural, que visa a obtenção do hiperplano de máxima distância das amostras de classificação, denominado hiperplano de separação ótima.

Para a busca da solução ótima, a função $f(x) = w \cdot x + b$ é descrita pelo seguinte problema de otimização:

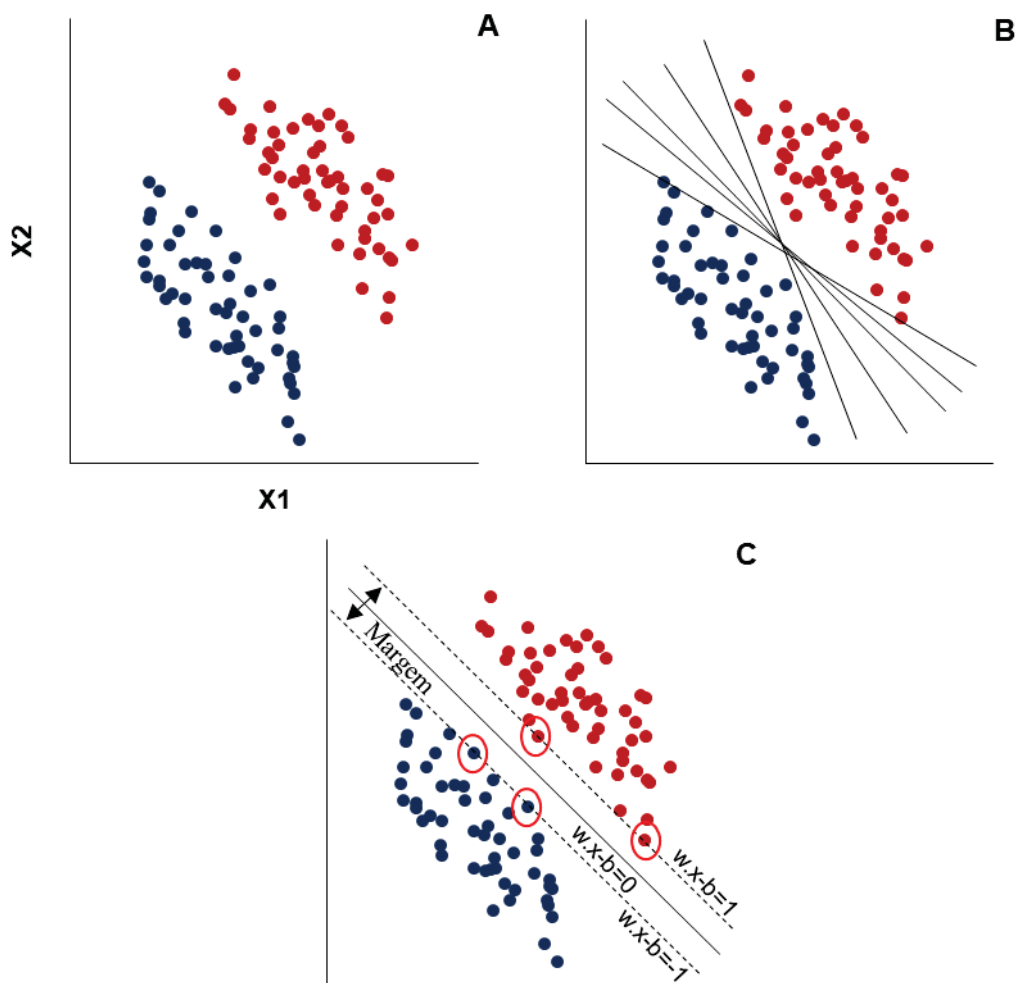
$$\begin{aligned} & \text{minimize } \frac{1}{2} \|w\|^2 \\ & \text{sujeito a } \begin{cases} w \cdot x + b \geq +1 & \forall y \in \{+1\} \\ w \cdot x + b \leq -1 & \forall y \in \{-1\} \end{cases} \end{aligned}$$

em que $\|w\|^2$ é o produto escalar do vetor w com ele mesmo, ou seja $(w \cdot w)$. As restrições do problema garantem que os dados das observações utilizadas no

conjunto de treinamento não apareçam entre as margens de separação, o que lhes confere o nome de Margens Rígidas.

São denominados vetores de suporte as amostras que delimitam as margens de separação entre as classes, as quais assumem valor $w \cdot x + b = +1$ quando pertencentes à classe +1, e $w \cdot x + b = -1$ quando pertencentes à classe -1 (Figura 3C).

FIGURA 3 - ILUSTRAÇÃO DE UM PROBLEMA DE CLASSIFICAÇÃO BINÁRIA LINEARMENTE SEPARÁVEL (A); SOLUÇÕES POSSÍVEIS PARA O PROBLEMA DA SEPARAÇÃO DOS CONJUNTOS ATRAVÉS DE UM HIPERPLANO (B); SEPARAÇÃO DOS DADOS PELO HIPERPLANO DE SEPARAÇÃO ÓTIMA (C).



FONTE: O autor (2019).

O exemplo dado consiste em um problema de classificação de amostras linearmente separáveis, porém na maioria das vezes duas classes não são separáveis por um hiperplano, o que levou Cortes e Vapnik (1995) a desenvolverem o conceito de variável de folga ξ_i , o qual permite a existência de erros de classificação.

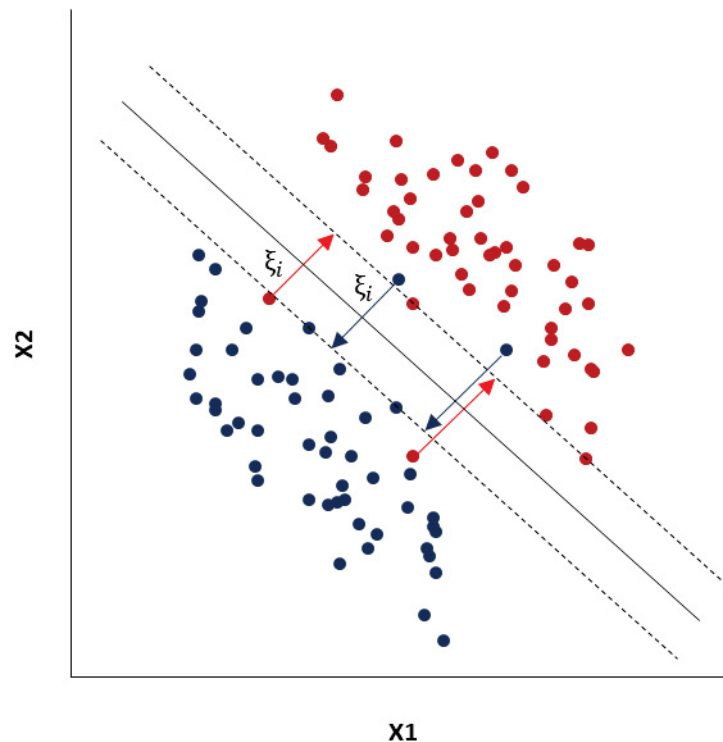
O erro é proporcional à distância da amostra à margem definida pelos vetores de suporte da classe à qual pertencem. Com a flexibilização da largura entre margens, os erros de classificação são controlados por uma constante C , que quando possui um valor baixo, admite que mais amostras possam ser classificadas erroneamente, enquanto um valor alto torna o problema menos tolerante a erros (Figura 4). O valor do parâmetro C está ligado à complexidade de um modelo de MVS, agindo na regulação de subajuste e sobreajuste (HUANG e WANG, 2006).

A constante C é adicionada ao problema de otimização juntamente à variável ξ_i , de modo a controlar a tolerância à erros, modificando a função objetivo que toma as seguintes expressões:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

$$\text{sujeito a } \begin{cases} w \cdot x + b \geq +1 - \xi_i \\ w \cdot x + b \leq -1 + \xi_i \\ \xi_i > 0 \end{cases}$$

FIGURA 4 - ERROS DE CLASSIFICAÇÃO EM UM CONJUNTO DE OBSERVAÇÕES PERTENCENTES A DUAS CLASSES.



FONTE: O autor (2019).

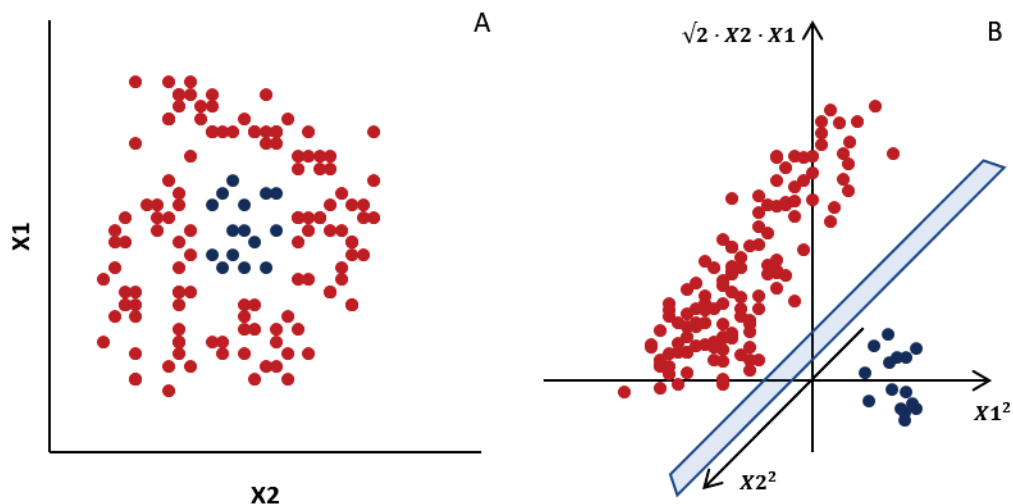
Em diversos casos a classificação de dados linearmente inseparáveis não pode ser realizada mesmo utilizando variáveis de folga. Nesses casos utiliza-se um artifício de transformação dimensional ao conjunto de entrada, o qual é mapeado para um espaço de maior dimensão, chamado de espaço de características, por meio de uma função *kernel* não linear (φ). (HEARST *et al.*, 1998).

A título de exemplo, as amostras ilustradas na Figura 5A pertencem a duas classes que não podem ser separadas linearmente em seu espaço original. Cada amostra possui valores para um par de variáveis (x_1, x_2), o que faz deste um problema bidimensional. Uma função *kernel* pode ser aplicada de modo a aumentar a dimensão original do problema, como exemplifica a expressão a seguir:

$$\varphi(x) = (x_1^2, x_2^2, \sqrt{2} \cdot x_1 \cdot x_2)$$

A função *kernel* recebe o par de variáveis no espaço original e calcula o produto escalar destes no espaço de características. O aumento da dimensão dos dados permite que estes sejam separáveis por um hiperplano em um espaço de características de alta dimensão (Figura 5B).

FIGURA 5 - EXEMPLO DE UM PROBLEMA DE CLASSIFICAÇÃO BINÁRIA NÃO SEPARÁVEL LINEARMENTE NO ESPAÇO ORIGINAL (A). MAPEAMENTO DAS VARIÁVEIS ORIGINAIS NO ESPAÇO DE CARACTERÍSTICAS DE ALTA DIMENSÃO POR MEIO DE UMA FUNÇÃO KERNEL (B).



FONTE: O autor (2019).

Existem diversas funções *kernel*, como a linear, a sigmoideal, a polinomial e a função de base radial (FBR), sendo esta última a mais popular. As funções *kernel* mais utilizadas são apresentadas no Quadro 2.

QUADRO 3 - FUNÇÕES KERNEL MAIS UTILIZADAS EM MÁQUINAS DE VETORES DE SUPORTE.

Função <i>kernel</i>	Expressão matemática
Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polinomial	$K(x_i, x_j) = (\gamma \cdot x_i \cdot x_j + \beta)^d$
Sigmoidal	$K(x_i, x_j) = \tanh(\gamma \cdot x_i \cdot x_j + \beta)$
Função de base radial	$K(x_i, x_j) = \exp[-\gamma \cdot \ x_i - x_j\ ^2]$

FONTE: O autor (2019).

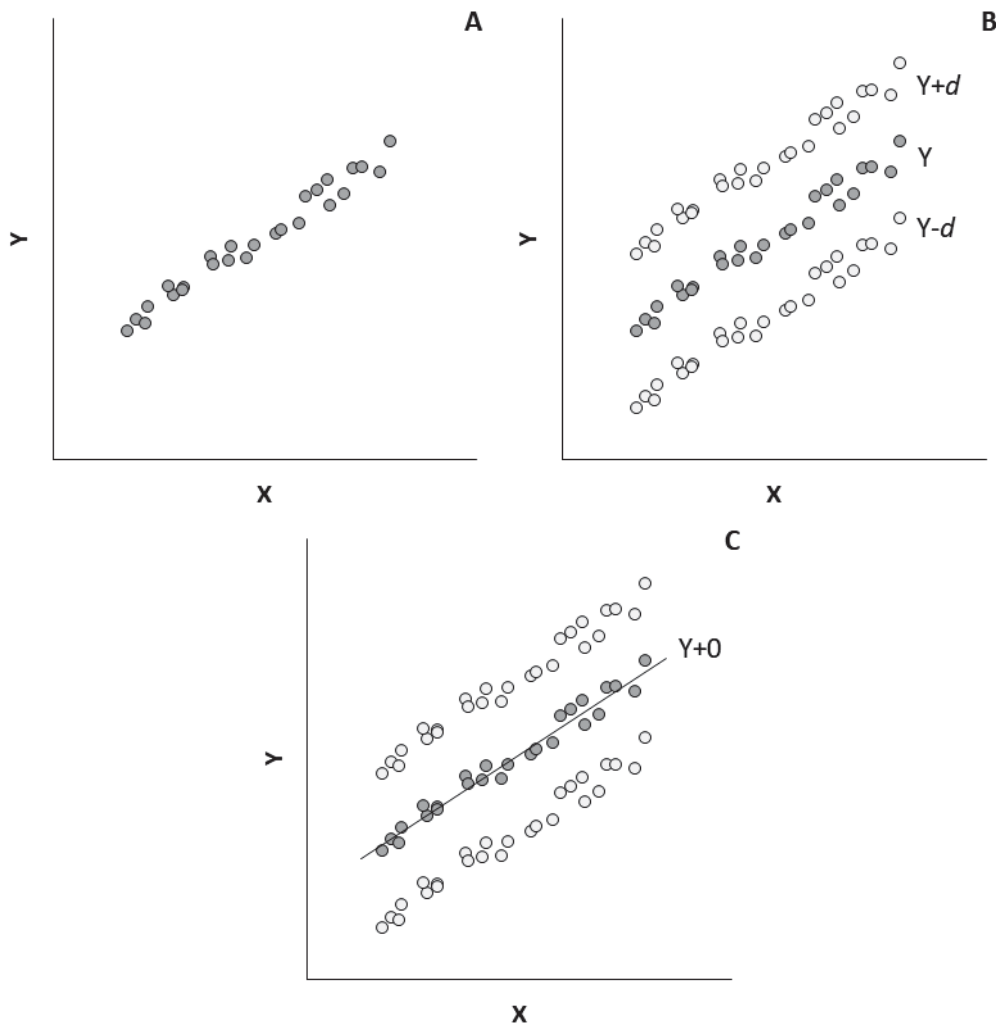
Ao selecionar a FBR, é necessário otimizar o hiperparâmetro γ (gamma), uma vez que este, quando subestimado, reduz a capacidade de não linearidade do mapeamento das variáveis no espaço de características, e quando superestimado, torna o modelo hipersensível a ruídos e pode ocasionar um sobreajuste aos dados de treinamento. São otimizáveis também os parâmetros γ e β na função sigmoide, e, ainda, o parâmetro d na polinomial (SOUZA, 2010).

Para a resolução de problemas de regressão, algumas alterações são realizadas no método original. Considerando o conjunto de dados ilustrado na Figura 6A, em que y_i é uma variável contínua, uma constante d deve ser somada e subtraída de y_i , de modo a formar uma classe positiva e outra negativa (Figura 6B). Deste modo, o conjunto formado pelos valores $y_i + d$ é o positivo e $y_i - d$ é o negativo, permitindo que o problema possa ser tratado como uma classificação binária em que o hiperplano de máxima separação passe por $y_i + 0$, ou seja, pelos dados originais (Figura 6C).

Esta abordagem requer alterações na função objetivo, a qual recebe duas variáveis de folga: ξ_i e ξ_i^* , que representam os erros positivos e negativos respectivamente, e que são delimitados por uma tolerância ε , a qual forma margens no entorno do hiperplano de separação (linha de regressão) que delimitam a área onde os erros de predição são tolerados (Figura 7). Nesse caso, para resolução do problema podem ser utilizadas as seguintes expressões:

$$\begin{aligned} & \text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ & \text{sujeito a } \begin{cases} w \cdot x + b \geq +1 - \xi_i \\ w \cdot x + b \leq -1 + \xi_i \\ \xi_i > 0 \end{cases} \end{aligned}$$

FIGURA 6 - CONJUNTO DE DADOS PARA REGRESSÃO (A); CRIAÇÃO DOS CONJUNTOS NEGATIVO E POSITIVO A PARTIR DA SUBTRAÇÃO E DA ADIÇÃO DA CONSTANTE D (B); HIPERPLANO DE MÁXIMA SEPARAÇÃO (C).

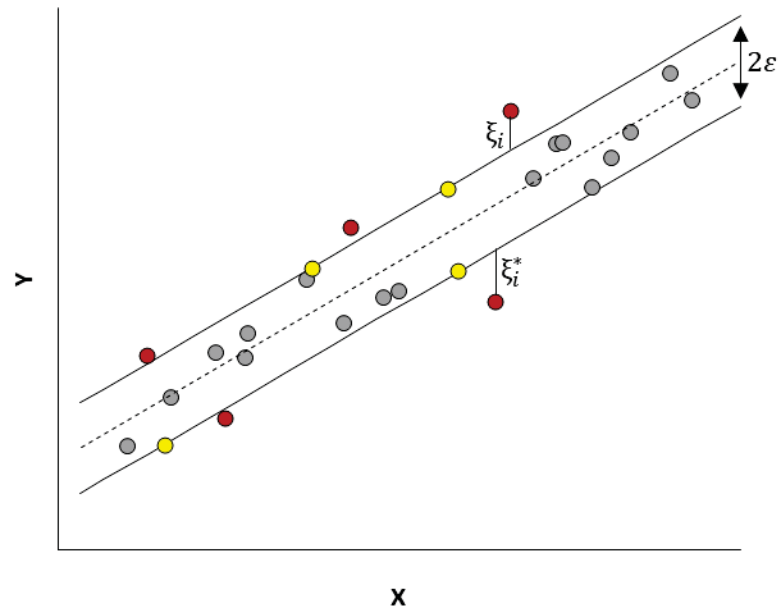


FONTE: O autor (2019).

A função objetivo apresentada anteriormente, refere-se ao algoritmo conhecido como *e*-regression (*epsilon*-regression), o qual permite, também, a otimização do das margens ε , que influenciam diretamente no número de vetores de suporte. Há ainda uma reformulação do método conhecida como *v*-regression (*nu*-regression), que foi desenvolvida com o objetivo de controlar o número de vetores de suporte por meio do parâmetro ν , que representa o limite inferior de proporção entre a quantidade de vetores de suporte e número total de amostras. Ao ser acrescentado à função objetivo, esta toma a seguinte expressão:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \left[\nu \varepsilon + \frac{1}{N} \sum_{i=1}^N (\xi_i + \xi_i^*) \right]$$

FIGURA 7 - LIMITE DE TOLERÂNCIA DETERMINADO POR ξ EM MÁQUINAS DE VETORES DE SUPORE.



FONTE: O autor (2019).

Montaño (2016), analisando o desempenho de algoritmos de aprendizagem de máquina e comparando-os aos modelos alométricos tradicionais para estimativas de atributos florestais, observou que as máquinas de vetores de suporte resultaram em desempenho superior para estimativas volumétricas de plantios de *Pinus taeda* e avaliação de biomassa em plantios de Acácia-negra, assim como para relações hipsométricas em ambas as espécies. Entretanto, a utilização dessa tecnologia para estimativas de biomassa florestal é incipiente e não existem muitos trabalhos em torno dessa matéria, embora os trabalhos existentes apresentem resultados encorajadores (GLEASON e IM, 2012).

3.2.3 FLORESTA ALEATÓRIA

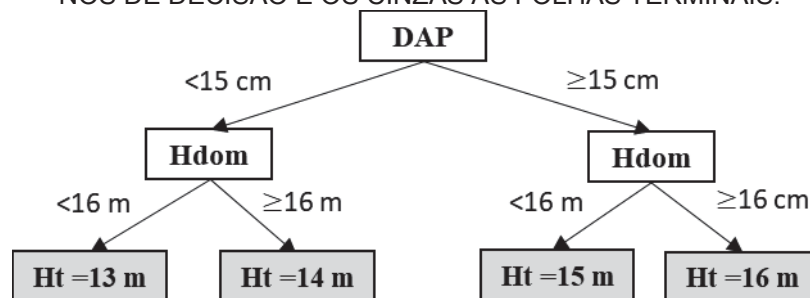
O algoritmo Floresta Aleatória (FA), criado por Breiman (2001), é um método não paramétrico para regressão ou classificação, onde um conjunto de árvores de regressão ou decisão é construído a partir de dados de treinamento.

As árvores de regressão são modelos hierárquicos de aprendizado supervisionado compostos internamente por nós de decisão e folhas terminais (Figura 8). Cada nó de decisão aplica um teste aos dados de entrada, de modo a seguir por

um dos ramos que levará a um novo nó de decisão ou à uma folha terminal, a qual determina o valor da variável de interesse (ALPAYDIN, 2010).

O algoritmo de construção de árvores de regressão funciona de modo a dividir recursivamente o conjunto de treinamento, selecionando a cada nó de divisão uma ou mais variáveis preditoras para particionamento, de modo a minimizar o erro de predição. O algoritmo é executado até que atinja um critério de parada, como o número máximo de nós de divisão (CHEN e ISHWARAN, 2012).

FIGURA 8 - EXEMPLO DE ÁRVORE DE REGRESSÃO PARA PREDIÇÃO DE ALTURA TOTAL, EM QUE AS SETAS REPRESENTAM OS RAMOS, OS RETÂNGULOS BRANCOS REPRESENTAM OS NÓS DE DECISÃO E OS CINZAS AS FOLHAS TERMINAIS.

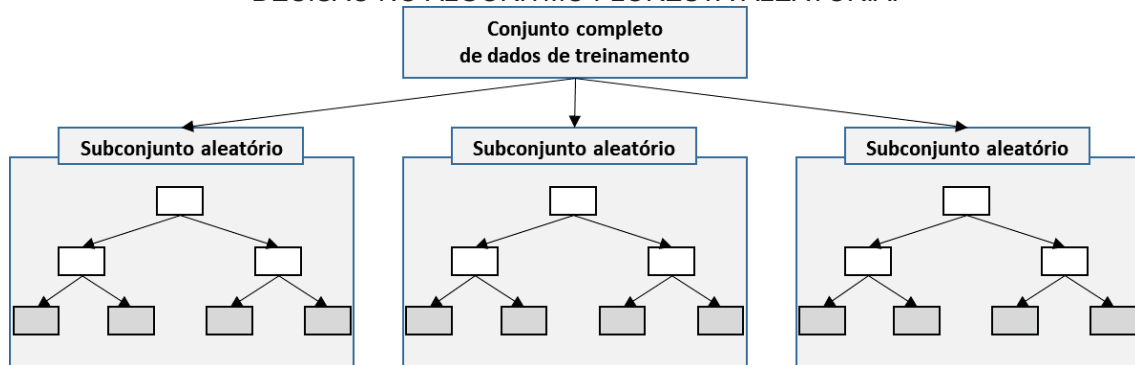


FONTE: O autor (2019).

O algoritmo Floresta Aleatória consiste na geração de N árvores de regressão ou decisão, de modo que cada árvore é treinada a partir de um subconjunto de dados amostrados aleatoriamente e com reposição a partir do conjunto de treinamento (Figura 9). A predição da variável dependente para novos dados é realizada por meio da inserção dos dados de entrada em todas as árvores criadas e consolidação dos resultados a partir da média, para os casos de regressão, e a partir de votação, para os casos de classificação. Ao método de construção de vários regressores ou classificadores por meio do particionamento da base de dados, e consolidação dos resultados pela média ou por votação, dá-se o nome de *bagging*.

Na Floresta Aleatória, dois hiperparâmetros devem ser ajustados a priori. Primeiramente, deve-se decidir quantas árvores construir (n) e quantas variáveis preditivas utilizar em cada nó (m). Um terceiro hiperparâmetro pode ainda ser otimizado: o número máximo de nós (ms), o qual determinará o crescimento máximo da árvore de regressão.

FIGURA 9 - ILUSTRAÇÃO DO MÉTODO DE CRIAÇÃO DE ÁRVORES DE REGRESSÃO OU DECISÃO NO ALGORITMO FLORESTA ALEATÓRIA.



FONTE: O autor (2019).

Devido à eficiência computacional e potenciais aplicações para regressão, análise de sobrevivência, agrupamentos e classificação, o algoritmo Floresta Aleatória tem sido estudado para modelagem de distribuição de espécies (CLUTER *et al.*, 2007), detecção e mapeamento de plantas invasoras em plantios comerciais (PEERBHAY *et al.*, 2015) e atributos de árvores individuais como DAP, altura e volume em ecossistemas florestais, associados a dados obtidos por sensoriamento remoto (LAFITI *et al.*, 2010; YU *et al.*, 2011).

3.2.4 REDES NEURAIS ARTIFICIAIS

A partir de estudos sobre neurofisiologia, McCulloch e Pitts (1943) introduziram um algoritmo de aproximação universal de funções. Esse algoritmo, conhecido como Redes Neurais Artificiais (RNA), consiste de diversos processadores conectados, ou neurônios, que produzem uma sequência de ativações de valores reais sobre uma rede. Neurônios de entrada são ativados através de sensores que percebem o ambiente e outros neurônios da rede são ativados através de conexões, com pesos diferentes, provenientes de neurônios anteriores da sequência (SCHMIDHUBER, 2015). Desde suas aplicações em finanças na década de 1990, as RNA têm se tornado bastante populares, pois são consideradas modelos não paramétricos de um ponto de vista estatístico, o que as torna bastante flexíveis na modelagem de fenômenos do mundo real, onde as observações geralmente estão disponíveis (RAMOS e MARTÍNEZ, 2013).

Nas Redes Neurais Artificiais, cada neurônio é uma unidade de processamento com várias entradas e uma saída. Considerando um neurônio da rede

em que $x = x_1, x_2, \dots, x_n$ é um vetor sinalizador com n canais de entrada, então a soma ponderada das entradas x é calculada pela multiplicação de cada elemento x_i por um peso w_i , tal que a contribuição do canal de entrada i seja adequadamente expressa.

A ativação do neurônio α na rede, é dada por:

$$\alpha = \sum_{i=1}^n w_i x_i + b$$

onde b é o *bias*, uma constante corretiva que impede α de assumir um valor negativo quando todos os elementos de x são iguais a zero. O valor da saída α é calculado através de uma função de ativação $f(\alpha)$.

No Quadro 3 estão apresentadas funções de ativação comumente utilizadas em Redes Neurais Artificiais.

QUADRO 4 - FUNÇÕES DE ATIVAÇÃO COMUMENTE UTILIZADAS NO TREINAMENTO DE REDES NEURAS ARTIFICIAIS.

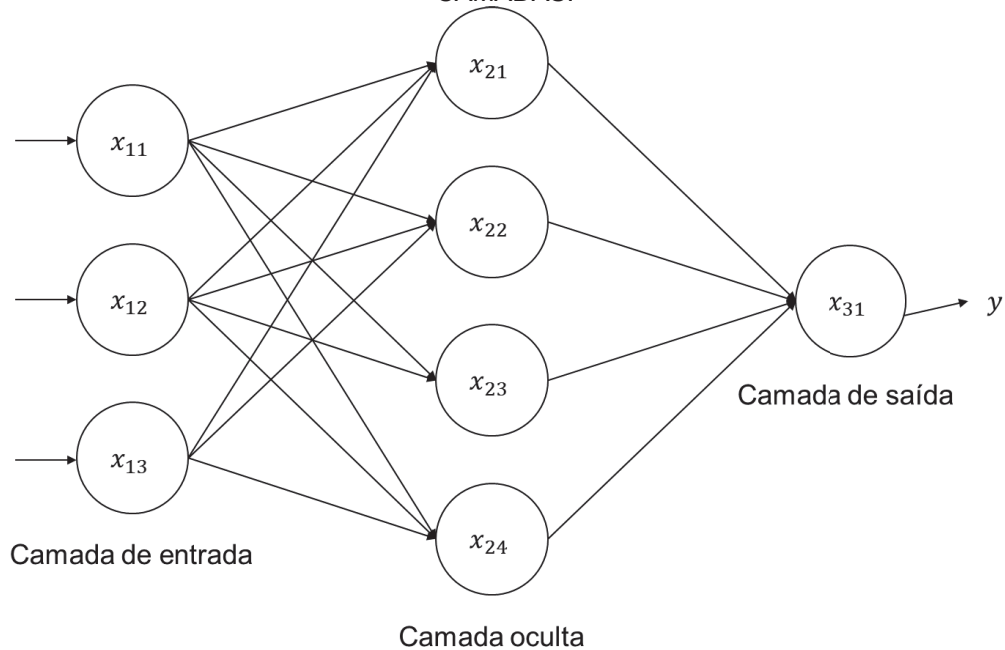
Função de ativação	Expressão
Sigmóide	$f(\alpha) = \frac{1}{1 + e^\alpha}$
Tangente hiperbólica	$f(\alpha) = \frac{e^\alpha - e^{-\alpha}}{e^\alpha + e^{-\alpha}}$
Unidade linear retificada (ReLU)	$f(\alpha) = \max(0, \alpha)$
Unidade linear retificada com vazamento (Leaky ReLu)	$f(\alpha) = \max(\theta, \alpha)$
<i>Maxout</i>	$f(\alpha) = \max(\theta_1, \theta_2)$

FONTE: O autor (2019).

NOTA: α = entrada da função de ativação; θ = inclinação da porção negativa da função ReLu.

Uma variação bastante utilizada do algoritmo de RNA é o *Perceptron* de Múltiplas Camadas, o qual consiste em uma camada de entrada, uma camada de saída e uma ou mais camadas de neurônios entre a entrada e a saída. O número de camadas e a quantidade de neurônios na arquitetura da rede depende de vários fatores, e não existem regras gerais que determinam sua morfologia, sendo esse processo muitas vezes executado por tentativa e erro (CHOJACZYK, 2015). Na Figura 10 está representada uma arquitetura hipotética de uma Rede Neural Artificial de Múltiplas Camadas.

FIGURA 10 - ARQUITETURA HIPOTÉTICA DE UMA REDE NEURAL ARTIFICIAL DE MÚLTIPLAS CAMADAS.



FONTE: O autor (2019).

O método mais utilizado na construção de RNA de múltiplas camadas consiste na utilização de um algoritmo introduzido por Rumelhart *et al.* (1986). Esse algoritmo, conhecido como algoritmo de retropropagação dos erros, utiliza o procedimento de gradiente descendente para atualização dos pesos sinápticos.

Dado um conjunto de dados, onde atributos e rótulos são conhecidos para treinamento da rede, pode-se variar os valores dos pesos ao longo dos números reais e computar o erro correspondente, expressado como a distância entre o rótulo e o valor estimado pela rede. Nesse caso, uma superfície tridimensional é construída, onde os pesos compõem os eixos da abscissa e da ordenada, e o erro compõe a cota. O gradiente descendente é a aplicação de derivadas parciais que buscam encontrar o mínimo global dessa superfície de maneira a obter pesos que produzam os menores erros possíveis (GARDNER e DORLING, 1998).

Diamantopoulou e Milios (2010), estudando o desenvolvimento de modelos para volume total de árvores dominantes de Pinus em um plantio na Grécia, identificaram nas RNA's erros bastante reduzidos nos estimadores gerados. Particularmente, nesse estudo, tecnologias de análise multivariada foram utilizadas para apreender as variáveis que melhor explicavam o comportamento do volume e, a partir dessas variáveis, foram gerados modelos de Redes Neurais Artificiais. A diferença entre os modelos de RNA com todas as variáveis obtidas do campo e os

modelos de RNA com as variáveis resgatadas das análises multivariadas não foi significativa, entretanto, a redução do número de variáveis por efeito do processamento multivariado foi da ordem de 60%, o que implica em uma redução significativa de possíveis mensurações em campo. Nesse cenário, as RNA's, embora ainda tenham uma amplitude modesta nas ciências florestais, têm sido empregadas em diferentes estudos de modelagem florestal e apresentado resultados que encorajam a pavimentação de investigações científicas sobre suas potenciais aplicações nas diferentes faces do setor florestal (SILVA *et al.*, 2009; MIGUEL *et al.*, 2015; CAMPOS *et al.*, 2016; REIS *et al.*, 2018).

3.3 MODELOS NÃO LINEARES

Os sistemas de origem biológica exibem comportamentos que são descritos frequentemente por modelos de regressão não-lineares (SOUZA, 1998). Por definição, esses modelos são aqueles que apresentam não linearidade em seus parâmetros, por exemplo:

$$y_i = \beta_1 e^{\left(\frac{-\beta_2}{x_i} + \beta_3\right)} + \varepsilon_i$$

Em que:

- y_i = variável dependente;
- e = base do logaritmo natural;
- x_i = variável independente;
- ε_i = erro associado;
- β_1, β_2 e β_3 = parâmetros do modelo.

A tarefa de estimação dos parâmetros de um modelo não-linear é mais difícil e intensiva do ponto de vista computacional que os modelos lineares e, assim, usualmente opta-se por linearizar os modelos não-lineares usando transformações logarítmicas ou outras (PAYANDEH, 1983). Entretanto, muitos inconvenientes emergem da linearização desses modelos; dificuldades nas estimativas do coeficiente de determinação e no controle do erro padrão da estimativa, além da geração de modelos tendenciosos. Deve-se ressaltar que, ademais, em alguns casos, a

linearização de modelos não-lineares não é possível, sendo esses modelos intrinsecamente não-lineares (PAYANDEH, 1981), como o modelo acima descrito.

Os métodos de estimativa dos parâmetros de um modelo não-linear são diversos. Os mais conhecidos são os métodos de Gauss-Newton, *Steepest-Decent*, ou método do gradiente, e o método de Marquardt (BATES e WATTS, 1988). Mazucheli e Achcar (2002) argumentam que pelo fato de que esses métodos utilizam derivadas parciais da função esperança, em relação a cada parâmetro, suas aplicações podem ser restringidas, uma vez que, usualmente, a função esperança é bastante complexa e, assim, soluções numéricas para as derivadas são uma opção factível.

Os modelos não-lineares possuem características que os possibilitam descrever comportamentos biológicos de modo satisfatório. Um exemplo disso diz respeito aos processos de crescimento biológico, os quais apresentam um padrão sigmoidal no tempo, e podem ser descritos, por exemplo, através da função de crescimento de Richards (1959), descrita pela seguinte expressão.

$$y_i = \beta_0(1 - e^{-\beta_1 x_i})^{\beta_2} + \varepsilon_i$$

Em que:

y_i = o crescimento de uma variável de interesse;

x_i = o tempo ou idade;

É comum na literatura a utilização desses modelos para estudos de crescimento de povoamentos florestais em relação a classes diamétricas, alturas dominantes e áreas basais que, em última análise, permitem a construção de inferências sobre os volumes de madeira futuros (ZHU *et al.*, 2003; EISFELD *et al.*, 2005; TONINI *et al.*, 2009; RICKEN *et al.*, 2011; MCCULLAGH *et al.*, 2017). A título de exemplo, Kudus *et al.* (1999) demonstram que os modelos de regressão não-lineares podem ser utilizados, também, nas estimativas dos parâmetros da função de distribuição probabilística Johnson S_B que, no que concerne a modelagem de distribuição diamétrica de povoamentos florestais, é amplamente utilizada (PALAHÍ *et al.*, 2007; MACHADO *et al.*, 2010; MARCON *et al.*, 2014), e na estimativa da altura de árvores (SENA *et al.*, 2015).

3.4 RELAÇÃO HIPSOMÉTRICA

A relação existente entre o diâmetro e altura é reconhecida em todas as plantas terrestres, sendo a altura do vegetal expressa em função de seu diâmetro de maneiras diversas para diferentes grupos taxonômicos (NIKLAS, 1993). Do ponto de vista ecológico, essa relação é usualmente explicada, nas plantas vasculares, em função de restrições hidráulicas e biomecânicas (NIKLAS, 1993; HENRY e AARSEN, 1999; BECKER *et al.*, 2000). Particularmente, nas ciências florestais, a relação entre diâmetro e altura é conhecida como relação hipsométrica e possui importância central nos estudos de povoamentos florestais (MACHADO *et al.*, 2008; FIGUEIREDO FILHO *et al.*, 2010; SANQUETTA *et al.*, 2013).

Ribeiro *et al.* (2010) discutem a importância de estudos sobre mensuração florestal em um mercado de produtos florestais que está em plena expansão nacional e internacionalmente. Desse modo, os estudos sobre comportamento dos diferentes modelos hipsométricos, nos diferentes contextos em que os povoamentos florestais estão inseridos, são fundamentais para o processo de planejamento de qualquer empresa florestal. De fato, os modelos volumétricos frequentemente possuem, como uma de suas variáveis dependentes, a altura, a qual é estimada a partir modelos hipsométricos em grandes povoamentos, onde a medição de todas as alturas é impraticável (LEITE e ANDRADE, 2003; RIBEIRO *et al.*, 2010; ANDRADE e LEITE, 2011; MENDONÇA *et al.*, 2015).

Os modelos hipsométricos são diversos nas ciências florestais, assumindo características lineares e não-lineares (BARROS *et al.*, 2002; VENDRUSCOLO *et al.*, 2015). Embora o diâmetro seja a variável tradicionalmente reconhecida como fundamental para estimativa da altura, outras variáveis são frequentemente aditadas aos modelos hipsométricos como, por exemplo, a altura dominante (BARTOSZECK, 2004; RETSLAFF *et al.*, 2015). Essa variável demonstra significativa relevância na melhoria de modelos hipsométricos, principalmente em situações de inventário florestal onde a altura dominante tem variabilidade evidente entre diferentes unidades amostrais (LEITE e ANDRADE, 2003).

Recentemente, principalmente durante a última década, tem-se observado um gradual aumento no interesse sobre o uso de técnicas de inteligência artificial para modelagem florestal, sendo as relações hipsométricas um tema de debate científico

bastante frequente à luz desse novo paradigma (ARAÚJO, 2015; CAMPOS *et al.*, 2016; VENDRUSCOLO *et al.*, 2017). Como argumenta Binoti (2012), a possibilidade do uso de variáveis categóricas, usualmente incluídas nos modelos de regressão como variáveis binárias, é uma vantagem evidente do uso de técnicas de inteligência artificial na construção de modelos hipsométricos. De fato, trabalhos recentes na literatura científica têm apresentado resultados promissores para a modelagem das relações hipsométricas por meio dessas técnicas (BINOTI, 2012; BINOTI *et al.*, 2013a; BINOTI *et al.*, 2013b; CAMPOS *et al.*, 2016; VENDRUSCOLO *et al.*, 2017; MENDONÇA *et al.*, 2018).

4 REFERÊNCIAS

- ALPAYDIN, E. **Introduction to machine learning**. 3. Ed. Boston: MIT press, 2009. 616 p.
- ALVES, J. de A. *et al.* Equações hipsométricas para *Toona ciliata* com inclusão de covariantes. **Ciência Florestal**, Santa Maria, v. 27, n. 2, p. 581-595, 2017.
- ANDRADE, V. C. L. D.; LEITE, H. G. Hipsometric relationship modeling using data sampled in tree scaling and inventory plots. **Revista Árvore**, v. 35, n. 1, p. 157-164. 2011.
- ARAÚJO, B. H. P. D. **Modelagem da altura, volume e afilamento do fuste de *Calycophyllum Spruceanum* Benth. empregando regressão e redes neurais artificiais**. Dissertação (Mestrado em Recursos Florestais e Engenharia Florestal) – Instituto Nacional de Pesquisas da Amazônia – INPA, Manaus, 2015.
- BARTOSZECK, A. C. P. S.; MACHADO, S.A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E.B. Dinâmica da relação hipsométrica em função da idade, do sítio e da densidade inicial de povoamentos de bracatinga da Região Metropolitana de Curitiba, PR. **Revista Árvore**, v. 28, n. 4, p. 517-533. 2004.
- BARROS, D. D.; MACHADO, S. D. A.; ACERBI JUNIOR, F. W.; SCOLFORO, J. R. S. Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. **Boletim de Pesquisa Florestal**, n. 45, p. 3-28. 2002.
- BARTOSZECK, A. C. P. S.; *et al.* Dinâmica da relação hipsométrica em função da idade, do sítio e da densidade inicial de povoamentos de bracatinga da região metropolitana de Curitiba, PR. **R. Árvore**, v. 28, n. 4, p. 517-533, 2004.
- BATES, D. M.; WATTS, D. G. Nonlinear regression: iterative estimation and linear approximations. In: **Nonlinear Regression Analysis and Its Applications**. New York: Wiley, 1988. cap. 2, p. 32-66.
- BEAUDOIN, A. *et al.* Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. **Canadian Journal of Forest Research**, v. 44, n. 5, p. 521-532, 2014.
- BECKER, P.; MEINZER, F. C.; WULLSCHLEGER, S. D. Hydraulic limitation of tree height: a critique. **Functional Ecology**, v. 14, n. 1, p. 4-11. 2000.
- BHATIA, N. Survey of nearest neighbor techniques. **International Journal of Computer Science and Information Security**, v. 8, n. 2, p. 302-305, 2010.
- BINOTI, D. H. B., BINOTI, M L. M. da S.; LEITE, H. G.; SILVA, A. Redução dos custos em inventário de povoamentos equiâneos. **Revista Brasileira de Ciências Agrárias**, v. 8, n.1, p.125-129, 2013a.

BINOTI, M. **Emprego de redes neurais artificiais em mensuração e manejo florestal**. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 2012.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, v. 37, n. 4, p. 639-645. 2013b.

BREIMAN, L. Stacked regressions. **Machine learning**, v. 24, n. 1, p. 49-64, 1996.

BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5-32, 2001.

CALEGARIO, N. *et al.* Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não-lineares generalizados. **Scientia Forestalis**, v. 69, p. 38-50, dez., 2005.

CAMPOS, B. P. F. *et al.* Predição da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. **Pesquisa Florestal Brasileira**, v. 36, n. 88, p. 375-385, 2016.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 3. ed. Viçosa, MG: UFV, 2009. 548 p.

CHEN, X.; ISHWARAN, H. Random forests for genomic data analysis. **Genomics**, v. 99, n. 6, 323-329, 2012.

CHIRICI, G. *et al.* A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. **Remote Sensing of Environment**, v. 176, p. 282-294, 2016.

CHOJACZYK, A. A.; TEIXEIRA, A. P.; NEVES, L. C.; CARDOSO, J. B.; SOARES, C. G. Review and application of artificial neural networks models in reliability analysis of steel structures. **Structural Safety**, v. 52, p. 78-89, 2015.

CORTES, C; VAPNIK, V. N. Support-Vector Networks. **Machine Learning**. 2 ed. New York: Springer, 1995, 314 p.

CUTLER, D. R.; EDWARDS JR, T. C.; BEARD, K. H.; CUTLER, A.; HESS, K. T.; GIBSON, J.; LAWLER, J. J. Random forests for classification in ecology. **Ecology**, v. 88, n. 11, p. 2783-2792, 2007.

DIETTERICH, T. G. Ensemble methods in machine learning. In: International workshop on multiple classifier systems, 2010. **Multiple Classifier Systems**. Berlin: Springer, p. 1-15, 2010.

GARDNER, M. W.; DORLING, S. R. Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. **Atmospheric environment**, v. 32, n. 14-15, p. 2627-2636, 1998.

DIAMANTOPOULOU, M. J.; MILIOS, E. Modelling total volume of dominant pine trees in reforestations via multivariate analysis and artificial neural network models. **Biosystems engineering**, v. 105, n. 3, p. 306-315, 2010.

EISFELD, R.L.; SANQUETTA, C. R.; ARCE, J. E.; MAESTRI, R.; WEBER, K. S. Modelagem do crescimento e da produção de *Pinus taeda* L. por meio de função probabilística. **Floresta**, v. 35, n. 2, p. 167-177. 2005.

FIGUEIREDO FILHO, A.; DIAS, A. N.; KOHLER, S. V.; VERUSSA, A. A.; CHIQUETTO, A. L. Evolution of the hypsometric relationship in *Araucaria angustifolia* plantations in the mid-south region of Paraná state. **Cerne**, v. 16, n. 3, p. 347-357. 2010.

GLEASON, C. J.; IM, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. **Remote Sensing of Environment**, v. 125, p. 80-91, 2012.

HANSEN, L. K.; SALAMON, P. Neural network ensembles. **Transactions on pattern analysis and machine intelligence**, v. 12 n. 10, p. 993-1001, 1990.

HEARST, M. A. *et al.* Support vector machines. **IEEE Intelligent Systems**, IEEE Computer Society, v. 13, n. 4, p. 18–28, 1998.

HECHENBICHLER, K.; SCHLIEP, K. **Weighted k-nearest-neighbor techniques and ordinal classification**. Munchen: Ludwig-Maximilians-Universitat, 2004. 17 p.

HENRY, H. A. L.; AARSSSEN, L. W. The interpretation of stem diameter–height allometry in trees: biomechanical constraints, neighbour effects, or biased regressions? **Ecology letters**, v. 2, n. 2, p. 89-97. 1999.

HUANG, C. L.; WANG, C. J. A GA-based feature selection and parameters optimization for support vector machines. **Expert Systems with applications**, v. 31, n. 2, p. 231-240, 2006.

KUDUS, K. A.; AHMAD, M. I.; LAPONGAN, J. Nonlinear regression approach to estimating Johnson SB parameters for diameter data. **Canadian Journal of Forest Research**, v. 29, n. 3, p. 310-314. 1999.

LARRANAGA, P. *et al.* Machine learning in bioinformatics. **Briefings in bioinformatics**, v. 7, n.1, p. 86-112, 2006.

LATIFI, H.; NOTHDURFT, A.; KOCH, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. **Forestry**, v. 83 n. 4, p. 395-407, 2010.

LEITE, H. G.; ANDRADE, V. D. Importância das variáveis altura dominante e altura total em equações hipsométricas e volumétricas. **Revista Árvore**, v. 27, n. 3, p. 301-310. 2003.

LORENA, A. C.; CARVALHO, A. C. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43-67, 2007.

MACHADO, S. A.; NASCIMENTO, R. G. M.; AUGUSTYNCZIK, A. L. D.; DA SILVA SILVA, L. C. R.; FIGURA, M. A.; PEREIRA, E. M.; TÉO, S. J. Comportamento da relação hipsométrica de *Araucaria angustifolia* no capão da Engenharia Florestal da UFPR. **Pesquisa Florestal Brasileira**, n. 56, p. 5, 2008.

MACHADO, S. A.; DOS SANTOS, A. A. P.; DO NASCIMENTO, R. G. M.; AUGUSTYNCZIK, A. L. D.; ZAMIN, N. T. Modelagem da distribuição diamétrica de quatro espécies de Lauraceae em um fragmento de Floresta Ombrófila Mista. **RECEN-Revista Ciências Exatas e Naturais**, v. 12, n. 1, p. 91-105, 2010.

MARCON, F.; TÉO, S. J.; SCHNEIDER, C. R.; DOS SANTOS, F. B.; CHIARELLO, K. M. A.; FIORENTIN, L. D. Modelagem da distribuição diamétrica de um fragmento de Floresta Ombrófila Mista em Lebon Régis, SC. **Floresta**, v. 45, n. 2, p. 337-348. 2014.

MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. **Acta Scientiarum**, v. 24, n. 6, p. 1761-1770. 2002.

MCCULLAGH, A.; BLACK, K.; NIEUWENHUIS, M. Evaluation of tree and stand-level growth models using national forest inventory data. **European journal of forest research**, v. 136, n. 2, p. 251-258, 2017.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MENDONÇA, A. R.; DE PÁDUA CHAVES, S.; CARVALHO, N. C. Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. **Cerne**, v. 21, n. 1, p. 107-115. 2015.

MENDONÇA, A. R.; DA SILVA, J. C.; AOZAI, T. S.; DA SILVA, E. R.; SANTOS, J. S.; BINOTI, D. H. B.; DA SILVA, G. F. Estimação da altura total de árvores de ipê felpudo utilizando modelos de regressão e redes neurais. **Revista Brasileira de Biometria**, v. 36, n. 1, p. 128-139. 2018.

MEZA, A. M. A.; SANTACOLOMA, G. D.; MEDINA, C. D. A.; DOMÍNGUEZ, G. C. Parameter selection in least squares-support vector machines regression oriented, using generalized cross-validation. **Dyna**, v. 79, n. 171, p. 23-30, 2012.

MONTAÑO, R. A. N. R. **Aplicação de técnicas de aprendizado de máquina na mensuração florestal**. Tese (doutorado). Universidade Federal do Paraná, 2016.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. In: REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. Editora Manole Ltda., 2003, p. 32.

MIGUEL, E. P.; REZENDE, A. V.; LEAL, F. A.; MATRICARDI, E. A. T.; VALE, A. T. D.; PEREIRA, R. S. Artificial neural networks for modeling wood volume and aboveground

biomass of tall Cerrado using satellite data. **Pesquisa Agropecuária Brasileira**, v. 50, n. 9, p. 829-839, 2015.

NIKLAS, K. J. The scaling of plant height: a comparison among major plant clades and anatomical grades. **Annals of botany**, v. 72, n. 2, p. 165-172, 1993.

PALAHÍ, M.; PUKKALA, T.; BLASCO, E.; TRASOBARES, A. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). **European Journal of Forest Research**, v. 126, n. 4, p. 563-571. 2007.

PAYANDEH, B. Some applications of nonlinear regression models in forestry research. **The Forestry Chronicle**, v. 59, n. 5, p. 244-248. 1983.

PAYANDEH, B. Choosing regression models for biomass prediction equations. **The Forestry Chronicle**, v. 57, n. 5, p. 229-232. 1981.

PEERBHAY, K. Y.; MUTANGA, O.; ISMAIL, R. Random forests unsupervised classification: The detection and mapping of *Solanum mauritianum* infestations in plantation forestry using hyperspectral data. **Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 8 n. 6, p. 3107-3122, 2015.

RAMOS, E. G.; MARTÍNEZ, F. V. A Review of Artificial Neural Networks: How Well Do They Perform in Forecasting Time Series. **Análitika: revista de análisis estadístico**, v. 6, p. 7-18, 2013.

RETSLAFF, F. A. S.; FIGUEIREDO FILHO, A.; DIAS, A. N.; BERNETT, L. G.; FIGURA, M. A. Curvas de sítio e relações hipsométricas para *Eucalyptus grandis* na região dos Campos Gerais, Paraná. **Cerne**, v. 21, n. 2, p. 219-225. 2015.

REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. Editora Manole Ltda, 2003.

REIS, L. P. *et al.* Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. **Ecological Engineering**, v. 112, p. 140-147, 2018.

REESE, H.; GRANQVIST-PAHLÉN, T.; EGBERTH, M.; NILSSON, M.; OLSSON, H. Automated estimation of forest parameters for Sweden using Landsat data and the kNN algorithm. In: 31st International Symposium on Remote Sensing of Environment, 2005.

RIBEIRO, A., FERRAZ FILHO, A. C., MELLO, J. M. D., ZÉLIA FERREIRA, M., MARTINS LISBOA, P. M., & SOARES SCOLFORO, J. R. Estratégias e metodologias de ajuste de modelos hipsométricos em plantios de *Eucalyptus* sp. **Cerne**, v. 16, n. 1, p. 22-31. 2010.

RICKEN, P.; MATTOS, P. P.; BRAZ, E.; KALIL FILHO, A. N. **Crescimento diamétrico de povoamento de *Toona ciliata* var. *australis* em Adrianópolis, PR**. Paraná: Embrapa Florestas, 2011. Relatório técnico.

RICHARDS, F. J. A flexible growth function for empirical use. **Journal of experimental Botany**, v. 10, n. 2, p. 290-301. 1959.

RUMELHART, D.E.; HINTON, G.E.; WILLIAMS, R.J. Learning internal representations by backpropagating errors. **Nature**, v. 323, p. 533–536, 1986.

SANQUETTA, C. R.; DALLA CORTE, A. P.; ROGL, A.; PIMENTEL, A. Relações diâmetro-altura para espécies lenhosas em um fragmento de floresta Ombrófila Mista no Sul do Paraná. **Série Botânica**, v. 68, n. 1, p. 103-114. 2013.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural networks**, v. 61, p. 85-117, 2015.

SCHNEIDER, P. R. Manejo florestal I – Tópicos para o planejamento da produção florestal. Santa Maria: Universidade Federal de Santa Maria, 1986. 289p.

SCOLFORO, J. R. S. **Biometria florestal**: parte I: modelos de regressão linear e não linear: parte II: modelos para relação hipsométrica, volume, afilamento e peso da matéria seca. Lavras: UFLA/FAEPE, 2005. 352 p.

SENA, A. L. M.; SILVA NETO, A. J. da; OLIVEIRA G. M. V.; CALEGARIO, N. Modelos não lineares com uso de covariantes para relação hipsométrica de duas espécies de *Pinus* locais. **Ciência Florestal**, Santa Maria, v. 25, n. 4, p. 969-980, out./dez., 2015.

SILVA, G. F. da.; XAVIER, A. C.; RODRIGUES, F. L.; PETERNELLI, L. A. Análise da influencia de diferentes tamanhos e composições de amostras no ajuste de uma relação hipsométrica. **Revista Árvore**, v.31, n.4, p.685-694, 2007.

SILVA, C. A. *et al.* Estimating Stand Height and Tree Density in *Pinus taeda* plantations using in-situ data, airborne LiDAR and k-Nearest Neighbor Imputation. **Anais da Academia Brasileira de Ciências**, v. 90, n. 1, p. 295-309, 2018.

SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Adjustment of the Schumacher and Hall model and application of artificial neural networks to estimate volume of eucalypt trees. **Revista Árvore**, v. 33, n. 6, 1133-1139, 2009.

SOUZA, C. R. Kernel Functions for Machine Learning Applications. César Souza. 2010. Disponível em: <<http://crsouza.com/2010/03/17/kernel-functions-for-machine-learning-applications/>>. Acesso em: 7 nov. 2018.

SOUZA, G. S. **Introdução aos modelos de regressão linear e não-linear**. Brasília: Embrapa-SPI/Embrapa-SEA, 1998.

TONINI, H.; COSTA, M. C. G.; SCWENGBER, L. A. M. Crescimento da teca (*Tectona grandis*) em reflorestamento na Amazônia Setentrional. **Pesquisa Florestal Brasileira**, n. 59, p. 05, 2009.

VAPNIK, V. N. **The nature of Statistical learning theory**. Springer-Verlag, New York, 1995.

VENDRUSCOLO, D. G. S.; CHAVES, A. G. S.; MEDEIROS, R. A.; SILVA, R. D.; SOUZA, H. S.; DRESCHER, R.; LEITE, H. G. Estimativa da altura de árvores de *Tectona grandis* Lf utilizando regressão e redes neurais artificiais. **Nativa, Sinop**, v. 5, n. 1, p. 52-58. 2017.

VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S.; MOURA, J. P. V. M.; MAMORÉ, F. M. D.; SIQUEIRA, T. D. S. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v. 33, n. 4, p. 556-569. 2015.

WOJCIECHOWSKI, J. JCarbon – *Software na Web com Data Mining* para estimativas de volume, biomassa e carbono em florestas. Tese (Doutorado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba, 2015.

YU, X.; HYYPPÄ, J.; VASTARANTA, M.; HOLOPAINEN, M.; VIITALA, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. **Journal of Photogrammetry and remote sensing**, v. 66, n. 1, p. 28-37, 2011.

ZHU, J. J.; FAN, Z. P.; ZENG, D. H.; JIANG, F. Q.; MATSUZAKI, T. Comparison of stand structure and growth between artificial and natural forests of *Pinus sylvestris* var. *mongolica* on sandy land. **Journal of Forestry Research**, v. 14, n. 2, p. 103-111. 2003.

CAPÍTULO 1 – TÉCNICAS DE APRENDIZADO DE MÁQUINA E MODELOS NÃO LINEARES COM INCLUSÃO DE COVARIANTES NA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE *Pinus taeda* L.

RESUMO

No presente estudo foram investigadas técnicas de aprendizado de máquina e modelos não lineares com uso de covariantes para construção de modelos hipsométricos genéricos em povoamentos de *Pinus taeda* L. Para tanto, foram utilizados dados de plantios em diferentes idades, localizados na região de Itapeva, São Paulo. Foram aplicados os algoritmos k-Vizinhos Mais Próximos (k-VMP), Floresta Aleatória (FA), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA). O treinamento dos algoritmos de aprendizado de máquina considerou a otimização de hiperparâmetros pelo método de busca em grade combinada à validação cruzada *k-fold*. Os modelos Monomolecular e de Gompertz foram modificados pela decomposição de seus parâmetros para inserção de covariantes. As variáveis idade e altura dominante da parcela foram testadas como covariantes dos parâmetros, utilizando o método *backward* de seleção. Para cada técnica, os melhores modelos gerados foram utilizados para predição da altura total de indivíduos pertencentes a um conjunto de dados independente e os resultados foram comparados aos obtidos por modelos lineares genéricos. As técnicas foram avaliadas quanto ao desempenho de generalização pelos indicadores: erro absoluto médio (EAM) e erro quadrático médio (EQM) percentuais, além de análises de resíduos. Os modelos de aprendizado de máquina resultaram indicadores de desempenho similares aos modelos de regressão estatística, no entanto os resultados foram mais favoráveis à utilização das Redes Neurais Artificiais e Máquinas de Vetores de Suporte. Os modelos não lineares com inserção de covariantes também se destacaram por apresentarem indicadores de desempenho superiores às técnicas k-Vizinhos Mais Próximos e Floresta Aleatória. As técnicas RNA, MVS e os modelos não lineares modificados pela inserção de covariantes demonstraram ser mais apropriados para representar a relação hipsométrica em povoamentos heterogêneos em comparação aos modelos lineares genéricos.

Palavras-chave: Máquinas de Vetores de Suporte; Redes Neurais Artificiais; k-Vizinhos Mais Próximos, Floresta Aleatória.

CHAPTER 1 – MACHINE LEARNING ALGORITHMS AND NON-LINEAR REGRESSION WITH COVARIATE INCLUSION IN THE MODELING OF THE HEIGHT-DIAMETER RELATIONSHIP OF *Pinus taeda* L.

ABSTRACT

The present study aimed to investigate machine learning algorithms and nonlinear models with covariates inclusion to construct generic height-diameter models for *Pinus taeda* L. stands. For this purpose, a biometric data set obtained in stands at different ages, located in the region of Itapeva, São Paulo, was used. The k-Nearest Neighbors (k-VMP), Random Forest (FA), Support Vector Machines (MVS) and Artificial Neural Networks (RNA) algorithms were applied. The optimization of hyperparameters by the grid search method and the k-fold cross validation were applied in the training phase of the algorithms. The Monomolecular and Gompertz models were modified by the decomposition of its parameters for the inclusion of covariates. The age and dominant height variables were tested as covariates, using the backward variable selection method. For each technique, the best models were used to predict the total height of individuals in an independent data set and the results were compared with those obtained by generic linear models. The techniques generalization performances were evaluated by the mean absolute percentage error (EAM%) and mean squared percentage error (EQM%). The machine learning models resulted in similar performance indicators to the statistical regression models, however the results were more favorable to the use of Artificial Neural Networks and Support Vector Machines. The nonlinear models with covariates inclusion also presented favorable results, presenting better performance indicators than the k-Nearest Neighbors and Random Forest techniques. The RNA, MVS and nonlinear models modified by the inclusion of covariates were most suitable to represent the height-diameter relation in heterogeneous stands than generic linear models.

Keywords: Support Vector Machines; Artificial Neural Networks; k-Nearest Neighbors; Random Forest.

1.1 INTRODUÇÃO

Na mensuração florestal, os modelos estatísticos são amplamente utilizados para a predição ou explicação de fenômenos físicos. As observações de um fenômeno geralmente são consideradas representáveis por uma função determinística, na qual uma ou mais variáveis explicativas são empregadas, visando prever o comportamento de uma variável dependente (ROBINSON *et al.*, 2014). Uma categoria de modelos mais sofisticados, conhecidos como algoritmos de aprendizado de máquina, têm sido objeto de estudos para modelagem de variáveis de interesse na mensuração florestal, como altura (BINOTI *et al.*, 2013; VENDRUSCOLO *et al.*, 2015; VENDRUSCOLO *et al.*, 2017) e volume (GORGENS *et al.*, 2014; CORDEIRO *et al.*, 2015).

O aprendizado de máquina (AM) é uma área da inteligência artificial voltada à construção de sistemas capazes de induzir hipóteses ou aproximar funções a partir da experiência acumulada em problemas anteriores (FACELI *et al.*, 2011). As decisões tomadas por algoritmos de AM são baseadas no aprendizado indutivo, que pode ser subdividido em supervisionado, quando o objetivo é a resolução de problemas de classificação ou de regressão, e não supervisionado, quando a tarefa é de agrupamento ou associação (MONARD e BARANAUSKAS, 2003).

A maioria das pesquisas na área de mensuração florestal tem utilizado modelos de AM em tarefas de regressão, sendo as Redes Neurais Artificiais (RNA) as mais estudadas. Na maioria das vezes os resultados obtidos por RNA possuem performance superior aos modelos de regressão tradicionais (BINOTI *et al.*, 2014; GORGENS *et al.*, 2014; MARTINS *et al.*, 2016; VENDRUSCOLO *et al.*, 2016).

Existem técnicas de AM cujas aplicações na mensuração florestal ainda são pouco estudadas. Essas têm sido amplamente investigadas na área de sensoriamento remoto, tendo apresentado resultados satisfatórios.

Por exemplo, k-Vizinhos mais Próximos, Florestas Aleatórias e Máquinas de Vetores de Suporte (MVS) foram utilizadas por Abdollahnejad *et al.* (2017) para predição de espécies de árvores dominantes em florestas naturais no Irã. Já Gorgens *et al.* (2015) empregou os algoritmos Floresta Aleatória (RF, *random forest*), RNA e MVS na predição do volume de povoamentos de eucalipto a partir de dados obtidos por sistema LiDAR.

Outro procedimento tem conferido à modelos não lineares melhor desempenho na predição da altura de árvores: a decomposição de parâmetros para

inserção de covariantes (ALVES *et al.*, 2015; SENA *et al.*, 2015). A técnica de adição de covariantes à modelos biológicos confere melhor desempenho às equações ajustadas, por considerar que os coeficientes de uma equação são influenciados por fatores que afetam diretamente a relação hipsométrica como idade e qualidade de sítio, permitindo também a construção de modelos mais genéricos que representem a relação do DAP com altura de povoamentos com características diferentes.

O uso de técnicas de aprendizado de máquina ou de regressão não linear na modelagem da relação hipsométrica já foi tema de alguns estudos, no entanto ainda carece de discussões aprofundadas, especialmente quanto à configuração dos algoritmos de AM, ao método de seleção de covariantes para os modelos não lineares, e à abordagem comparativa entre as duas técnicas.

Assim, este trabalho teve como objetivos: aplicar modelos não lineares modificados pela inclusão de covariantes e técnicas de aprendizado de máquina à modelagem da relação hipsométrica de povoamentos de *Pinus taeda* com diferentes características; otimizar os hiperparâmetros dos algoritmos de AM visando a maximização do desempenho das técnicas; e comparar o desempenho dos modelos de predição gerados com modelos lineares genéricos.

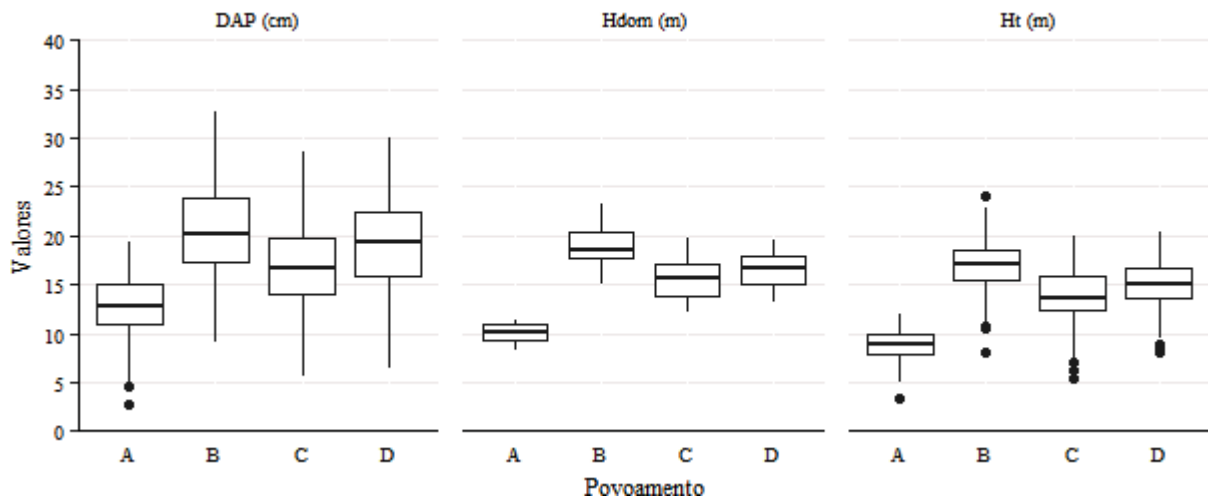
1.2 MATERIAIS E MÉTODOS

1.2.1 BASE DE DADOS

Os dados utilizados neste estudo foram obtidos no ano de 2016 em plantios comerciais de *Pinus taeda*, conduzidos sem desbaste e em espaçamento 3x2 metros, localizados na região de Itapeva, estado de São Paulo. Na ocasião do inventário florestal, foram amostrados povoamentos com quatro idades por meio de parcelas temporárias, sendo mensurados o diâmetro à 1,3 m do solo (DAP) com fita métrica e a altura total (*Ht*) com hipsômetro digital, em 162 unidades amostrais circulares de 200 m². Foram medidas as alturas de 10 árvores por parcela mais duas árvores dominantes.

Na Figura 1.1 está apresentado um resumo estatístico das variáveis dos povoamentos que compõem a base de dados.

FIGURA 1.1 - CARACTERÍSTICAS DA BASE DE DADOS OBTIDA EM POVOAMENTOS DE *Pinus taeda* NA REGIÃO DE ITAPEVA/SP.



NOTA: Idades dos povoamentos: A=6,7; B=7,9; C=9,7; D=11,8.

FONTE: O autor (2019).

Para a avaliação da capacidade de generalização dos modelos ajustados, realizou-se o particionamento dos dados por meio do método *hold-out*, o qual consistiu em dividir a base de dados em dois conjuntos independentes, sendo um para treinamento e outro para teste dos modelos (SHALEV-SHWARTZ e BEN-DAVID, 2014). A razão utilizada para a divisão foi de 1/3 dos dados para teste e 2/3 para treinamento/ajuste dos algoritmos e modelos de regressão.

Para que a separação resultasse em dois grupos de amplitude e variação similares, os diâmetros à altura do peito foram ordenados do maior para o menor dentro de cada povoamento e a divisão foi realizada sistematicamente ao separar a última árvore a cada três.

1.2.2 MODELOS DE REGRESSÃO

1.2.2.1 Modelos lineares

Foram selecionados dois modelos lineares (Quadro 1.1), sendo o ajuste realizado no *software* R. Os modelos selecionados são denominados genéricos, uma vez que geram equações que permitem gerar estimativas para conjuntos de dados originados de povoamentos florestais com características silviculturais diferentes. Esses modelos resultam em precisão superior aos modelos hipsométricos tradicionais, devido contemplarem variáveis do povoamento, como idade e altura dominante, as quais influenciam na relação altura/diâmetro (BARROS *et al.*, 2002).

QUADRO 1.1 - MODELOS LINEARES SELECIONADOS PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA EM POVOAMENTOS DE *Pinus taeda*.

Modelo	Expressão
Linear 1 (Barros <i>et al.</i> , 2002)	$Ht = \beta_0 + \beta_1 DAP + \beta_2 Hdom + \beta_3 id + \beta_3 DAP^2 + \beta_5 id^2 + \varepsilon$
Linear 2 (<i>stepwise</i>)	$ln(Ht) = \beta_0 + \beta_1 \frac{1}{DAP} + \beta_2 ln(Hdom) + \beta_3 id + \beta_4 id^2 + \varepsilon$

FONTE: O autor (2019).

NOTA: *ln* = logaritmo neperiano; *Ht* = altura total; *DAP* = diâmetro a 1,3m do solo; *Hdom* = altura média das árvores dominantes; *id* = idade; ε = erro aleatório inerente à regressão.

1.2.2.2 Modelos não lineares com adição de covariantes

Dois modelos não lineares foram selecionados (Quadro 1.2). Considerando a influência das variáveis idade e altura dominante, os parâmetros dos modelos foram decompostos e procedeu-se à inserção de covariantes, de modo a estimar os parâmetros por meio de funções lineares. As covariantes para cada parâmetro foram selecionadas por procedimento *backward*, sendo mantidas no modelo somente as significativas a 1% de probabilidade.

Os modelos não lineares foram ajustados por meio da função *nls.lm* do pacote *minpack.lm* do R, que utiliza uma modificação do algoritmo Levenberg-Marquadt para resolução de problemas de mínimos quadrados não lineares. No Quadro 1.2 estão apresentados os modelos selecionados para ajuste na forma original e modificada.

QUADRO 1.2 - MODELOS NÃO LINEARES SELECIONADOS PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE *Pinus taeda*.

Modelo	Forma	Expressão
Gompertz	Orig.	$Ht = \beta_0 e^{-\beta_1 e^{-\beta_2 DAP}} + \varepsilon$
	Modif.	$Ht = (\beta_{00} + \beta_{01} Hdom) * e^{-(\beta_{10} + \beta_{11} id) * e^{-(\beta_{20} + \beta_{21} Hdom) DAP}} + \varepsilon$
Monomolecular	Orig.	$Ht = \beta_0 (1 - e^{\beta_1 DAP}) + \varepsilon_i$
	Modif.	$Ht = (\beta_{01} Hdom + \beta_{02} id) * (1 - e^{(\beta_{10} + \beta_{11} Hdom) DAP}) + \varepsilon$

FONTE: O autor (2019).

1.2.3 ALGORITMOS DE APRENDIZADO DE MÁQUINA

Foram testados quatro algoritmos de aprendizado de máquina: k-Vizinhos Mais Próximos (k-VMP), Floresta aleatória (FA), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA), cujos treinamentos foram realizados no *software* R Studio 1.0.136. No treinamento os dados foram normalizados entre 0 e 1.

As variáveis de entrada utilizadas em todos os algoritmos foram DAP, idade e altura dominante.

1.2.3.1 k-Vizinhos Mais Próximos (k-VMP)

O treinamento de k-VMP realizado para o presente estudo considerou a ponderação das distâncias entre as instâncias por meio de funções *kernel*. As funções *kernel*, quando aplicadas a uma distância calculada por uma métrica definida, estabelecem o peso de cada vizinho na determinação do valor de uma nova instância. O método foi descrito com maiores detalhes por Hechenbichler e Schliep (2004).

O número de vizinhos mais próximos (k) testados variou de 1 a 50. Também foram testadas duas métricas de distância: euclidiana e de Manhattan, dadas pelas expressões apresentadas no Quadro 1.3.

QUADRO 1.3 - MÉTRICAS DE DISTÂNCIA TESTADAS PARA O ALGORITMO K-VMP.

Métrica de distância	Expressão matemática
Euclidiana	$d(x_i, x_j) = \sqrt{\sum_{s=1}^n (x_{is} - x_{js})^2}$
Manhattan	$d(x_i, x_j) = \sum_{s=1}^n x_{is} - x_{js} $

FONTE: O autor (2019).

NOTA: d = distância entre as instâncias x_i ($x_{i1}, x_{i2} \dots x_{in}$) e x_j ($x_{j1}, x_{j2} \dots x_{jn}$); x_i e x_j = instâncias diferentes com n variáveis explicativas.

Foram testadas também diversas funções *kernel* para ponderação das distâncias, apresentadas no Quadro 1.4.

QUADRO 1.4 - FUNÇÕES KERNEL TESTADAS PARA K-VIZINHOS MAIS PRÓXIMOS.

Função kernel	Expressão matemática
Retangular	$K = \frac{1}{2}$
Triangular	$K(d) = (1 - d)$
Epanechnikov	$K(d) = \frac{3}{4}(1 - d^2)$
Quadrática	$K(d) = \frac{15}{16}(1 - d^2)^2$
Cúbica	$K(d) = \frac{35}{32}(1 - d^2)^3$
Cosseno	$K(d) = \frac{\pi}{4} \cos\left(\frac{\pi}{2}d\right)$
Gauss	$K(d) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d^2}$
Inversa	$K(d) = \frac{1}{ d }$

FONTE: O autor (2019).

1.2.3.2 Floresta Aleatória (FA)

O treinamento do algoritmo FA foi realizado por meio do método *bagging*, que consiste na utilização de subconjuntos de dados da base de treinamento para geração das árvores. O número de árvores de regressão (*Ntree*) e o número de variáveis selecionadas em cada nó (*Mtry*) foram os parâmetros testados no treinamento dos modelos de FA. Testaram-se valores de 10 a 500 para o parâmetro *Ntree*, com saltos de 10 unidades a cada treinamento, enquanto para *Mtry* foram testados os valores 2 e 3. Determinou-se a quantidade mínima de cinco folhas terminais por árvore (valor padrão do pacote R *randomForest*) e sem valor máximo definido.

1.2.3.3 Máquinas de Vetores de Suporte (MVS)

Foi configurado a MVS com função *kernel* de base radial, a qual possui o parâmetro γ (*gamma*). Para o treinamento das MVS foi utilizada a função de erro do tipo II, denominada ν -regression (SMOLA, 1996), a qual possui o parâmetro ν utilizado para determinar a proporção entre o número de vetores de suporte e o número total de instâncias na base de treinamento. No treinamento utilizou-se $\nu = 0,5$, valor padrão do pacote R *e1071*.

A busca pelos parâmetros ótimos para treinamento do algoritmo MVS foi realizada conforme o método sugerido por Hsu *et al.* (2003). A primeira busca, denominada exploratória, consistiu na utilização de uma sequência de valores espaçados para os parâmetros C e γ , sendo de 2^1 a 2^8 em intervalos de 2^1 para o parâmetro C, e de 2^{-8} a 2^{-1} para γ . Após a identificação de uma região promissora onde os erros tendiam a valores mais baixos, foi realizada a segunda busca, denominada refinada, utilizando sequências de valores com intervalos menores para cada parâmetro, em que, considerando que o valor para qualquer um dos parâmetros a serem otimizados seja 2^x , a busca refinada testará valores de $2^{x-0,5}$ a $2^{x+0,5}$ em intervalos de $2^{0,2}$.

1.2.3.4 Redes Neurais Artificiais (RNA)

Foram configuradas Redes Neurais Artificiais multicamada ântero-alimentadas (*feedforward*), com algoritmo de treinamento de retropropagação dos erros, o qual utiliza o método do gradiente descendente estocástico para atualização

dos pesos sinápticos (CANDEL *et al.*, 2017). O número de neurônios testados na camada intermediária foi de 1 a 15 e o limite de ciclos de treinamento foi 3.000. O treinamento foi automaticamente interrompido pelo algoritmo quando a precisão do modelo não foi incrementada por 20 ciclos. A função de ativação utilizada na camada de saída foi a linear, ao passo que na camada oculta, foram testadas as funções tangente hiperbólica, linear retificada e *maxout*, cujas descrições matemáticas são apresentadas no Quadro 1.5.

QUADRO 1.5 - FUNÇÕES DE ATIVAÇÃO TESTADAS NO TREINAMENTO DE REDES NEURAIS ARTIFICIAIS.

Função de ativação	Expressão
Tangente hiperbólica	$f(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}$
Unidade linear retificada (<i>ReLU</i>)	$f(\alpha) = \max(0, \alpha)$
<i>Maxout</i>	$f(\alpha) = \max(\theta_1, \theta_2)$

FONTE: O autor (2019).

O parâmetro taxa de aprendizado, que determina a taxa de atualização dos pesos sinápticos a cada nova iteração, e o termo momentum, que auxilia na tarefa de evitar mínimos locais e instabilidades relacionadas, não foram testados neste estudo, uma vez que o pacote *h2o* permite a utilização de aprendizado adaptativo, o qual combina o método de arrefecimento da taxa de aprendizado (*rate annealing*) e de treinamento do termo momentum, as quais adaptam o valor desses parâmetros durante o treinamento (CANDEL *et al.*, 2017).

1.2.4 OTIMIZAÇÃO DOS HIPERPARÂMETROS DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA

Para obtenção dos hiperparâmetros que maximizam o desempenho dos algoritmos de AM, empregou-se o método de busca em grade. Este método requer a definição de um conjunto ou sequência de valores para cada parâmetro, para que a busca seja realizada por meio de ajustes sucessivos, utilizando todas as combinações de parâmetros possíveis (BERGSTRA e BENGIO, 2012).

Para evitar a seleção de modelos sobreajustados aos dados de treinamento, a avaliação dos algoritmos de AM sob diferentes configurações de hiperparâmetros foi realizada por meio da validação cruzada *k-fold* combinada ao método de busca em

grade. Segundo Shalev-Shwartz e Ben-David (2014), o método *k-fold* consiste em particionar a base de dados aleatoriamente em *k* subconjuntos de dimensões similares e, para cada configuração do algoritmo, realizar *k* treinamentos, separando um subconjunto por vez para validação.

Desse modo, todas as instâncias são utilizadas *k-1* vezes para treinamento e uma vez para validação. A medida de desempenho de cada configuração é obtida pela média do erro quadrático médio das *k* validações. Após a etapa de validação cruzada, o algoritmo efetua o treinamento de um modelo final para a configuração testada, utilizando toda a base de dados de treinamento.

Para cada algoritmo de AM, a melhor configuração foi a que conferiu ao modelo o menor erro quadrático médio obtido por validação cruzada. Neste estudo, utilizou-se o método *3-fold*. Adicionalmente, o tempo de busca para os parâmetros por meio do método *3-fold* foi computado para cada algoritmo.

1.2.5 AVALIAÇÃO DO DESEMPENHO DOS MODELOS

A pré-avaliação dos modelos de regressão ajustados foi realizada pelo cálculo dos indicadores usualmente utilizados para comparação de modelos estatísticos: o coeficiente de determinação ajustado (R^2_{aj}) e o erro padrão da estimativa ($S_{yx\%}$), são calculados pelas expressões a seguir:

$$erro_i = Ht_{i_{obs}} - Ht_{i_{est}}$$

$$R^2 = 1 - \frac{\sum(erro_i)^2}{\sum(Y_i - \bar{Y})^2} \quad R^2_{aj} = 1 - [(1 - R^2) * \left(\frac{n - 1}{n - p - 1}\right)]$$

$$S_{yx} = \sqrt{\frac{\sum(erro_i)}{n - p}} \quad S_{yx\%} = \frac{S_{yx}}{\bar{Y}} * 100$$

Em que:

\bar{Y} = média dos valores observados;

n = número de observações;

p = número de parâmetros do modelo.

Adicionalmente, a significância dos parâmetros ajustados para os modelos de regressão estatística foi obtida pelo teste *t*. Para tanto, avaliou-se a premissa de homoscedasticidade dos resíduos pelo teste finBreusch-Pagan (BREUSCH e PAGAN, 1979) ao nível de 5% de significância.

Os modelos de AM e estatísticos foram avaliados conjuntamente quanto à performance de ajuste aos dados de treinamento e teste. A qualidade dos ajustes foi avaliada por análises gráficas de resíduos e pelos indicadores erro quadrático médio percentual (EQM%) e erro absoluto médio percentual (EAM%), representados matematicamente pelas seguintes fórmulas:

$$EQM\% = \frac{\sqrt{\sum_{i=1}^n erro_i^2}}{n} * \frac{100}{\overline{Ht}}$$

$$EAM\% = \frac{\sum_{i=1}^n |erro_i|}{n} * \frac{100}{\overline{Ht}}$$

Sendo:

Ht_{i_obs} = altura observada para a árvore *i*;

Ht_{i_est} = altura estimada para a árvore *i*;

\overline{Ht} = altura média;

n = número de observações.

Todas as técnicas foram submetidas à análise gráfica de resíduos sobre a linha de ajuste e sobre as variáveis preditoras.

1.3 RESULTADOS E DISCUSSÃO

1.3.1 AJUSTE DOS MODELOS DE REGRESSÃO

Na Tabela 1.1 estão apresentados os coeficientes dos modelos de regressão ajustados. Por meios dos indicadores estatísticos, observa-se que todos os modelos foram apropriados para estimativa da altura das árvores. Os modelos não lineares modificados apresentaram estatísticas de ajuste melhores que o modelos lineares, já

que resultaram em menor erro padrão das estimativas ($S_{yx\%}$) e maior coeficiente de determinação ajustado (R^2_{aj}).

TABELA 1.1 - COEFICIENTES DE REGRESSÃO E INDICADORES ESTATÍSTICOS DOS MODELOS HIPSONÔMTRICOS AJUSTADOS.

Base de ajuste	Parâmetro	Estimativa	Erro padrão	t_c	Valor-p	$S_{yx\%}$	R^2_{aj}
Linear 1	β_0	-7,2897	1,7260	-4,223	< 0,001	9,02	0,863
	β_1	0,7526	0,0534	14,091	< 0,001		
	β_2	0,4479	0,0206	21,7945	< 0,001		
	β_3	0,8579	0,3867	2,218	0,027		
	β_4	-0,0101	0,0014	-7,014	< 0,001		
	β_5	-0,0384	0,0200	-1,917	0,046		
Linear 2	β_0	1,0213	0,1383	7,384	< 0,001	9,35	0,859
	β_1	-6,0671	0,1687	-35,969	< 0,001		
	β_2	0,5755	0,0241	23,916	< 0,001		
	β_3	0,0810	0,0312	2,589	< 0,001		
	β_4	-0,0039	0,0016	-2,391	0,017		
Gompertz	β_{00}	1,4869	0,6379	2,331	0,020	8,31	0,884
	β_{01}	1,0082	0,0413	24,399	< 0,001		
	β_{10}	2,2944	0,2542	9,025	0,027		
	β_{12}	-0,0512	0,0178	-2,875	0,004		
	β_{20}	0,2391	0,0215	11,118	< 0,001		
	β_{21}	-0,0067	0,0009	-7,679	< 0,001		
Monomolecular	β_{01}	1,0704	0,02359	45,371	< 0,001	8,27	0,875
	β_{02}	0,1572	0,0370	4,251	< 0,001		
	β_{10}	-0,1648	0,0073	-22,630	< 0,001		
	β_{11}	0,0047	0,0003	15,459	< 0,001		

FONTE: O autor (2019).

O teste Breusch-Pagan foi significativo para os dois modelos lineares, indicando heterogeneidade de variâncias ao longo da linha de ajuste. Os valores-p foram < 0,001, 0,0395, 0,0821 e 0,0808 para os modelos Linear 1, Linear 2, Gompertz e Monomolecular, respectivamente. A heterocedasticidade dos modelos lineares invalida as estatísticas de teste-t para avaliação da significância dos parâmetros, uma vez que os erros-padrão dos coeficientes são viesados. Isto pode ser corrigido pelo emprego da teoria dos mínimos quadrados generalizados (GUJARATI, 2006), o que no entanto, não foi objeto de estudo na presente pesquisa.

O atendimento à premissa de homogeneidade dos resíduos por parte dos modelos não lineares valida o teste de significância das covariantes adicionadas. Em ambos modelos pelo menos uma das variáveis *idade* e *Hdom* demonstrou ser determinante na obtenção de cada parâmetro.

Segundo Sena *et al.* (2015), modelos não lineares são mais adequados para explicação da relação hipsométrica de povoamentos florestais que modelos lineares, devido à interpretação biológica dos parâmetros e por possuírem maior capacidade de representar a relação não linear entre altura e diâmetro. O modelo não linear que apresentou o menor erro padrão (Monomolecular) apresentou um erro 8,3% menor que o modelo Linear 1, que apresentou o menor $S_{yx\%}$ dentre os lineares, destacando a maior capacidade dos modelos não lineares modificados em explicar a alta variabilidade dos dados.

1.3.2 BUSCA EM GRADE

1.3.2.1 K-vizinhos Mais Próximos

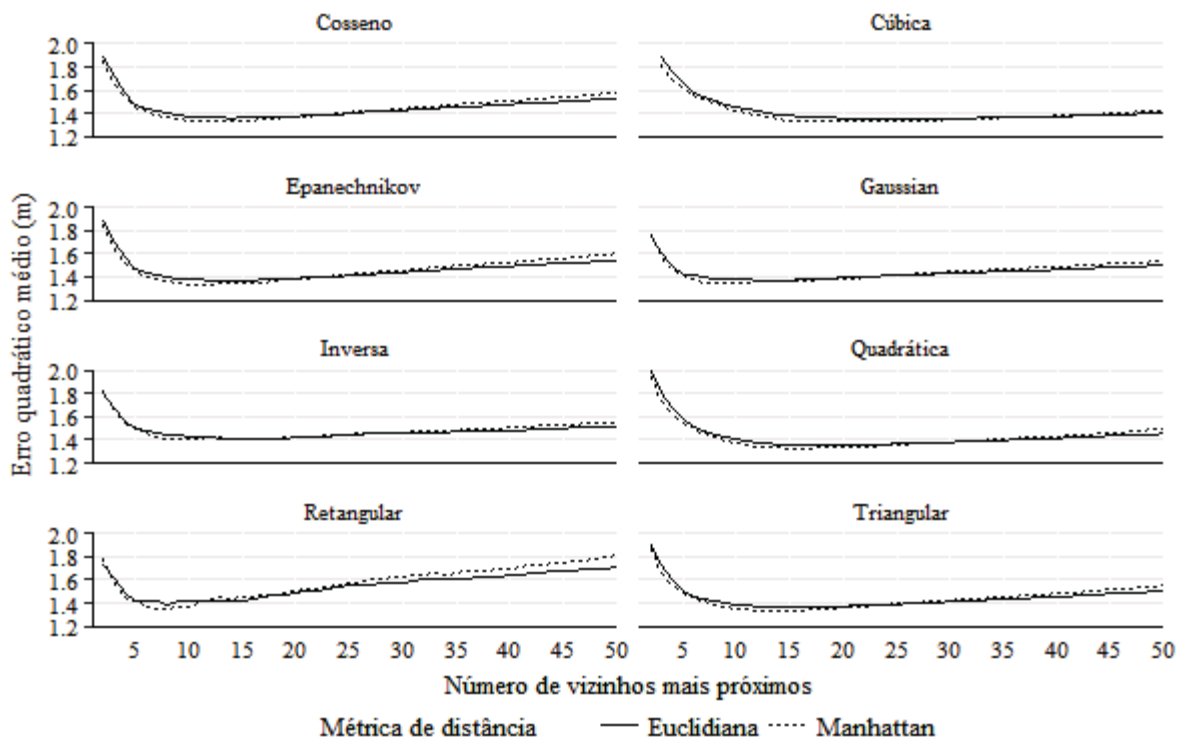
A melhor configuração obtida para k-VMP foi $k = 16$, função de ponderação quadrática e métrica de distância de Manhattan. A métrica de distância e a função de ponderação tiveram pouca influência na performance dos resultados, sendo o número de vizinhos mais próximos o parâmetro determinante para qualidade das estimativas (Figura 1.2). A função quadrática combinada com a métrica de distância de Manhattan foi a que minimizou o erro quadrático; no entanto, a diferença entre o MSE obtido pelo modelo de melhor desempenho e o modelo que ocupou a 800ª posição foi inferior a 1%.

A função inversa de ponderação é comumente utilizada nos estudos aplicados à mensuração florestal (HAARA *et al.*, 1997; TOMMOLA *et al.*, 1999; SANQUETTA *et al.*, 2013). Contudo, no presente estudo, o melhor modelo gerado pela função inversa obteve desempenho inferior aos melhores modelos das demais funções, porém com diferença máxima de 0,23%. Esse resultado corrobora a afirmação de Hechenbichler e Schliep (2004), que a escolha da função *kernel* não é crucial e que maiores diferenças são obtidas somente quando é escolhida a função retangular, a qual aplica o mesmo peso para todos os vizinhos, independente da distância.

Sanquetta *et al.* (2013) ao analisarem a performance de k-VMP para estimar o estoque de carbono em indivíduos de *Araucaria angustifolia* (Bertol.) Kuntze, observaram similaridade entre os resultados obtidos por três configurações distintas, em que variaram o número de vizinhos (1 ou 3) e a métrica de distância (euclidiana e euclidiana ao quadrado). No presente estudo, também foi observada elevada

similaridade entre diversos modelos gerados com diferentes configurações, no entanto, os modelos treinados com $k \leq 5$ e com elevado número de vizinhos apresentaram precisão inferior aos demais, independente da métrica de distância e da função de ponderação escolhida.

FIGURA 1.2 - ERRO QUADRÁTICO MÉDIO EM FUNÇÃO DO NÚMERO DE VIZINHOS MAIS PRÓXIMOS NO TREINAMENTO DE K-VIZINHOS MAIS PRÓXIMOS COM DIFERENTES FUNÇÕES *KERNEL* E MÉTRICAS DE DISTÂNCIA PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE POVOAMENTOS DE *Pinus taeda* COM IDADES VARIANDO 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

De modo geral, o valor ótimo para k variou entre 10 e 20, considerando as diferentes funções *kernel* e as métricas de distância. De modo similar ao observado neste estudo, Tommola *et al.* (1999) afirmaram que um baixo valor para k resulta em sobreajuste aos dados de treinamento, enquanto um elevado valor provoca um subajuste, quando não há ponderação pela distância, originando curvas mais suaves e fazendo com que as estimativas tendam aos valores próximos a média do conjunto de dados de ajuste. Esses autores observaram aumento no percentual de erro na estimativa da razão entre sortimentos de um povoamento florestal na Finlândia quando k foi superior a 20, tal como ocorrido no presente estudo.

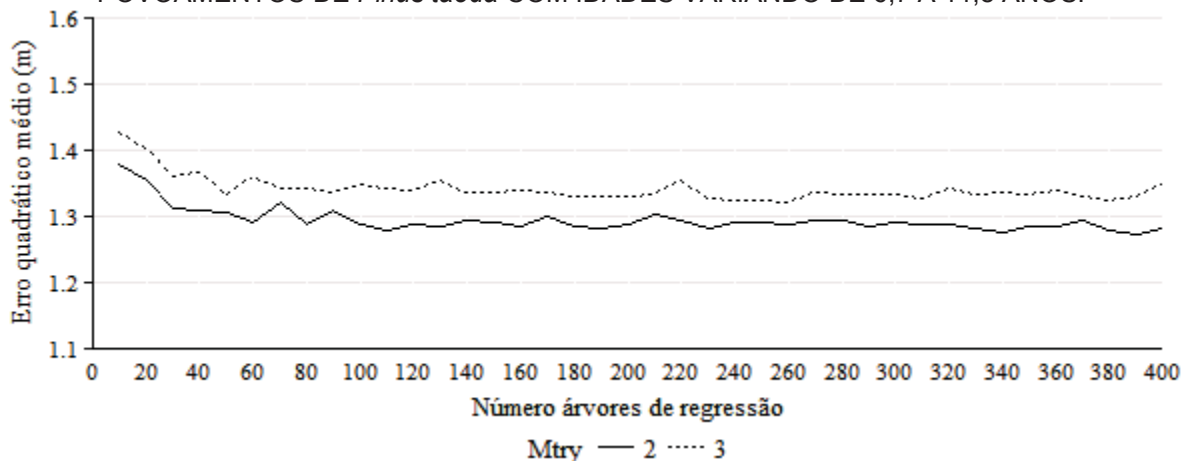
1.3.2.2 Floresta Aleatória

Para o algoritmo FA, os melhores valores para os parâmetros $Mtry$ e $Ntree$ foram 390 e 2, respectivamente. Os treinamentos realizados utilizando duas variáveis por nó foram superiores, de modo geral, aos realizados utilizando três. O valor recomendado por Breiman (2001) para o parâmetro $Mtry$ é $p/3$, em que p é o número de variáveis de entrada. Esse valor não deve ser considerado como regra, uma vez que no presente estudo p foi igual a 3 e os melhores resultados foram obtidos com $Mtry = 2$.

Não foram observadas tendências ou regiões ótimas para o parâmetro n , uma vez que a diferença entre o melhor modelo ($Ntree = 390$ e $Mtry = 2$) e o pior ($Ntree = 10$ e $Mtry = 3$) foi de 0,5%. No entanto, foi observado que os modelos treinados com $Ntree \leq 100$ foram mais instáveis e frequentemente apresentaram maiores valores de erro quadrático médio, indicando que os treinamentos realizados com poucas árvores tendem a apresentar performance de generalização inferior, embora um elevado número de árvores não necessariamente implique em melhores resultados (Figura 1.3).

Segundo Hastie (2003), o algoritmo FA estabiliza a partir de 200 árvores e o valor ótimo de $Mtry$ depende da tarefa em questão, devendo ser tratado como um parâmetro a ser ajustado. No presente estudo, a estabilização do erro foi observada a partir de aproximadamente 100 árvores.

FIGURA 1.3 – ERRO QUADRÁTICO MÉDIO EM FUNÇÃO DO NÚMERO DE ÁRVORES DE REGRESSÃO ($NTREE$) PARA DIFERENTES VALORES DE $MTRY$ NO TREINAMENTO DE FLORESTAS ALEATÓRIAS PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE POVOAMENTOS DE *Pinus taeda* COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.

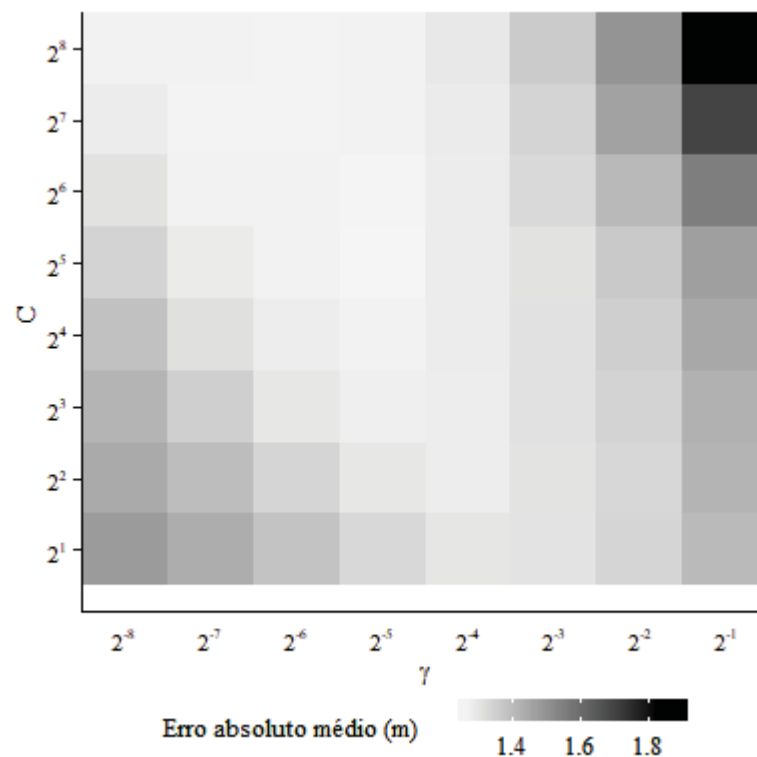


FONTE: O autor (2019).

1.3.2.3 Máquinas de Vetores de Suporte

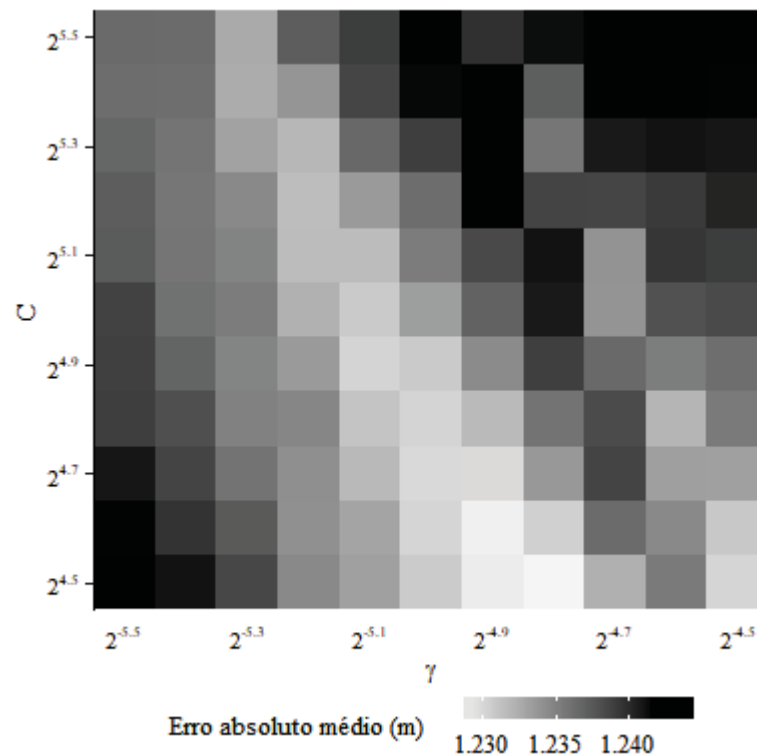
A busca exploratória por uma região promissora para a configuração do algoritmo MVS obteve como melhor resultado $C = 2^5$ e $\gamma = 2^{-5}$ (Figura 1.4). A partir desse resultado, foi realizada uma busca mais restrita, onde o espaço de possibilidades compreendeu aos valores entre $2^{4.5}$ e $2^{5.5}$ para C e entre $2^{-5.5}$ e $2^{-4.5}$ para γ , sendo encontrados os valores $2^{4.5}$ e $2^{-4.8}$ como os ótimos, respectivamente (Figura 1.5).

FIGURA 1.4 – RESULTADO DA BUSCA EXPLORATÓRIA PARA O ERRO ABSOLUTO MÉDIO EM FUNÇÃO DOS PARÂMETROS C E γ NO TREINAMENTO DE MÁQUINAS DE VETORES DE SUORTE APLICADAS À MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE POVOAMENTOS DE *Pinus taeda* COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

FIGURA 1.5 - RESULTADO DA BUSCA RESTRITA PARA O ERRO ABSOLUTO MÉDIO EM FUNÇÃO DOS PARÂMETROS C E γ NO TREINAMENTO DE MÁQUINAS DE VETORES DE SUORTE APLICADAS MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE POVOAMENTOS DE *Pinus taeda* COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

1.3.2.4 Redes Neurais Artificiais

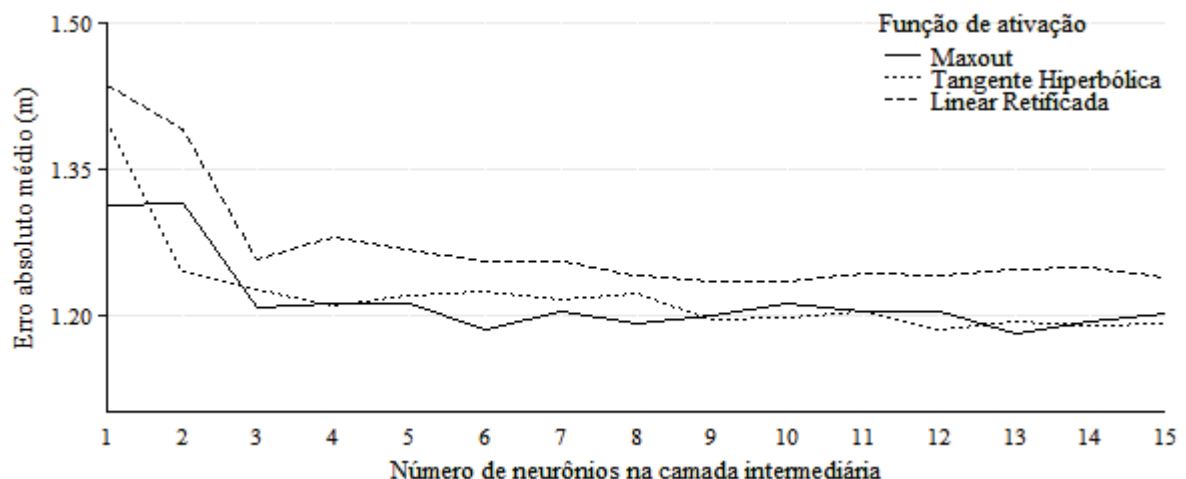
A rede que apresentou os melhores resultados foi treinada utilizando 13 neurônios na camada oculta e função de ativação *maxout*. As funções *maxout* e tangente hiperbólica apresentaram resultados similares e foram superiores à função linear retificada em todos os treinamentos. No entanto, destaca-se que a diferença de desempenho observada entre as funções de ativação foi pouco significativa.

As redes treinadas com apenas 1 neurônio apresentaram os piores resultados em comparação às redes treinadas com mais neurônios e, de modo geral, maior precisão foi observada em redes treinadas com pelo menos 3 neurônios (Figura 1.6).

No estudo de Özçelik *et al.* (2013), que testou RNA para a estimativa de altura de árvores de zimbro na Criméia empregando a função de ativação sigmoidal combinada às diferentes configurações de parâmetros, os resultados mais satisfatórios foram obtidos por redes treinadas com um ou dois neurônios na camada

oculta. Por outro lado, Vendruscolo *et al.* (2016), utilizando a mesma função de ativação para estimativa de alturas individuais em povoamentos de *Tectona grandis* L., observaram melhor performance de generalização em redes treinadas com seis a dez neurônios na camada oculta. Outros autores ao testarem diferentes configurações de RNA obtiveram resultados que demonstram a ausência de uma configuração mais indicada para sua utilização na modelagem da relação hipsométrica, ou mesmo de outros atributos florestais (BINOTI *et al.*, 2014; VENDRUSCOLO, 2015; MARTINS *et al.*, 2016) e que, desse modo, diferentes parâmetros devem ser testados a fim de se obter a configuração mais adequada para cada caso.

FIGURA 1.6 – ERRO ABSOLUTO MÉDIO EM FUNÇÃO DO NÚMERO DE NEURÔNIOS NA CAMADA INTERMEDIÁRIA DE REDES NEURAIS ARTIFICIAIS TREINADAS COM DIFERENTES FUNÇÕES DE ATIVAÇÃO PARA MODELAGEM DA RELAÇÃO HIPSOMÉTRICA DE POVOAMENTOS DE *Pinus taeda* COM IDADES VARIANDO 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

1.3.3 AVALIAÇÃO DOS MODELOS

Na Tabela 1.2 estão apresentados os indicadores de avaliação dos modelos ajustados para os conjuntos de treinamento e teste. Os métodos MVS e RNA e os modelos não lineares modificados foram os que apresentaram indicadores mais favoráveis à utilização, embora todos os métodos testados tenham apresentado resultados satisfatórios.

Os métodos k-VMP e FA apresentaram os menores EAM% e EQM% para o grupo de ajuste/treinamento, no entanto estes métodos foram inferiores à maioria dos demais métodos no grupo de teste. Os métodos RNA, MVS e modelo Monomolecular

apresentaram os menores valores de EAM% e EQM% no grupo de teste, destacando-se a RNA com o menor erro médio e MVS com menor variância dos erros.

Apesar de terem apresentado indicadores estatísticos similares aos modelos de AM na etapa de teste, os modelos estatísticos apresentaram desempenho inferior aos modelos de AM na dispersão de resíduos, com tendências à subestimativa da altura das árvores mais altas e mais baixas (Figura 1.7). Os modelos lineares apresentaram elevada tendenciosidade em comparação aos demais métodos, enquanto a tendência nos modelos não lineares foi pouco expressiva. Ademais, todos os algoritmos de AM, especialmente a RNA, apresentaram dispersão de resíduos mais homogênea em comparação aos modelos de regressão estatística.

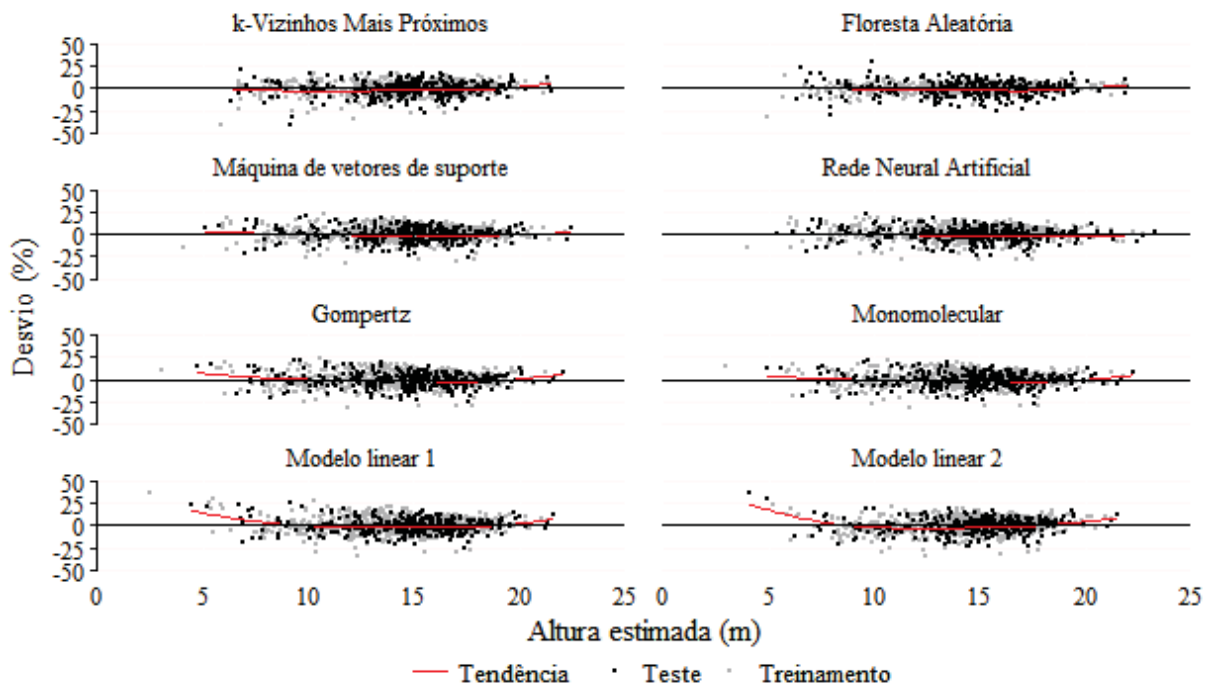
TABELA 1.2 - INDICADORES DE DESEMPENHO DOS MODELOS ESTATÍSTICOS E DE APRENDIZADO DE MÁQUINA NA ESTIMATIVA DA ALTURA TOTAL DE ÁRVORES DE *Pinus taeda* COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.

Modelo	Ajuste/Treinamento		Teste	
	EAM%	EQM%	EAM%	EQM%
RNA	5,53	7,17	6,10	7,89
k-VMP	5,02	6,61	6,30	8,22
FA	3,05	3,96	6,39	8,08
MVS	5,63	7,36	6,12	7,85
Gompertz	5,75	7,52	6,21	7,95
Monomolecular	5,75	7,50	6,19	7,90
Linear 1	6,04	7,83	6,35	7,95
Linear 2	6,13	7,98	6,48	8,19

FONTE: O autor (2019).

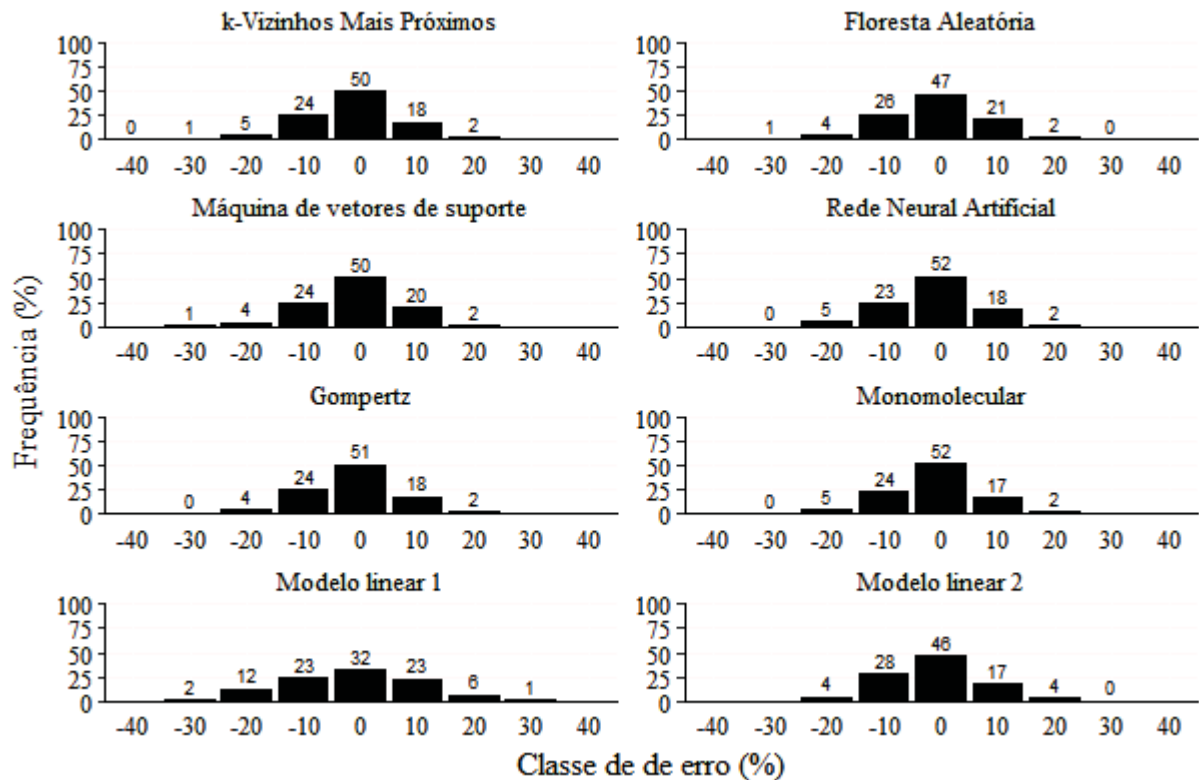
Todas as técnicas testadas quando utilizadas para estimar as alturas na base de teste produziram resíduos com distribuição simétrica e com maior frequência na classe zero (Figura 1.8). Os modelos lineares e o método Floresta Aleatória apresentaram menor concentração de resíduos em torno de zero em comparação aos demais métodos.

FIGURA 1.7 - DISPERSÃO DOS RESÍDUOS PARA A ESTIMATIVA DA ALTURA TOTAL DE ÁRVORES DE *Pinus taeda* COM IDADES DE 6,7 A 11,8 ANOS, OBTIDOS POR MODELOS LINEARES, NÃO LINEARES E DE APRENDIZADO DE MÁQUINA.



FONTE: O autor (2019).

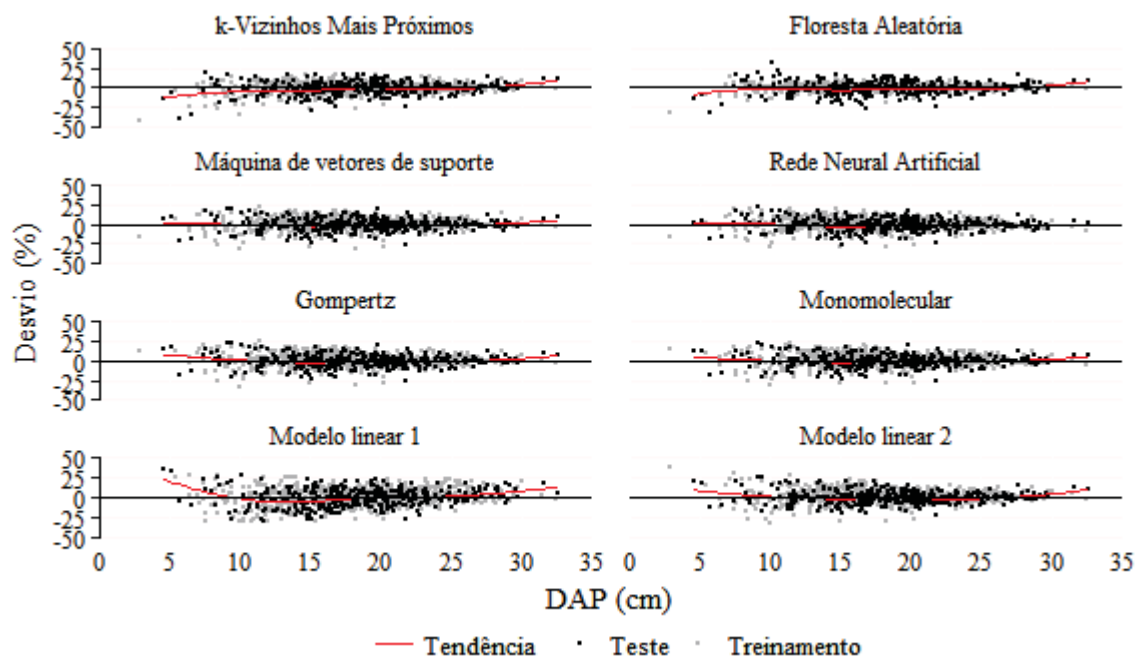
FIGURA 1.8 - FREQUÊNCIA DE RESÍDUOS RELATIVOS DA ALTURA ESTIMADA RESULTANTES DA APLICAÇÃO DOS MÉTODOS NO CONJUNTO DE TESTE, EM POVOAMENTOS DE *Pinus taeda* COM IDADES DE 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

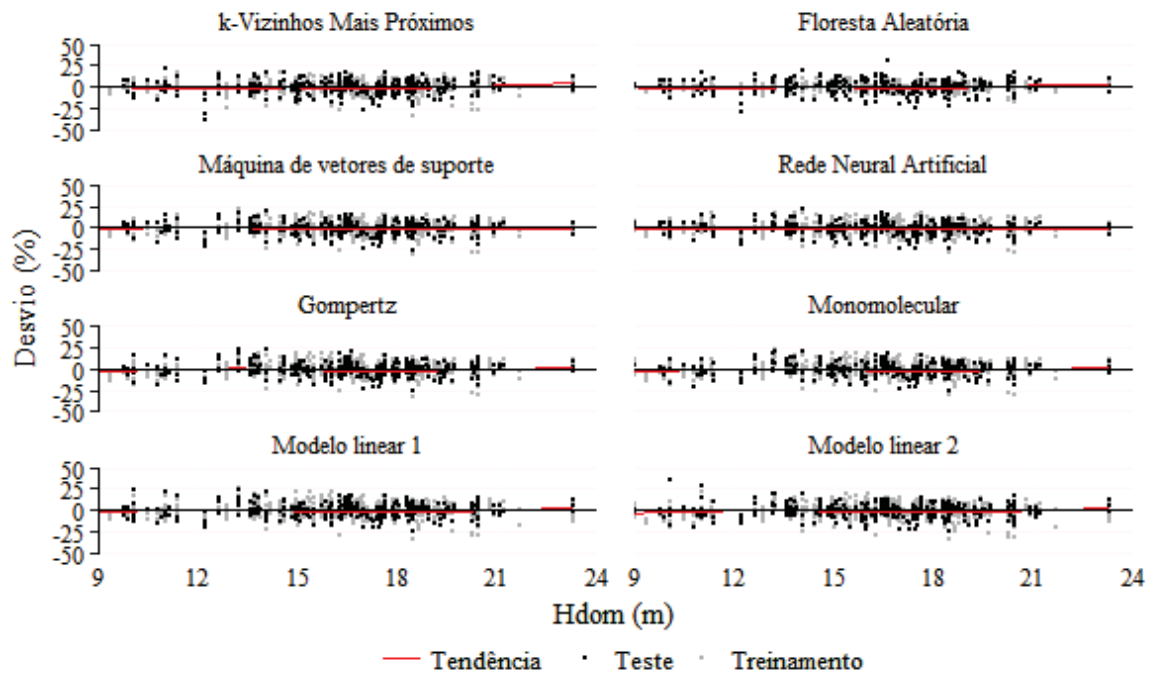
Os métodos RNA e MVS, nesta ordem, apresentaram a dispersão dos resíduos sobre o diâmetro à altura do peito mais satisfatória. Os métodos k-Vizinhos mais próximos e Florestas aleatória apresentaram tendências a superestimar a altura de árvores de pequeno diâmetro e subestimar as de maior diâmetro, enquanto os modelos não lineares e lineares subestimaram ambos os extremos (Figura 1.9). Os resíduos plotados sobre as variáveis altura dominante (Figura 1.10) e idade (Figura 1.11) foram satisfatórios para todos os métodos. A RNA e a MVS foram os métodos que apresentaram maior robustez, uma vez que a dispersão dos resíduos para estes foi menos sensível aos desvios das variáveis DAP e Hdom, especialmente nas extremidades da linha de ajuste.

FIGURA 1.9 - DISPERSÃO DOS RESÍDUOS DA ALTURA ESTIMADA POR MODELOS LINEARES, NÃO LINEARES E DE APRENDIZADO DE MÁQUINA EM RELAÇÃO À VARIÁVEL DIÂMETRO À ALTURA DO PEITO DE ÁRVORES DE *PINUS TAEDA* COM IDADES DE 6,7 A 11,8 ANOS.



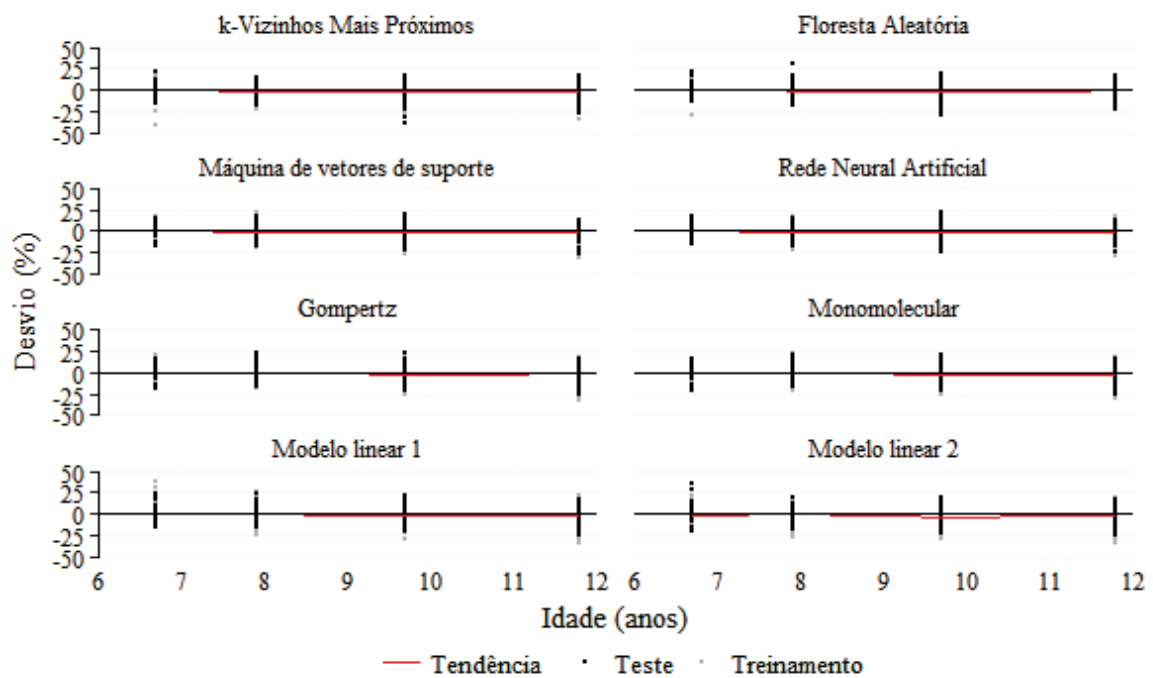
FONTE: O autor (2019).

FIGURA 1.10 - DISPERSÃO DOS RESÍDUOS DA ALTURA ESTIMADA POR MODELOS LINEARES, NÃO LINEARES E DE APRENDIZADO DE MÁQUINA EM RELAÇÃO À VARIÁVEL ALTURA DOMINANTE DE ÁRVORES DE *Pinus taeda* COM IDADES DE 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

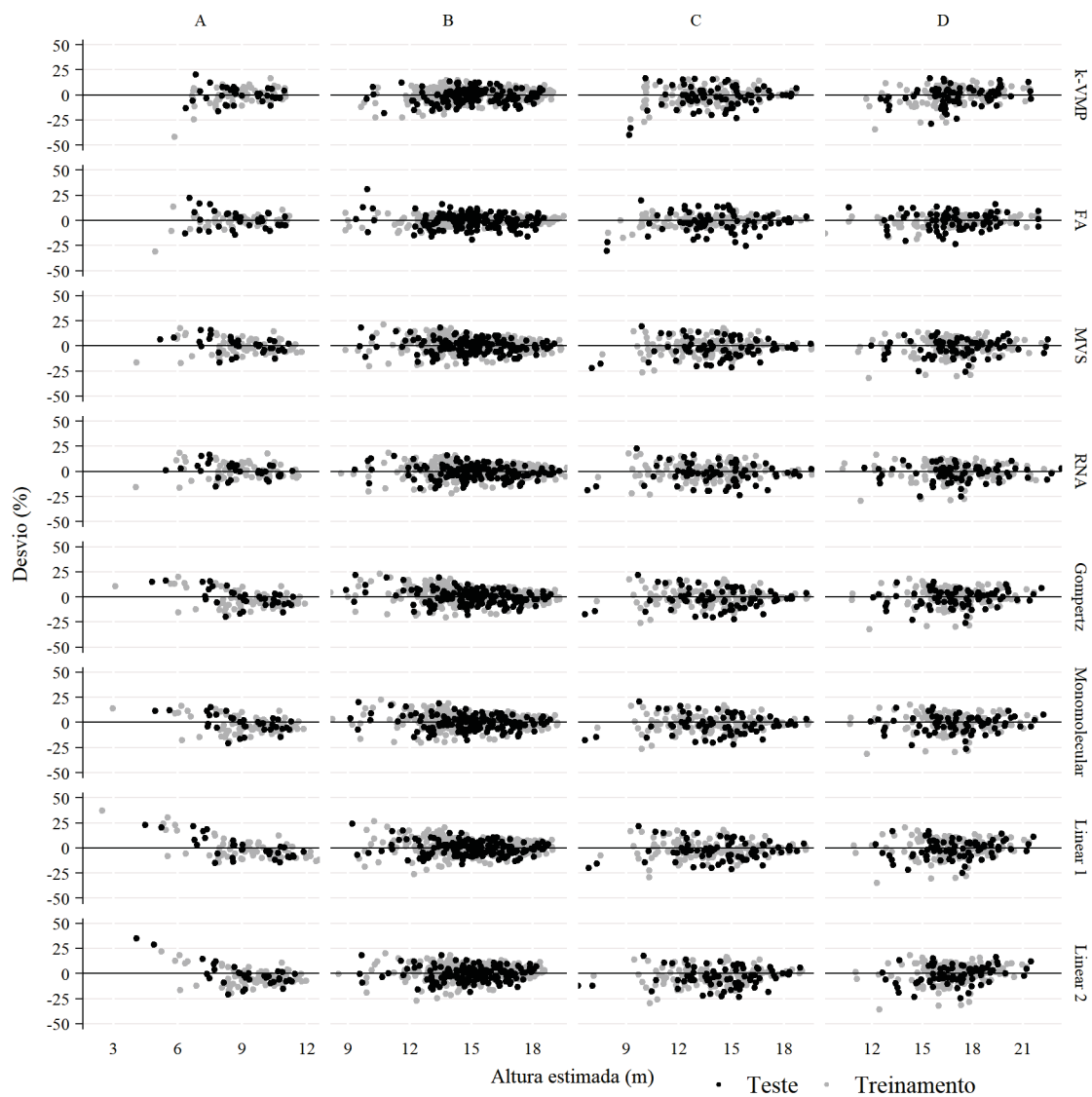
FIGURA 1.11 - DISPERSÃO DOS RESÍDUOS DA ALTURA ESTIMADA POR MODELOS LINEARES, NÃO LINEARES E DE APRENDIZADO DE MÁQUINA EM RELAÇÃO À VARIÁVEL IDADE DE ÁRVORES DE *Pinus taeda* COM IDADES VARIANDO DE 6,7 A 11,8 ANOS.



FONTE: O autor (2019).

Ao considerar a análise de resíduos sobre a linha de ajuste por povoamento, observa-se que para o povoamento A os modelos lineares e não lineares apresentaram elevado viés (Figura 1.12). Isso pode ser resultado da baixa quantidade de observações em comparação aos demais povoamentos. Neste caso, a performance superior dos métodos de AM pode sugerir que estes apresentam melhor desempenho na regressão de dados desbalanceados.

FIGURA 1.12 - DISPERSÃO DOS RESÍDUOS RELATIVOS ESTIMADOS PELOS ALGORITMOS DE APRENDIZADO DE MÁQUINA E MODELOS DE REGRESSÃO LINEAR E NÃO LINEAR COM ESTRATIFICAÇÃO DE POVOAMENTOS DE *Pinus taeda* POR IDADE: A (6,7 ANOS), B (7,9 ANOS), C (9,7 ANOS) E D (11,8 ANOS).



No povoamento D os modelos lineares e o algoritmo k-VMP apresentaram tendenciosidade. Para o povoamento B os modelos não lineares apresentaram maior

tendenciosidade na estimativa da altura das árvores de menor diâmetro e para o C todos os métodos apresentaram algum nível de tendenciosidade na estimativa das menores árvores, sendo que as técnicas k-VMP e FA apresentaram os erros mais expressivos.

Os modelos de AM podem apresentar vantagens quanto à dispersão dos resíduos em bases de dados heterogêneas e desbalanceadas, em comparação a modelos de regressão lineares e não lineares.

Dentre os modelos de inteligência artificial empregados neste estudo, as MVS e RNA merecem destaque por terem apresentado os menores valores de EAM% e EQM%, além distribuição de resíduos mais satisfatória na etapa de predição do conjunto de teste. Os indicadores estatísticos e análise de resíduos foram mais favoráveis ao uso da RNA.

Isso também foi observado por Cordeiro *et al.* (2015), os quais concluíram que a RNA treinada superou a MVS e o modelo de Schumacher & Hall na estimativa do volume de *Acacia mangium* Willd. A conformidade entre resultados não é suficiente para atestar superioridade absoluta da RNA sobre a MVS para fins florestais, mas indica que ambos métodos tendem a apresentar elevada performance de ajuste, superior aos métodos estatísticos de regressão.

Por outro lado, os modelos não lineares com adição de covariantes apresentaram resultados próximos à RNA e MVS, superando as técnicas k-VMP e FA. Isto demonstra que o uso de covariantes confere grande flexibilidade aos modelos não lineares para explicar relações não lineares em conjuntos de dados de grande variabilidade. Segundo Calegario *et al.* (2005), o método de decomposição dos parâmetros possui grande flexibilidade devido ao fato de que as covariantes podem estar associadas à um parâmetro e não a outro, dependendo de sua significância.

1.3.4 CONSIDERAÇÕES SOBRE A OTIMIZAÇÃO DE HIPERPARÂMETROS DE ALGORITMOS DE APRENDIZADO DE MÁQUINA

A realização da busca em grade para otimização dos parâmetros dos algoritmos de AM foi importante para seleção de modelos de elevado desempenho. A obtenção dos parâmetros que minimizam os erros das estimativas requer um esforço computacional que, em alguns casos, pode inviabilizar a utilização de alguns desses métodos quando há necessidade de resultados imediatos.

Contudo, o custo computacional em termos de tempo e de capacidade de processamento pode ser reduzido a uma escala aceitável quando a grade de busca apresenta menores dimensões (BERGSTRA e BENGIO, 2012).

O tempo necessário para realização da busca em grade para cada algoritmo, considerando a grade de parâmetros utilizada para cada técnica e empregando o método de validação cruzada *3-fold* em uma máquina Intel® Core™ i7-5500U 2.40GHz com dois núcleos e 8Gb de RAM, foi: k-VMP = 6,7 segundos; FA = 5,4 minutos; MVS = 3,4 minutos (busca exploratória) e 5,5 minutos (busca refinada); RNA = 12,8 minutos. Para o treinamento do algoritmo FA, o tempo de busca de parâmetros pode ser reduzido drasticamente, uma vez que o erro estabilizou a partir da geração de 50 árvores. A título de exemplo, o tempo de processamento da busca em grade para treinamento do algoritmo com 500 árvores, e visando buscar o valor mais adequado para *Mtry* em um espaço de busca de 2 valores (2 ou 3), foi de 17,2 segundos.

O tempo gasto para treinamento de uma MVS com $C = 1$ e $\gamma = 0.1$, empregando o conjunto de treinamento utilizado neste estudo como entrada, foi de 1,7 segundos, enquanto para o mesmo valor de γ e com $C = 100$, o tempo de treinamento foi de 7,19 segundos, o que demonstrou que, para este caso, quanto maior o valor de C maior o tempo de treinamento. O conjunto de valores utilizado para busca de C na grade de busca exploratória foi amplo o suficiente para localizar uma região promissora dentro dos limites estabelecidos.

Para o treinamento de RNA, o teste de diferentes funções de ativação permitiu identificar que a precisão obtida pelas redes treinadas com função linear retificada foi inferior às treinadas com as outras funções. Candel *et al.* (2017) afirmam que é difícil determinar a melhor função de ativação a ser usada, e que uma pode ser superior às

demais em cenários diferentes, sendo adequada a aplicação da busca em grade para seleção da melhor função em cada problema.

É importante ressaltar que a avaliação de algoritmos de aprendizado de máquina deve ser realizada em um ou mais conjuntos de teste, e que a avaliação em um conjunto de treinamento não avalia a capacidade de generalização de um algoritmo treinado, especialmente quando são empregados k-VMP e FA, uma vez que esses modelos tenderam a se sobreajustar aos dados de treinamento no presente estudo. Para efeitos práticos, o particionamento da base de dados em conjuntos de treinamento e de teste deve ser realizada apenas para comparação entre modelos, devendo, após a escolha do melhor método, ser realizado um treinamento final com a base completa (ALPAYDIN, 2010).

Ainda, a avaliação dos resíduos estratificados por povoamento demonstrou ser importante para o detalhamento dos erros obtidos quando se utiliza uma base de dados heterogênea, uma vez que tendências dentro dos povoamentos podem não ser identificadas por meio da análise dos resíduos como um todo. Esse método de avaliação é fundamental quando a base de dados é constituída por diferentes estratos de povoamentos.

1.4 CONCLUSÕES

- O treinamento de algoritmos de aprendizado de máquina deve ser tratado como uma tarefa de otimização, cujo objetivo é a definição dos parâmetros que minimizam o erro das predições em dados desconhecidos;
- A melhor configuração para o algoritmo k-Vizinhos Mais Próximos (k-VMP) pode ser obtida utilizando somente a busca pelo número ótimo de vizinhos mais próximos.
- A melhor configuração para o algoritmo Floresta Aleatória (FA) pode ser obtida buscando o valor ótimo para o parâmetro *Mtry*, com um número de árvores (*Ntree*) superior a 50.
- Redes Neurais Artificiais (RNA) apresentam elevada precisão e baixa tendenciosidade nas estimativas de altura de árvores. O treinamento destas no pacote *h2o* deve ser realizado de modo a identificar a quantidade ideal de neurônios na camada oculta.
- Assim como as RNA, as Máquinas de Vetores de Suporte com *kernel* função de base radial geram modelos de elevada precisão e baixa tendenciosidade na estimativa de alturas de povoamentos florestais, sendo importante a utilização da busca em grade para identificação dos valores para os parâmetros *C* e *gamma*, que conferem maior precisão ao modelo.
- Os modelos de Máquina de Vetores de Suporte (MVS) e Rede Neural Artificial (RNA) obtiveram performances de generalização superiores aos demais modelos quanto aos indicadores estatísticos e dispersão de resíduos relativos à estimativa da altura total.
- Modelos não lineares com inserção de covariantes apresentam desempenho superior aos modelos lineares genéricos e podem superar técnicas de aprendizado de máquina como k-Vizinhos Mais Próximos (k-VMP) e Floresta Aleatória (FA).
- As técnicas RNA, MVS e Modelos não lineares com adição de covariantes são as mais adequadas para gerar modelos hipsométricos genéricos.

1.5 REFERÊNCIAS

- ABDOLLAHNEJAD, A. *et al.* Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. **Forests**, [s.l.], v. 8, n. 2, p.42-60, 2017.
- ALPAYDIN, E. **Introduction to machine learning**. Boston: MIT press, 2009.
- ALVES, J. de A. *et al.* Equações hipsométricas para *Toona ciliata* com inclusão de covariantes. **Ciência Florestal**, Santa Maria, v. 27, n. 2, p. 581-595, 2017
- BARROS, D. A. *et al.* Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. **Boletim de Pesquisa Florestal**, Colombo, n. 45, p. 03-28, 2002.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **The Journal of Machine Learning Research**, Montréal, v. 13, n. 1, p.281-305, 2012.
- BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G. Configuração de Redes Neurais Artificiais para Estimação do Volume de Árvores. **Revista Ciência da Madeira**, Pelotas, v. 5, n. 1, p.58-67, 2014.
- BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v. 37, n. 4, p. 639-645. 2013.
- BREIMAN, L. Random Forests. **Machine Learning**, [s.l.], v. 45, n. 1, p.05-32, 2001.
- BREUSCH, T.; PAGAN, A. A simple test for heteroscedasticity and random coefficient variation. **Econometrica**, Menasha, v.47, p.1287–1294, 1979.
- CALEGARIO, N. *et al.* Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não-lineares generalizados. **ScientiaForestalis**, Piracicaba, v. 69, p. 38-50, 2005.
- CANDEL, A. *et al.* **Deep Learning with H2O**. Disponível em: < <http://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/DeepLearningBooklet.pdf> >. Acesso em: 02 jul. 2017.
- CORDEIRO, M. A., *et al.* Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. **Pesquisa Florestal Brasileira**, Colombo, v. 35, n. 83, 255-261, 2015.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. **An introduction to support vector machines and other kernel-based learning methods**. New York: Cambridge University Press, 2000, 189 p.
- FACELI, K. *et al.* **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. 378p.

GORGENS, E. B. *et al.* Influência da arquitetura na estimativa de volume de árvores individuais por meio de redes neurais artificiais. **Revista Árvore**, Viçosa, v. 38, n. 2, p.289-295, abr. 2014.

GORGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. **Computers and Electronics in Agriculture**, v. 116, p. 221-227, 2015.

GUJARATI, D. **Econometria básica**. São Paulo: Elsevier; 2006.

HAARA, A.; MALTAMO, M.; TOKOLA, T. The K-nearest-neighbour method for estimating basal-area diameter distribution. **Scandinavian Journal of Forest Research**, v. 12, n. 2, p. 200-208, 1997.

HECHENBICHLER, K; SCHLIEP, K. **Weighted k-Nearest-Neighbor Techniques and Ordinal Classification**. Discussion paper 399, Ludwig-Maximilians University Munich, Munich, 2004. Disponível em: < https://epub.ub.uni-muenchen.de/1769/1/paper_399.pdf>. Acesso em: 20 jun. 2017.

HSU, C. W; CHANG, C. C; LIN, C. J. **A Practical Guide to Support Vector Classification**. Technical report, Department of Computer Science, National Taiwan University, 2010

MARTINS, E. R. *et al.* Configuração de redes neurais artificiais para estimação da altura total de árvores de eucalipto. **Revista Brasileira de Ciências Agrárias (Agrária)**, Recife, v. 11, n. 2, p. 117-123, 2016.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**. v. 1, n. 1, p. 32, 2003.

ÖZÇELİK, R. *et al.* Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. **Forest Ecology And Management**, [s.l.], v. 306, p.52-60, out. 2013.

ROBINSON, A. P.; LANE, S. E.; THÉRIEN, G. Fitting forestry models using generalized additive models: a taper model example. **Canadian Journal of Forest Research**, [s.l.], v. 41, n. 10, p.1909-1916, out. 2011.

SANQUETTA, C. R. *et al.* On the use of data mining for estimating carbon storage in the trees. **Carbon Balance And Management**, [s.l.], v. 8, n. 1, p.6-14, 2013.

SENA, A. L. M. *et al.* Modelos não lineares com uso de covariantes para relação hipsométrica de duas espécies de *Pinus* topicais. **Ciência Florestal**, Santa Maria, v. 25, n. 4, p. 969-980, 2015.

SHALEV-SHWARTZ, S.; BEM-DAVID, S. **Understanding Machine Learning: From Theory to Algorithms**. New York: Cambridge University Press, 1. ed., 2014.

SMOLA, A. Regression Estimation with Support Vector Learning Machines. 1996. 78 f. Master's Thesis – Physics, Technische Universität München, München, 1996.

- TOMMOLA, M., *et al.* Estimating the characteristics of a marked stand using k-nearest-neighbour regression. **Journal of Forest Engineering**, [s.l.], v. 10, p.75-81, 1999.
- VAPNIK, V N. **The nature of statistical learning theory**. 2. ed. New York: Springer-Verlag, 2000. 314 p.
- VENDRUSCOLO, D. G. S. *et al.* Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v. 33, n. 4, p. 556-569. 2015.
- VENDRUSCOLO, D. G. S. *et al.* Height prediction of *Tectona grandis* trees by mixed effects modelling and artificial neural networks. **International Journal of Current Research**, [s.l.], v. 8, n. 12, p.43189-43195, 2016.
- VENDRUSCOLO, D. G. S. *et al.* Estimativa da altura de árvores de *Tectona grandis* Lf utilizando regressão e redes neurais artificiais. **Nativa**, Sinop, v. 5, n. 1, p. 52-58. 2017.

CAPÍTULO 2 – REDUÇÃO NA QUANTIDADE DE MEDIÇÕES DE ALTURA EM INVENTÁRIOS DE *Pinus taeda* L. PELO EMPREGO DE MODELOS NÃO LINEARES E TÉCNICAS DE APRENDIZADO DE MÁQUINA

RESUMO

A tarefa de medição da altura das árvores é um componente expressivo nos custos de inventários florestais. Visando avaliar a viabilidade da redução do número de medições de altura a serem realizadas, foram gerados dois subconjuntos a partir de uma base de dados de inventário florestal realizado em plantios de *Pinus taeda*. A base completa totaliza 1.336 parcelas circulares de 600m² de área, instaladas em povoamentos entre 5 e 21 anos e espaçamento de 4 a 9 m² por árvore, sendo que em cada parcela foram medidas as alturas de em média 20 árvores (47 no máximo e 5 no mínimo). Subconjuntos denominados A3 e A5 foram gerados de modo a simular a redução do número de alturas medidas por parcela para 3 e 5 árvores, respectivamente. Três técnicas de regressão foram testadas na modelagem da relação hipsométrica para cada subconjunto de dados: Redes Neurais Artificiais (RNA), Máquinas de Vetores de Suporte (MVS) e Modelo não linear de Gompertz com adição de covariantes, totalizando seis testes. Como testemunha, foi aplicado o método tradicionalmente utilizado pela empresa que disponibilizou os dados, que consiste no ajuste de modelos por estrato, cujo critério de agrupamento é: idade, fase de manejo, localização e classe de sítio. Para ajuste/treinamento dos modelos, foram utilizadas como variáveis preditoras: diâmetro à 1,3 m do solo, altura dominante da parcela, diâmetro quadrático médio da parcela e idade. Os indicadores de qualidade (erro quadrático médio e erro absoluto médio), análises gráficas de resíduos, análise de variância e teste de comparação de médias de Dunnett indicaram que a redução do número de medições de altura por parcela para 5 permitiu que as três técnicas de regressão apresentassem resultados equivalentes ao método tradicional, o que não ocorreu para os ajustes ao conjunto com 3 árvores, que tiveram desempenho menos satisfatório. As técnicas de aprendizado de máquina apresentaram indicadores de desempenho mais favoráveis em comparação ao modelo de Gompertz modificado; no entanto, apresentam maior complexidade para replicação, representação e interpretação dos parâmetros em comparação ao modelo não linear. Concluiu-se que em inventários florestais de *P. taeda* a mensuração da altura de cinco árvores por unidade amostral é suficiente para descrever a relação hipsométrica, empregando RNA, MVS ou modelos não lineares com inclusão de covariantes.

Palavras-chave: Máquinas de Vetores de Suporte; Redes Neurais Artificiais; relação hipsométrica; inventário florestal.

CHAPTER 2 – REDUCTION IN THE NUMBER OF HEIGHT MEASUREMENTS IN *Pinus taeda* L. INVENTORIES USING NON-LINEAR MODELS AND MACHINE LEARNING ALGORITHMS

ABSTRACT

The task of measuring tree height is an expressive component in the costs of forest inventories. In order to evaluate the viability of reducing the number of height measurements, two subsets were generated from a forest inventory dataset of *Pinus taeda* stands. The full data set consists of 1,336 circular plots of 600 m² of area, plotted in 5 to 21 years old stands, with spacing from 4 to 9 m² per tree, where 20 height measurements were taken, on average (47 at most and 5 at least). The subsets A3 and A5, refer to the simulation of the reduction in the number of height measurements per plot to 3 and 5 trees, respectively. Three regression techniques were tested in the modeling of the height-diameter relationship for each subset of data: Artificial Neural Networks, Support Vector Machines and the Gompertz non-linear model with covariates inclusion, totalizing six tests. The traditional method used by the company that provided the data for this study was the control method, consisting of the fit of models by stratum, whose the grouping criteria were: age, management phase, location and site index. For the adjustment/training of the models, the following variables were used as predictors: diameter at breast height, plot dominant height, plot quadratic mean diameter and age. The goodness of fit indicators, residual plots, analysis of variance and Dunnett test indicated that the reduction in the number of height measurements per plot to 5 allowed the three regression techniques to present equivalent results to the traditional method, what did not occur for the A3 data set, which presented less accuracy. The machine learning techniques presented higher performance indicators than the modified Gompertz model; however, they present greater complexity for replication, representation and interpretation of the parameters in comparison to the nonlinear model. It was concluded that in *P. taeda* forest inventories the measurement of the height of five trees per plot is sufficient to describe the height-diameter relation, using RNA, MVS or non-linear models with covariates inclusion.

Keywords: Support Vector Machines; Artificial Neural Networks; artificial intelligence; forest inventory; height-diameter relationship

2.1 INTRODUÇÃO

A obtenção da altura das árvores em inventário florestais é fundamental para estimativa do volume da biomassa ou da produção de madeira. Pela sua difícil obtenção, na prática as empresas florestais mensuram a altura de apenas algumas árvores na unidade amostral, utilizando relações hipsométricas para estimativa das demais. No entanto, a medição por meio de instrumentos da altura de árvores em pé é uma operação onerosa e sujeita a erros.

O ajuste de modelos empíricos a nível de parcela é considerado o método mais apropriado para representação da relação hipsométrica (SOARES *et al.*, 2004). Modelos lineares genéricos que utilizam variáveis como idade e altura dominante para representar características específicas dos povoamentos, que constituem um conjunto de dados heterogêneo, podem substituir modelos ajustados a nível de parcela devido apresentarem boa performance de ajuste, tendo como benefício a substituição de diversos modelos por apenas um (BARROS *et al.*, 2002; SOARES *et al.*, 2004).

Apesar do tradicional uso de modelos lineares para representação da relação hipsométrica, modelos não lineares são mais apropriados e podem apresentar melhor desempenho na estimativa das alturas. Especialmente com a adição de covariantes, como idade e altura dominante para estimativa dos parâmetros do modelo (SENA *et al.*, 2015).

Recentemente, a utilização de técnicas de inteligência artificial para tarefas de regressão tem apresentado bons resultados na modelagem da relação hipsométrica a partir de conjuntos de dados que agrupam povoamentos com diferentes idades, espaçamentos e qualidades de sítio (BINOTI *et al.*, 2013b; VENDRUSCOLO *et al.*, 2017).

Binoti *et al.* (2012) avaliaram o efeito da redução gradativa do número de parcelas a compor o conjunto de treinamento de Redes Neurais Artificiais aplicadas à modelagem da relação hipsométrica de povoamentos de eucaliptos. Os autores concluíram que a utilização da técnica permite a redução do custo de inventários florestais por meio da redução no número de alturas a serem medidas, sem perda significativa na precisão das estimativas.

A tarefa de medição da altura das árvores está diretamente relacionada ao rendimento das equipes de campo, sendo um componente expressivo na formação do custo de inventários florestais (LEITE e ANDRADE, 2003).

É possível reduzir o número de alturas medidas nas parcelas sem perder precisão em relação ao método tradicional valendo-se de técnicas de AM e regressão com covariantes para *Pinus taeda*? Essa questão motivou a apresentação da presente pesquisa, a qual teve como objetivo geral avaliar a utilização de métodos de aprendizado de máquina e modelos não lineares na modelagem da relação hipsométrica, sob duas condições de intensidade de medições: 3 e 5 árvores por parcela.

Os objetivos específicos foram: avaliar o desempenho do modelo não linear de Gompertz, Redes Neurais Artificiais e Máquinas de Vetores de Suporte na representação da relação hipsométrica a partir de conjuntos de dados que simulam a medição da altura de 3 e de 5 árvores por unidade amostral; e comparar a performance dos métodos propostos e do método tradicionalmente utilizado pela empresa que disponibilizou os dados utilizados neste estudo.

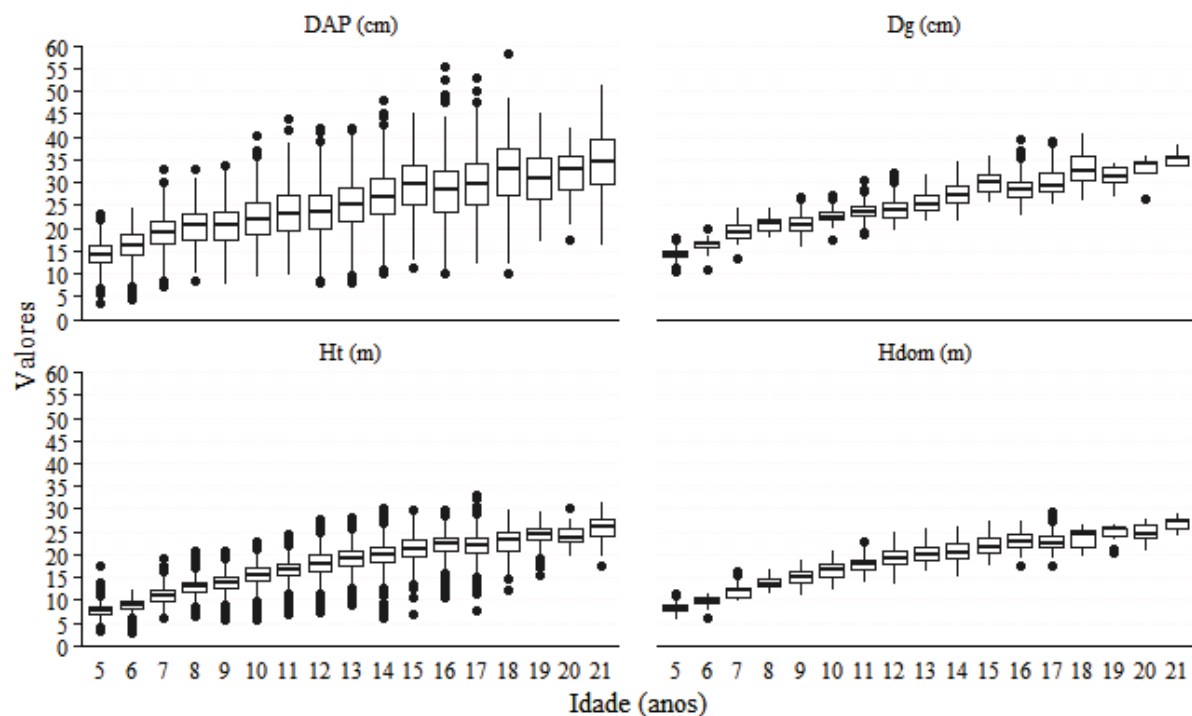
2.2 MATERIAIS E MÉTODOS

2.2.1 BASE DE DADOS

Os dados utilizados no presente estudo são oriundos de plantios comerciais de *Pinus taeda* localizados na região centro sul do Paraná, manejados sob regime de múltiplos produtos da madeira, com espaçamentos de plantio variados (de 4 a 9 m²/árvore). A base de dados para estudo da relação hipsométrica foi constituída por 23.754 árvores com suas respectivas medidas de diâmetro à 1,30 m do solo (DAP), altura total, altura dominante da parcela, amostradas em 1.336 parcelas circulares de 600 m² de área, instaladas em povoamentos entre 5 e 21 anos de idade e espaçamento de 4 a 9 m² por árvore. Em média, foram medidas as alturas de 20 árvores por parcela, no entanto a base de dados não considerou árvores quebradas, tortas, dominadas e rebrotas, o que reduziu à 18 árvores por parcela, em média.

O resumo estatístico das variáveis contidas na base de dados está apresentado na Figura 2.1.

FIGURA 2.1 - ESTATÍSTICA DESCRITIVA DA BASE DE DADOS OBTIDOS EM INVENTÁRIOS EM POVOAMENTOS DE *Pinus taeda* NO ESTADO DO PARANÁ.



FONTE: O autor (2019).

NOTA: Variáveis *Dg* e *Hdom* obtidas à nível de unidade amostral.

As unidades amostrais foram ordenadas em função da idade do povoamento, e separadas sistematicamente de modo a formar um conjunto de treinamento/ajuste (90%) e de teste (10%).

2.2.2 REDUÇÃO DO NÚMERO DE MEDIÇÕES DE ALTURAS

Foram simuladas duas situações de redução da intensidade amostral a partir de subamostragens do conjunto de treinamento/ajuste. Sendo elas:

I. Método A3: consistiu em subamostrar três árvores por parcela, correspondendo a primeira árvore mensurada, a de maior e a de menor diâmetro. Quando a primeira árvore coincidiu em ser a de menor ou maior diâmetro, então foi incluída na subamostragem a árvore subsequente.

II. Método A5: consistiu em uma subamostragem tal qual a realizada para gerar A3, porém garantindo sempre 5 árvores por parcela, sendo as 3 de maior diâmetro, a de menor diâmetro e a primeira diferente das quatro citadas.

Desse modo, foram geradas duas novas bases de treinamento/ajuste, as quais simularam a mensuração do método A3 e A5.

Nos conjuntos A3 e A5 a variável altura dominante ($Hdom$), que pelo método tradicional consiste na média das alturas das seis árvores de maior diâmetro, foi substituída pelas variáveis hd_1 e hd_3 , que representam a altura da árvore de maior diâmetro, e a média da altura das três árvores de maior diâmetro, respectivamente.

2.2.3 MÉTODO DE AJUSTE TRADICIONAL

Como testemunha para comparação aos métodos propostos, empregou-se o método de modelagem da relação hipsométrica utilizada pela empresa que disponibilizou os dados do estudo. O método consiste no ajuste de diferentes modelos por estrato, selecionando o que produziu o menor erro padrão (s_{yx}). A estratificação empregada pela empresa consiste em agrupar povoamentos de mesma idade e fase de manejo, que se situem em fazendas próximas e de índice de sítio similar. A base de dados utilizada conta com, ao todo, 146 estratos.

Para este estudo, foram ajustados três modelos por estrato os quais, segundo a empresa, são mais frequentemente utilizados na predição da altura dos povoamentos:

$$Ht = \beta_0 + \beta_1 DAP + \beta_2 DAP^2 + \beta_3 Hdom + \varepsilon$$

$$Ht = \beta_0 + \beta_1 DAP + \beta_2 Hdom + \varepsilon$$

$$\ln(Ht) = \beta_0 + \beta_1 \ln(DAP) + \varepsilon$$

em que:

β_1, β_2 e β_3 = parâmetros dos modelos;

DAP = diâmetro à 1,30m do solo;

$Hdom$ = altura dominante;

\ln = logaritmo neperiano;

ε = erro estatístico inerente à modelagem.

Para cada estrato, o modelo que conferiu ao ajuste o menor erro padrão foi utilizado para predição das alturas.

2.2.4 TREINAMENTO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA

Dois métodos de aprendizado de máquina foram aplicados no presente estudo, por terem apresentado melhores resultados no primeiro capítulo desta pesquisa: Máquinas de Vetores de Suporte (MVS) e Redes Neurais Artificiais (RNA). Foram utilizadas, como variáveis preditoras: altura total o diâmetro à altura do peito (DAP), média da altura das n árvores dominantes (Hd_n), idade (Id) e diâmetro médio quadrático da parcela (Dg).

Para evitar a seleção de modelos sobreajustados aos dados de treinamento e de baixa capacidade de generalização, o treinamento dos modelos de AM foi realizado por meio da aplicação do método de validação cruzada k -fold, combinado ao método de busca em grade. Segundo Shalev-Shwartz e Ben-David (2014), o método k -fold consiste em particionar a base de dados aleatoriamente em k subconjuntos de dimensões similares e, para cada configuração do algoritmo, realizar k treinamentos, separando um subconjunto por vez para validação. Desse modo, todas as observações são utilizadas $k-1$ vezes para treinamento e uma vez para validação. A medida de desempenho de cada configuração é obtida pela média do erro quadrático médio das k validações. Após a etapa de validação cruzada, o algoritmo efetua o treinamento de um modelo final para a configuração testada, utilizando toda a base de dados de treinamento.

Para cada algoritmo de AM, a melhor configuração foi a que conferiu ao modelo o menor erro quadrático médio obtido por validação cruzada. Os pacotes R utilizados realizam a validação cruzada k -fold automaticamente a partir da determinação do número de subconjuntos ($folds$) pelo usuário. Neste estudo, utilizou-se o método 3-fold. O treinamento dos algoritmos foi realizado no software R Studio 1.0.136.

2.2.4.1 Máquinas de Vetores de Suporte

Foram treinadas MVS utilizando *kernel* função de base radial e função de erro tipo II (regressão-v), a qual possui o parâmetro v utilizado para determinar a proporção

entre o número de vetores de suporte e o número total de instâncias na base de treinamento. Utilizou-se $\nu = 0,5$, valor padrão do pacote R utilizado (*e1071*).

Os parâmetros C e γ foram ajustados pelo método de busca em grade. A busca foi realizada conforme o método sugerido por Hsu *et al.* (2003), que consiste em, inicialmente, realizar uma busca exploratória através de valores intervalados para cada um dos parâmetros, e após a identificação de uma região promissora, onde os erros são menores, esses intervalos são reduzidos para realização de uma busca refinada. O treinamento das MVS foi realizado por meio da função *svm* presente no pacote *e1071* do R.

Na busca exploratória testou-se para a variável γ valores de 2^{-10} até 2^{-1} , em intervalos de 2^1 , enquanto para a variável C testou-se valores de 2^1 até 2^{10} , com o mesmo intervalo. Considerando que os melhores parâmetros obtidos pela busca exploratória sejam 2^{-n} e 2^n para γ e C , a busca refinada testou para γ valores de $2^{-n-0,5}$ até $2^{-n+0,5}$, e para C valores de $2^{n-0,5}$ até $2^{n+0,5}$, adotando em ambos os casos intervalo de $2^{0,2}$.

2.2.4.2 Redes Neurais Artificiais

Redes Neurais Artificiais do tipo *Perceptron* de múltiplas camadas foram treinadas por meio do algoritmo de retropropagação dos erros, utilizando-se a função *h2o.deeplearning*, do pacote *h2o* do R. Foi utilizada como função de ativação dos neurônios da camada oculta a Tangente Hiperbólica, descrita pela expressão a seguir.

$$f(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}$$

em que:

e = número de Euler (2,71828...); e

$$\alpha = \sum_i w_i x_i + b$$

em que:

w = vetor de pesos sinápticos;

x = vetor de entradas;

$b = bias$ ou viés da camada.

A quantidade de neurônios na camada oculta foi determinada a partir do método de busca em grade, por meio do treinamento de redes que continham de 1 a 10 neurônios na camada oculta.

2.2.4.3 Ajuste do modelo não linear

O modelo não linear foi ajustado por meio da função *nls.lm* do pacote *minpack.lm* do R, o qual utiliza uma modificação do algoritmo Levenberg-Marquadt para resolução de problemas de mínimos quadrados não lineares. Foi utilizado para ajuste o modelo de Gompertz modificado pela decomposição dos parâmetros e inserção das covariantes altura dominante, idade e diâmetro médio quadrático, além da variável DAP. As covariantes para cada parâmetro foram selecionadas por procedimento *backward*, sendo mantidas no modelo somente as significativas à 1% de probabilidade. A descrição matemática do modelo de Gompertz é dada por:

$$Ht = \beta_0 * e^{-\beta_1 * e^{-\beta_2 * DAP}} + \varepsilon$$

O modelo de Gompertz modificado pela inclusão de covariantes, o qual foi testado no presente estudo, pode ser descrito matematicamente por:

$$Ht = (\beta_0 + \beta_{01} * hd_n + \beta_{03} * dg) * e^{-(\beta_1 + \beta_{12} * id) * e^{-(\beta_2 + \beta_{21} * hd_n) * DAP}} + \varepsilon$$

em que:

$\beta_0, \beta_1, \beta_2, \beta_{01}, \beta_{03}, \beta_{12}, \beta_{21}$ = parâmetros do modelo;

Hd_n = altura da árvore dominante (A3) ou média das alturas das três árvores dominantes (A5).

2.2.5 AVALIAÇÃO DO DESEMPENHO DOS MÉTODOS

Após o treinamento dos algoritmos de aprendizado de máquina e ajuste do modelo não linear para as bases de dados A3 e A5, e o ajuste pelo método tradicional,

essas técnicas foram aplicadas à predição das alturas em um conjunto de dados desconhecidos.

Os métodos foram avaliados quanto ao desempenho na predição da altura total e quanto ao impacto nas predições dos volumes total e por sortimento.

2.2.5.1 Desempenho na estimativa da altura total

O desempenho dos métodos na estimativa da altura total de árvores desconhecidas foi avaliado pelo Erro Quadrático Médio (EQM%), e pelo Erro Absoluto Médio (EAM), descritos pelas expressões a seguir.

$$erro_i = Ht_{i_{obs}} - Ht_{i_{est}}$$

$$EQM\% = \frac{\sqrt{\sum_{i=1}^n erro_i^2}}{n} * \frac{100}{\overline{Ht}}$$

$$EAM\% = \frac{\sum_{i=1}^n |erro_i|}{n} * \frac{100}{\overline{Ht}}$$

em que:

$Ht_{i_{obs}}$ = altura total observada para a árvore i ;

$Ht_{i_{est}}$ = altura total estimada para a árvore i ;

n = número de observações;

\overline{Ht} = altura total média das árvores observadas.

Ainda, os modelos foram comparados por meio de análise de gráfica de resíduos e da curva de predição.

2.2.5.2 Impacto nos volumes total e por sortimento

O volume total e por sortimento para cada árvore do conjunto de teste foram estimados por meio de uma função de afilamento ajustada a partir de dados de cubagem disponibilizados pela empresa. Foi utilizado como função de afilamento o polinômio do 5º grau, que ajustado aos dados de cubagem, resultou a expressão que segue:

$$\frac{d_i}{DAP} = 1,24756 - 3,91343 * \left(\frac{h_i}{Ht}\right) + 15,44685 * \left(\frac{h_i}{Ht}\right)^2 - 32,73925 * \left(\frac{h_i}{Ht}\right)^3 + 30,79162 \left(\frac{h_i}{Ht}\right)^4 - 10,86377 \left(\frac{h_i}{Ht}\right)^5$$

em que:

d_i = diâmetro i ao longo do fuste;

h_i = altura i ao longo do fuste.

Os volumes estimados a partir das alturas observadas compuseram os valores para a testemunha, para que os volumes estimados a partir das alturas preditas pelos diferentes métodos pudessem ser avaliados por meio de uma modificação do EAM%, denominada diferença absoluta média percentual (DAM%), conforme as seguintes expressões:

$$diferença_i = V_{i\text{ obs}} - V_{i\text{ est}}$$

$$DAM\% = \frac{100}{\bar{V}} * \frac{\sum_{i=1}^n |diferença_i|}{n}$$

em que:

$V_{i\text{ obs}}$ = volume total ou do sortimento da árvore i estimado a partir da altura observada;

$V_{i\text{ est}}$ = volume total ou do sortimento da árvore i estimado a partir da altura estimada;

\bar{V} = média do volume total ou do sortimento observado.

A estimativa do volume total (Vt) considerou o fuste da árvore desde sua base até a ponta. Para a estimativa dos volumes por sortimento (Vs_i), foi considerada uma altura de toco de 10 centímetros e as medidas indicadas na Tabela 2.1.

TABELA 2.1 - SORTIMENTOS UTILIZADOS PARA ESTIMATIVA DOS VOLUMES COMERCIAIS.

Sortimento	Mínimo DPF (cm)	Máximo DPF (cm)	Comprimento (m)
VS_1	33	99	2,65
VS_2	23	33	2,65
VS_3	15	23	2,65
VS_4	8	15	2,4

FONTE: O autor (2019).

NOTA: DPF = diâmetro na ponta fina.

O volume das seções de uma árvore foi estimado a partir da integração da equação de afilamento. A integral definida do polinômio do 5º grau está implementada no Florexel, suplemento para MS Excel o qual foi utilizado neste estudo para estimativa dos volumes totais e por sortimento.

2.2.6 ANÁLISE DE VARIÂNCIA E TESTE DE COMPARAÇÃO DE MÉDIAS

Para a realização da análise de variância, optou-se pelo delineamento em blocos casualizados (DBC), de modo a isolar o efeito da variabilidade nas dimensões das árvores do grupo de teste. Cada método foi considerado um tratamento e cada árvore um bloco, cujos valores comparados se referem à diferença absoluta em metros entre a altura observada e a estimada para cada árvore do conjunto de teste. O método tradicional de modelagem foi considerado o controle, sendo comparado com todos os demais métodos pelo teste de comparação de médias de Dunnett a 5% de probabilidade.

2.3 RESULTADOS E DISCUSSÃO

2.3.1 AJUSTES DE MODELOS DE REGRESSÃO ESTATÍSTICA

O método tradicional contemplou o ajuste de 438 modelos, sendo três para cada um dos 146 estratos de plantio, sendo o melhor selecionado para modelagem da relação hipsométrica do estrato. O modelo linear 1 foi superior em 93 estratos (64%), enquanto o modelo linear 2 foi em 44 (30%), e o 3 em apenas 9 (6%). Os modelos escolhidos pelo método tradicional e os parâmetros estimados estão apresentados no Apêndice 1.

As estimativas dos parâmetros para os modelos não lineares ajustados as bases A3 e A5 estão apresentadas na Tabela 2.2. Todos os coeficientes dos dois modelos foram significativos a 5% de probabilidade.

TABELA 2.2 - ESTIMATIVAS DOS PARÂMETROS PARA OS MODELOS NÃO LINEARES AJUSTADOS PARA A RELAÇÃO HIPSOMÉTRICA DE ÁRVORES DE *Pinus taeda* NO PARANÁ.

Base de ajuste	Parâmetro	Estimativa	Erro padrão	t_c	Valor-p
A3	β_0	1,8706	0,25	7,38	> 0,001
	β_{01}	0,9941	0,02	65,43	> 0,001
	β_{03}	-0,0185	0,01	-2,22	0,027
	β_2	1,4933	0,08	19,65	> 0,001
	β_{22}	-0,0397	0,00	-9,42	> 0,001
	β_3	0,1330	0,01	16,47	> 0,001
	β_{31}	-0,0024	0,00	-8,50	> 0,001
A5	β_0	2,4640	0,28	8,92	> 0,001
	β_{01}	1,0409	0,01	71,03	> 0,001
	β_{03}	-0,0776	0,01	-10,79	> 0,001
	β_2	1,2200	0,06	20,38	> 0,001
	β_{22}	-0,0138	0,00	-3,33	> 0,001
	β_3	0,1094	0,01	13,53	> 0,001
	β_{31}	-0,0010	0,00	-2,96	0,003

FONTE: O autor (2019).

2.3.2 REDES NEURAIS ARTIFICIAIS

2.3.2.1 Conjunto de treinamento A3

Ao aplicar o treinamento no conjunto A3, a rede neural que produziu o menor erro quadrático médio na validação cruzada *k-fold* foi treinada com sete neurônios na camada intermediária.

A partir da obtenção dos pesos sinápticos e dos *bias* das camadas que compõem a RNA, essa pode ser reproduzida a partir de um conjunto de equações. As expressões a seguir, padronizam as entradas a partir da média e do desvio-padrão das variáveis no conjunto de treinamento.

$$Id_p = \frac{(Id - 12,03577)}{3,69174}$$

$$DAP_p = \frac{(DAP - 23,92868)}{8,82529}$$

$$Dg_p = \frac{(Dg - 23,86326)}{5,38666}$$

$$Hd_{1p} = \frac{(Hd_1 - 18,68979)}{4,87711}$$

De posse das variáveis padronizadas, estas devem ser inseridas nas expressões apresentadas a seguir, que representam as entradas α da função de ativação.

$$\begin{aligned}\alpha_1 &= -0,29807 * Id + 0,32751 * DAP - 0,11883 * Dg - 0,38466 * Hd_1 - 0,59788 \\ \alpha_2 &= -0,34853 * Id - 0,11141 * DAP - 0,16776 * Dg + 0,90006 * Hd_1 - 0,03791 \\ \alpha_3 &= -0,22528 * Id + 0,18346 * DAP - 0,28049 * Dg - 0,37735 * Hd_1 + 0,50759 \\ \alpha_4 &= -0,07518 * Id - 0,64691 * DAP - 0,05746 * Dg + 0,25849 * Hd_1 - 0,78055 \\ \alpha_5 &= -0,03691 * Id - 0,58995 * DAP + 0,13574 * Dg + 1,50851 * Hd_1 + 0,18029 \\ \alpha_6 &= 0,58179 * Id - 0,00044 * DAP - 0,33384 * Dg + 0,09494 * Hd_1 - 0,45018 \\ \alpha_7 &= -0,46758 * Id + 0,74055 * DAP - 0,19991 * Dg + 0,44884 * Hd_1 - 0,78298\end{aligned}$$

Os valores α devem ser inseridos na função de ativação tangente hiperbólica (Φ), conforme a expressão a seguir. Os resultados representarão as saídas dos sete neurônios da camada intermediária.

$$\Phi_n = \frac{e^{\alpha_n} - e^{-\alpha_n}}{e^{\alpha_n} + e^{-\alpha_n}}$$

Por fim, as saídas Φ serão utilizadas na equação abaixo, que descreve em uma única expressão o resultado da camada de saída que possui função de ativação linear e a despadronização do resultado em função da média e do desvio padrão da variável altura total (Ht) no conjunto de dados A3.

$$\begin{aligned}Ht &= (-0,84524 * \Phi_1 + 0,80079 * \Phi_2 - 0,62723 * \Phi_3 - 1,15929 * \Phi_4 - 0,25654 * \Phi_5 + 0,20118 * \Phi_6 \\ &\quad + 0,36673 * \Phi_7 - 0,45899) * 5,09869 + 17,21378\end{aligned}$$

2.3.2.2 Conjunto de treinamento A5

A rede treinada com treze neurônios na camada intermediária foi a que produziu o menor erro quadrático médio, dentre as 15 redes treinadas a partir do conjunto A5. Para replicação da rede, estão apresentadas equações que representam as operações matemáticas realizadas pela RNA para predição de novos valores. Inicialmente, as variáveis de entrada devem ser padronizadas conforme as expressões a seguir.

$$Id_p = \frac{(Id - 12,03588)}{3,69510}$$

$$DAP_p = \frac{(DAP - 26,19968)}{8,30855}$$

$$Dg_p = \frac{(Dg - 24,88186)}{5,39075}$$

$$Hd_{3p} = \frac{(Hd_1 - 18,52840)}{4,81253}$$

A seguir, as variáveis padronizadas devem ser substituídas nas seguintes expressões:

$$\alpha_1 = -0,03596 * Id_p + 0,22771 * DAP_p - 0,00008 * Dg_p - 0,50349 * Hd_{3p} - 0,44491$$

$$\alpha_2 = -0,24866 * Id_p + 0,01222 * DAP_p - 0,04202 * Dg_p + 0,60029 * Hd_{3p} - 0,11329$$

$$\alpha_3 = 0,00299 * Id_p - 0,05433 * DAP_p - 0,05151 * Dg_p - 0,4264 * Hd_{3p} + 0,4447$$

$$\alpha_4 = -0,05889 * Id_p - 0,5014 * DAP_p - 0,0728 * Dg_p + 0,18862 * Hd_{3p} - 0,67181$$

$$\alpha_5 = 0,16218 * Id_p - 0,26652 * DAP_p + 0,30383 * Dg_p + 1,17577 * Hd_{3p} + 0,20062$$

$$\alpha_6 = -0,00495 * Id_p - 0,00094 * DAP_p + 0,00756 * Dg_p + 0,00123 * Hd_{3p} + 0,0029$$

$$\alpha_7 = -0,34259 * Id_p + 0,3507 * DAP_p + 0,00568 * Dg_p + 0,31058 * Hd_{3p} + 0,00227$$

$$\alpha_8 = -0,12434 * Id_p - 0,1565 * DAP_p - 0,08218 * Dg_p - 0,08606 * Hd_{3p} + 0,03071$$

$$\alpha_9 = -0,47253 * Id_p - 0,7583 * DAP_p - 0,38195 * Dg_p + 0,88225 * Hd_{3p} - 0,04906$$

$$\alpha_{10} = 0,04397 * Id_p + 0,01724 * DAP_p - 0,05597 * Dg_p - 0,00464 * Hd_{3p} + 0,00124$$

$$\alpha_{11} = 0,35342 * Id_p - 0,10182 * DAP_p + 0,29108 * Dg_p + 0,02175 * Hd_{3p} - 0,1034$$

$$\alpha_{12} = 0,20917 * Id_p + 0,11896 * DAP_p - 0,22801 * Dg_p + 0,2265 * Hd_{3p} - 0,02547$$

$$\alpha_{13} = 0,09687 * Id_p - 0,02367 * DAP_p + 0,39888 * Dg_p - 0,70331 * Hd_{3p} + 0,10226$$

Assim como na RNA treinada a partir do conjunto A3, os valores α devem ser inseridos na função de ativação tangente hiperbólica apresentada anteriormente. Os resultados corresponderão às saídas dos 13 neurônios da camada intermediária e deverão ser inseridos na equação a seguir, a qual resultará no valor predito para a altura.

$$Ht = (-0,63243 * \Phi_1 + 0,51623 * \Phi_2 - 0,42580 * \Phi_3 - 1,13411 * \Phi_4 - 0,15334 * \Phi_5 - 0,00147 * \Phi_6$$

$$+ 0,07199 * \Phi_7 - 0,36652 * \Phi_8 + 0,18867 * \Phi_9 + 0,00213 * \Phi_{10} + 0,41231 * \Phi_{11}$$

$$+ 0,20806 * \Phi_{12} - 0,28901 * \Phi_{13} - 0,50597) * 5,06034 + 17,70355$$

2.3.3 MÁQUINAS DE VETORES DE SUPORTE

2.3.3.1 Conjunto de treinamento A3

A melhor configuração obtida para o conjunto de treinamento A3 foi $\gamma = 2^{-7,5}$ e $C = 2^{7,7}$. A obtenção de uma equação que represente a MVS treinada é uma tarefa mais complexa que a obtenção de equações de RNA. A quantidade de equações geradas é diretamente proporcional ao número de vetores de suporte utilizados para criação do modelo. A MVS treinada para o conjunto A3 utilizou 1.817 vetores de suporte, o que torna inviável a replicação do modelo em forma de equações. Para contornar esta limitação, a replicação da MVS pode ser realizada pelo armazenamento das matrizes de vetores de suporte e de coeficientes ajustados durante o treinamento, bem como o valor γ e o *bias*. A expressão a seguir representa matematicamente as operações para predição de novas observações.

$$y = w \cdot K(x, x') + b$$

Em que:

y = vetor de valores da variável dependente para as novas observações;

w = vetor de coeficientes ajustados no treinamento;

K = *kernel*;

x = matriz de variáveis preditoras das novas observações;

x' = matriz de vetores de suporte

O procedimento de utilização dos componentes armazenados para predição das alturas em novos dados no R está apresentado no Apêndice 2.

2.3.3.2 Conjunto de treinamento A5

A melhor configuração obtida para o conjunto de treinamento A5 foi $\gamma = 2^{-6,5}$ e $C = 2^{6,5}$. Foram utilizados 3.011 vetores de suporte. O procedimento de replicação da MVS está apresentada no Apêndice 2.

2.3.4 ANÁLISE DE INDICADORES ESTATÍSTICOS

As estatísticas de teste dos métodos empregados estão apresentadas na Tabela 2.3. De modo geral o modelo de Gompertz e os algoritmos de aprendizado de máquina foram similares quando foi utilizado para ajuste/treinamento o mesmo conjunto de dados.

Ao serem comparados em relação a EAM% , o ajuste a nível de estrato, que considera a medição das alturas de em média 18 indivíduos por parcela, foi superior aos métodos que contemplam a medição de 5 árvores por parcela. Essa diferença se acentua considerando a medição de 3 árvores por parcela.

Quanto ao EQM%, que dá maior peso a erros de maior magnitude, todos os métodos ajustados ao conjunto A5 foram equivalentes ou levemente superiores ao método tradicional.

A diferença média percentual (DM%) entre os volumes totais calculados a partir das alturas estimadas pelo método tradicional foi superior aos erros obtidos pelos métodos ajustados ao conjunto A5, e inferiores aos métodos ajustados à A3. Os resultados sugerem que os métodos que empregam mais árvores no ajuste/treinamento dos modelos apresentam menor DM% para o sortimento de maior valor (V_{S_1}). Estas diferenças foram mais evidentes neste sortimento que nas demais variáveis analisadas.

Os índices de qualidade indicam similaridade de desempenho entre o método tradicional e os três métodos de regressão aplicados ao conjunto A5. As alturas estimadas pelo modelo de Gompertz e as técnicas RNA e MVS, considerando a medição de 5 árvores por parcela, permitiram estimativas de volume total mais similares aos volumes estimados a partir das alturas observadas. Isso também ocorreu para os sortimentos V_{S_2} e V_{S_3} . Já para os sortimentos V_{S_1} e V_{S_4} , o método tradicional resultou volumes mais similares aos estimados pelas alturas reais.

TABELA 2.3 - ESTATÍSTICAS DE TESTE DOS MÉTODOS EMPREGADOS NA ESTIMATIVA DA ALTURA TOTAL E DOS VOLUMES TOTAL E POR SORTIMENTO PARA POVOAMENTOS DE *Pinus taeda* EM DIVERSAS IDADES NO ESTADO DO PARANÁ.

Método	EAM (m)	EQM (m)	DAM%				
			V _t	VS ₁	VS ₂	VS ₃	VS ₄
Tradicional	0,98 (6,0%)	1,41 (8,6%)	5,53%	7,22%	8,34%	12,35%	17,85%
Gompertz (A3)	1,13 (6,8%)	1,61 (9,7%)	6,17%	10,34%	9,22%	12,73%	18,72%
RNA (A3)	1,11 (6,7%)	1,59 (9,6%)	6,09%	9,34%	9,14%	12,74%	18,59%
MVS (A3)	1,11 (6,7%)	1,59 (9,6%)	6,10%	10,33%	8,99%	12,77%	18,59%
Gompertz (A5)	1,00 (6,1%)	1,41 (8,5%)	5,48%	8,81%	8,26%	12,11%	17,99%
RNA (A5)	1,00 (6,1%)	1,41 (8,6%)	5,49%	8,76%	8,15%	11,97%	18,01%
MVS (A5)	1,00 (6,0%)	1,40 (8,5%)	5,43%	8,78%	8,20%	12,01%	17,91%

FONTE: O autor (2019).

NOTA: EAM = erro absoluto médio na estimativa das alturas; EQM = erro quadrático médio na estimativa das alturas; DM% = diferença média percentual entre os volumes calculados a partir das alturas estimadas e observadas; VS₁, VS₂, VS₃, VS₄ = sortimentos pré-definidos.

Em termos relativos, as diferenças entre o método tradicional e as técnicas empregadas na regressão da relação hipsométrica do conjunto A5 foram baixas (Tabela 2.4). O procedimento utilizado pela empresa apresentou desempenho inferior a estes métodos para os indicadores EQM para altura e DAM para o volume total, e pouco superior para EAM, enquanto os métodos propostos considerando o conjunto A3 tiveram performance inferior ao método tradicional para todos os indicadores, com diferenças superiores a 10% nos indicadores estatísticos.

TABELA 2.4 - DIFERENÇAS PERCENTUAIS OBSERVADAS ENTRE OS INDICADORES ESTATÍSTICOS DOS MÉTODOS PROPOSTOS E DO MÉTODO TRADICIONAL PARA POVOAMENTOS DE *Pinus taeda* EM DIVERSAS IDADES NO ESTADO DO PARANÁ.

Método	ΔEAM	ΔEQM	ΔDAM V _t
Gompertz (A3)	-15,3%	-14,2%	-11,6%
RNA (A3)	-13,3%	-12,8%	-10,1%
MVS (A3)	-13,3%	-12,8%	-10,3%
Gompertz (A5)	-2,0%	0,0%	0,9%
RNA (A5)	-3,1%	0,0%	0,7%
MVS (A5)	-2,0%	0,7%	1,8%

FONTE: O autor (2019).

Os resíduos relativos para o conjunto de teste dos métodos que contemplaram a mensuração de 3 árvores por parcela apresentaram maior amplitude de dispersão e menor uniformidade em relação aos demais métodos, com tendência a subestimativa da altura de árvores de menor porte (Figura 2.2). Todos os métodos apresentaram algumas estimativas com resíduos de maior magnitude para árvores de porte médio, possivelmente resultantes da elevada variância dos dados. Ainda assim,

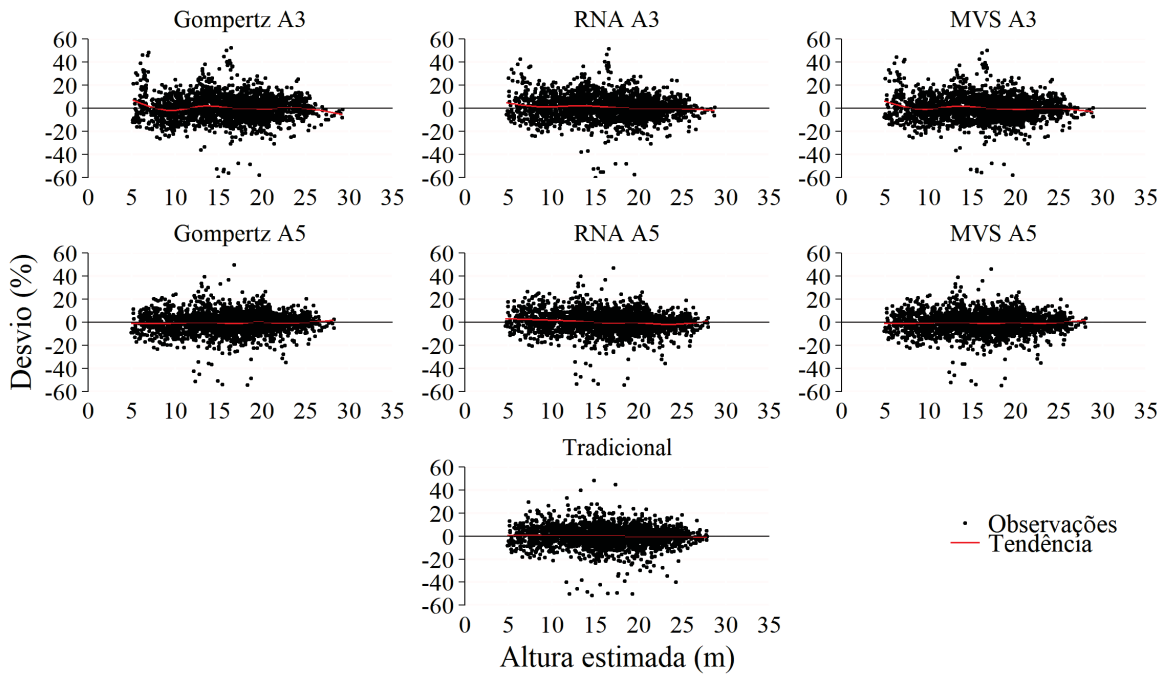
pode-se afirmar que o método tradicional e os métodos ajustados/treinados a partir do conjunto A5 foram satisfatórios quanto a capacidade de explicar a elevada variância dos dados, uma vez que as previsões realizadas nos dados de teste apresentam EAM inferior a 7%, resíduos livres de tendenciosidade, simetricamente distribuídos e concentrados em torno de zero (Figura 2.3), e com dispersão independente das variáveis preditoras DAP (Figura 2.4), *idade* (Figura 2.5), Hd_n (Figura 2.6) e Dg (Figura 2.7).

Considerando que o que diferencia o conjunto A5 do A3 é a presença dos valores de altura medidos de mais duas árvores dominantes, e que os modelos de predição ajustados ao conjunto A5 foram claramente superiores aos ajustados ao A3, independente da técnica de regressão, verifica-se que a variável que representa a qualidade do local é de grande contribuição para a qualidade do modelo de predição. A importância da variável altura dominante na modelagem da relação hipsométrica já foi citada por alguns autores (CAMPOS *et al.*, 1984; LEITE e ANDRADE, 2003), e especialmente neste estudo demonstrou ser fundamental para garantir indicadores estatísticos satisfatórios em face a redução da intensidade amostral.

Usualmente é utilizada para representação da qualidade de sítio o conceito de altura dominante de Assmann, o qual considera como índice a média da altura das cem árvores de maior diâmetro por hectare. Na redução do número de medições de altura representada pelo conjunto A5, a variável Hd_3 composta pela média das três árvores dominantes de cada unidade amostral de 600 m², representa a média da altura das 50 árvores dominantes por hectare. Esta variável consiste em um conceito de altura dominante alternativo ao de Assmann. Segundo a definição de Husch (1982), a altura dominante é a altura média das n árvores de maior diâmetro, ou a média da altura das n árvores mais altas de um povoamento, ou ainda a altura média das árvores do estrato dominante, por uma unidade de área definida.

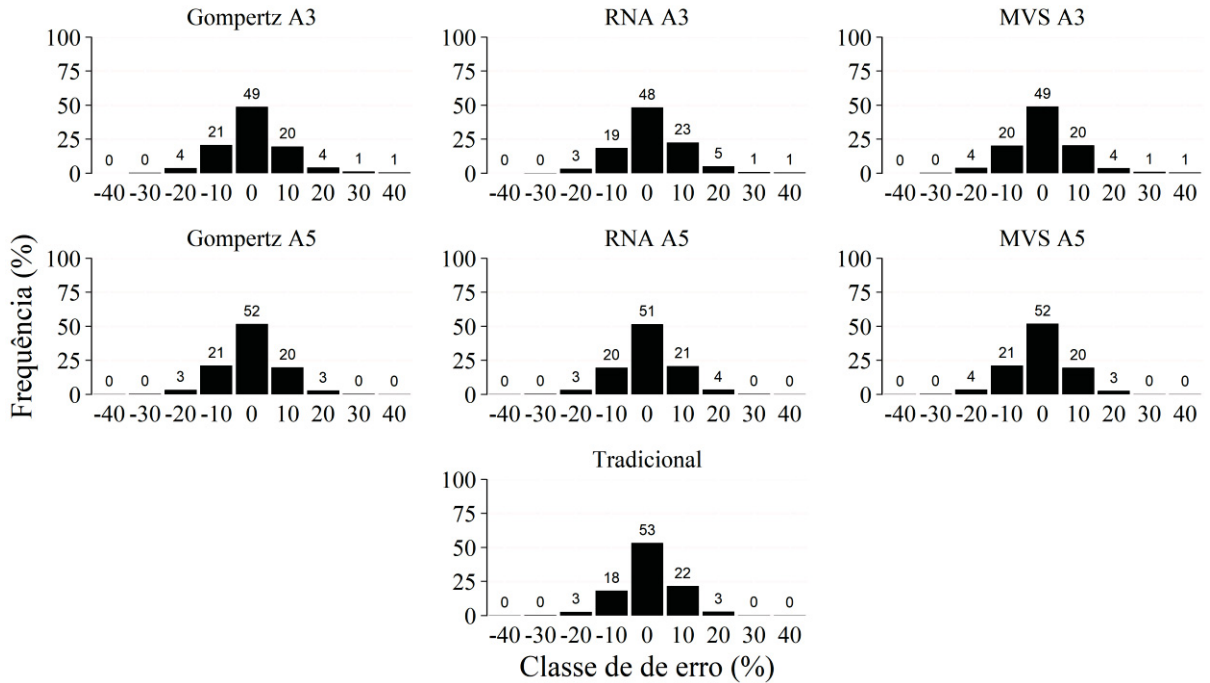
A Figura 2.8 demonstra o comportamento das curvas de predição em função do diâmetro em parcelas de diferentes idades, diâmetros médios quadráticos e alturas dominantes. Nota-se que todos os métodos foram capazes de descrever a relação hipsométrica observada nas parcelas, exceto na unidade amostral 5, em que os métodos ajustados à base A3 e o modelo de Gompertz ajustado à base A5 tenderam a superestimar as alturas.

FIGURA 2.2 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DAS ALTURAS ESTIMADAS PARA POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ.



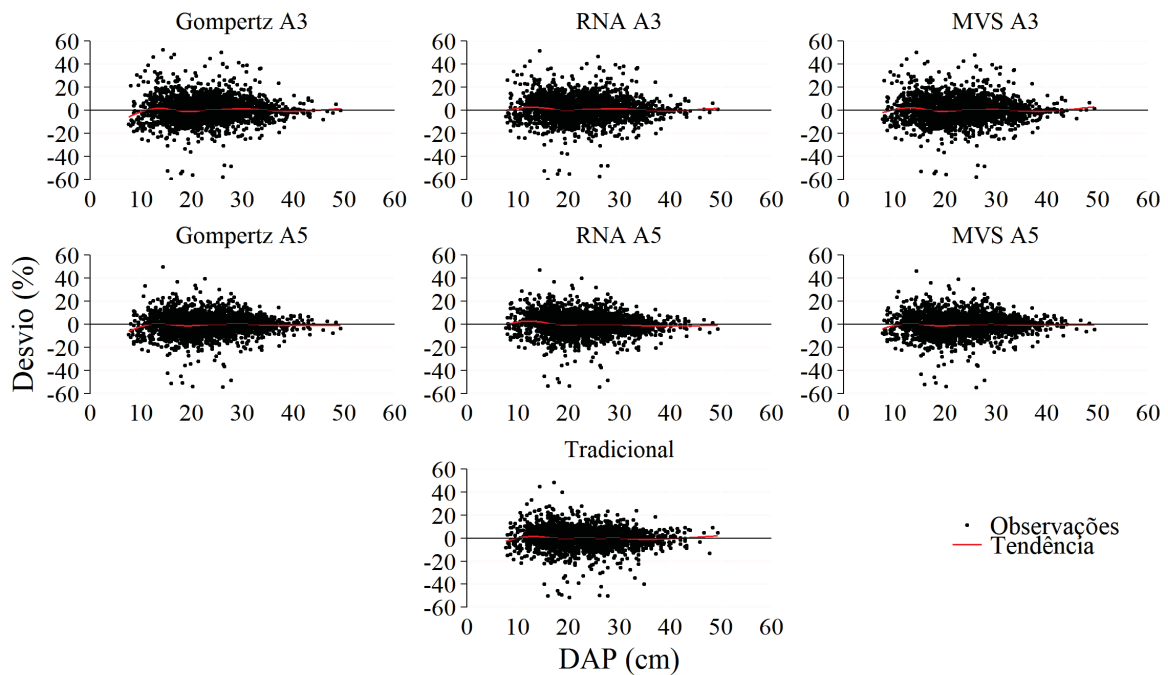
FONTE: O autor (2019).

FIGURA 2.3 - FREQUÊNCIA DE RESÍDUOS DAS PREDIÇÕES REALIZADAS PELOS MÉTODOS ESTUDADOS NO CONJUNTO DE TESTE DE POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ.



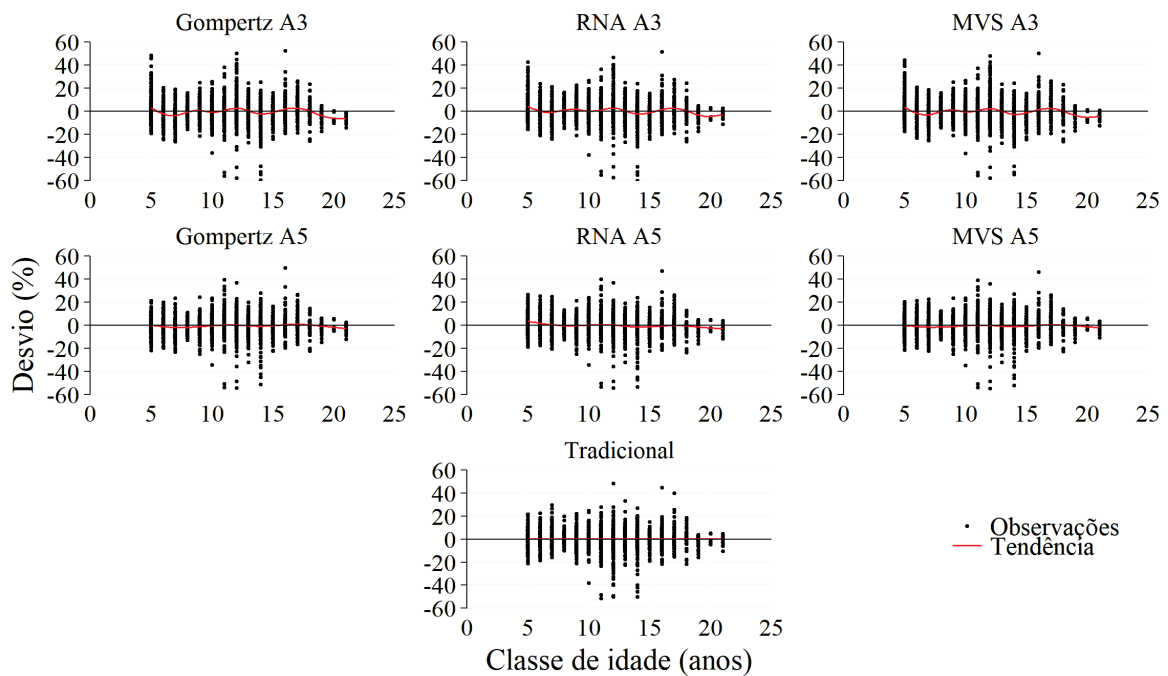
FONTE: O autor (2019).

FIGURA 2.4 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL EM FUNÇÃO DOS DIÂMETROS À ALTURA DO PEITO (DAP) REALIZADAS PELOS MÉTODOS ESTUDADOS, PARA POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ.



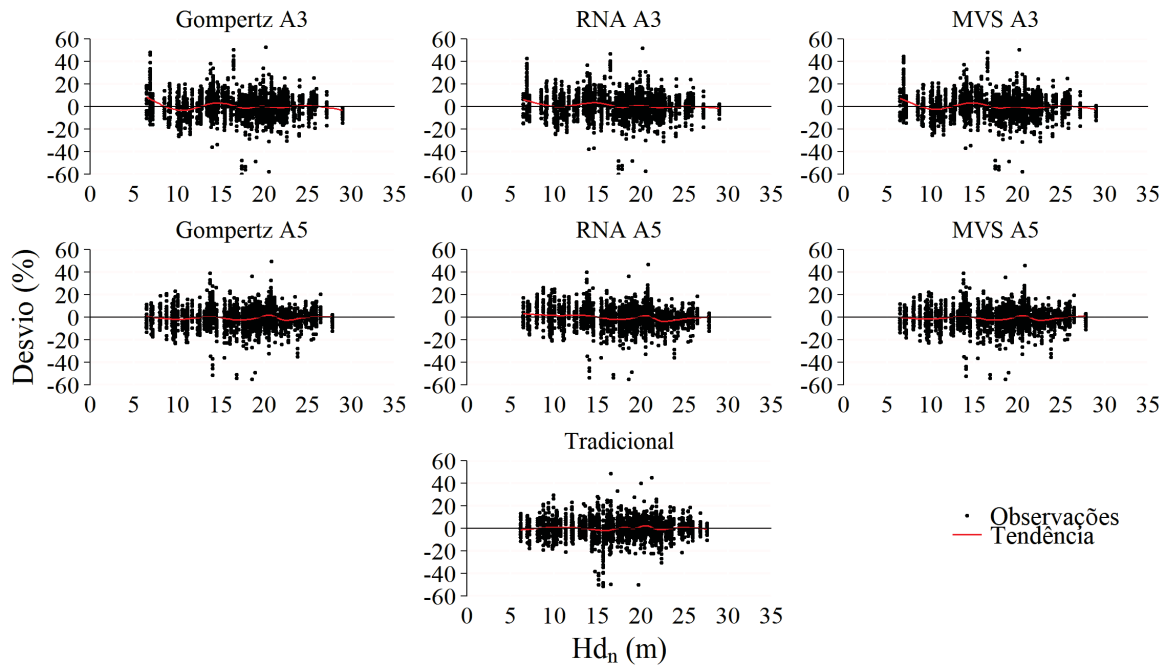
FONTE: O autor (2019).

FIGURA 2.5 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL EM FUNÇÃO DA IDADE, REALIZADA PELOS MÉTODOS ESTUDADOS, PARA POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ.



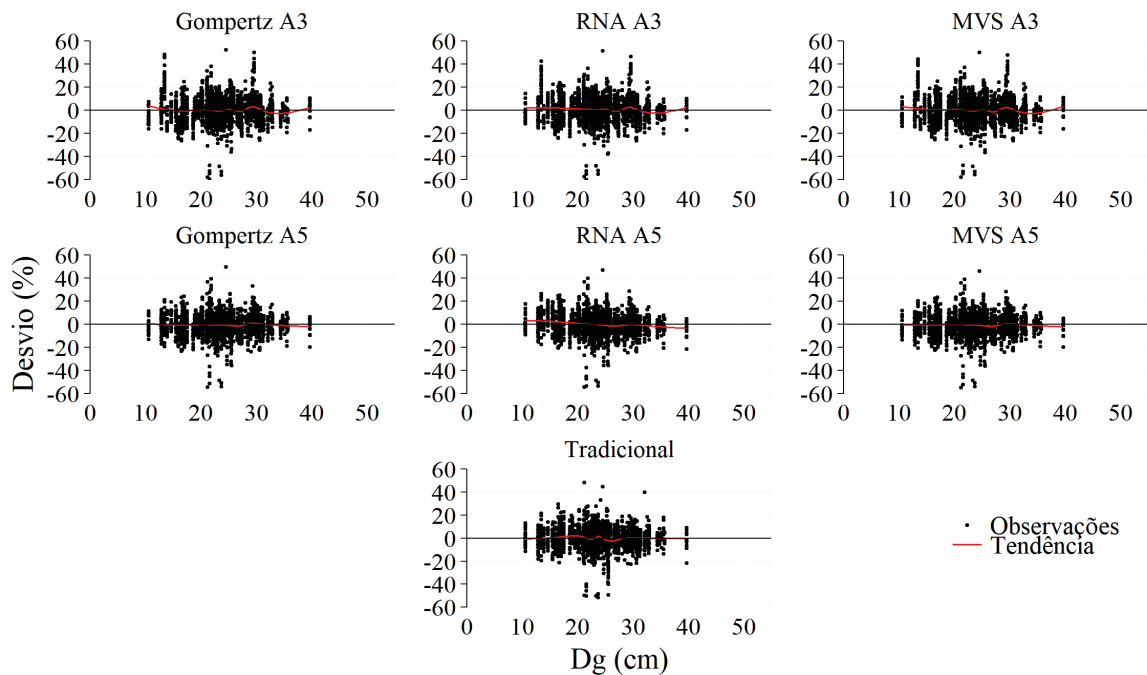
FONTE: O autor (2019).

FIGURA 2.6 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL PARA POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DA VARIÁVEL Hd_n , EM QUE $n = 6$ NO MÉTODO TRADICIONAL, $n = 3$ NO CONJUNTO A3 E $n = 5$ NO CONJUNTO A5.



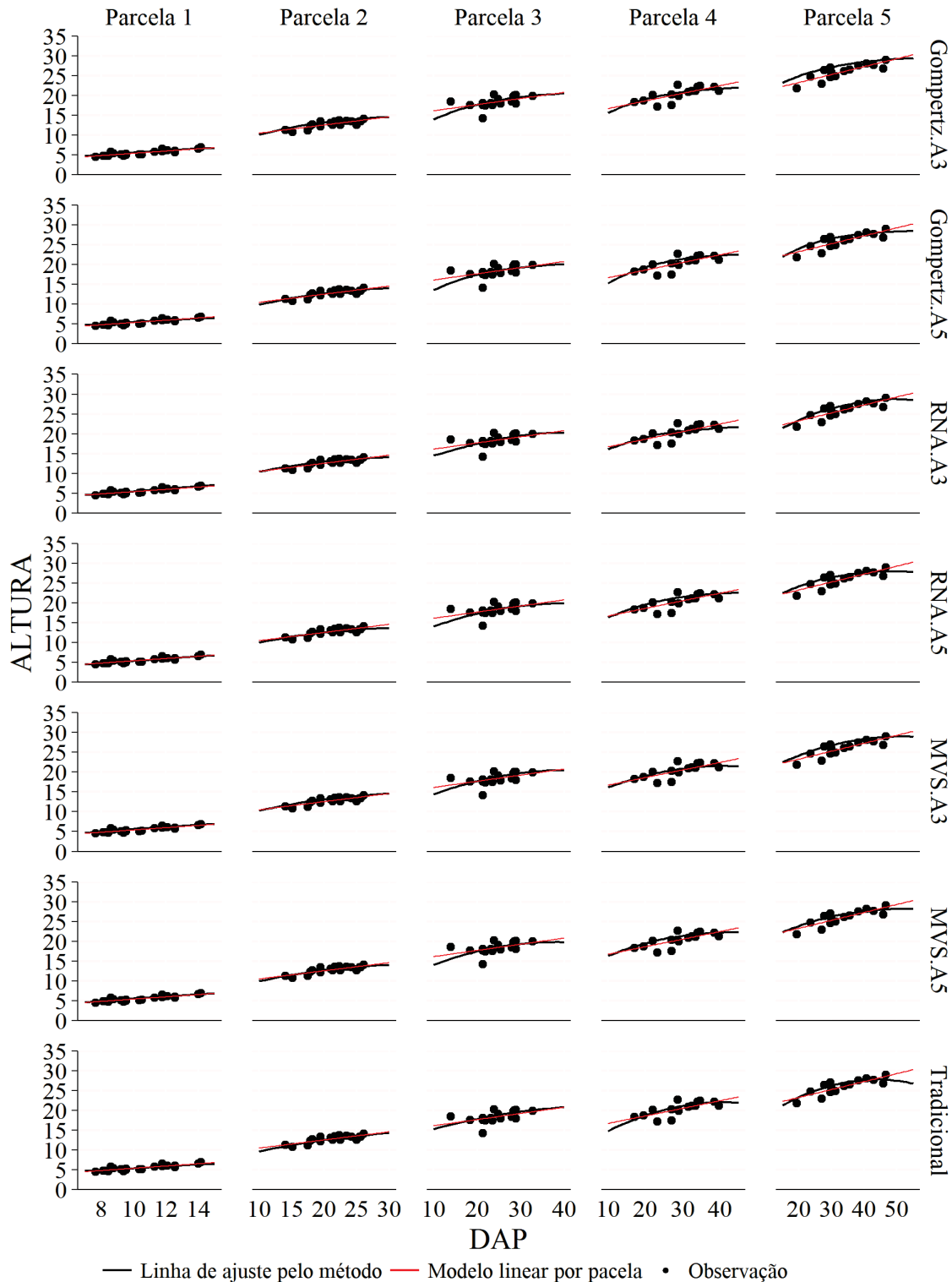
FONTE: O autor (2019).

FIGURA 2.7 - DISPERSÃO DOS RESÍDUOS RELATIVOS DAS PREDIÇÕES DE ALTURA TOTAL PARA POVOAMENTOS DE *Pinus taeda* L. NO PARANÁ REALIZADA PELOS MÉTODOS ESTUDADOS, EM FUNÇÃO DO DIÂMETRO MÉDIO QUADRÁTICO.



FONTE: O autor (2019).

FIGURA 2.8 - ALTURAS OBSERVADAS E ESTIMADAS POR DIFERENTES MÉTODOS PARA *Pinus taeda* L., PARA 5 PARCELAS SELECIONADAS ALEATORIAMENTE NO CONJUNTO DE TESTE.



FONTE: O autor (2019).

LEGENDA: (PARCELA 1: IDADE = 5, h_{dom} = 6,2, hd_1 = 6,9, hd_3 = 6,5, Dg = 10,58; PARCELA 2: IDADE = 9, h_{dom} = 13,4, hd_1 = 14,1, hd_3 = 13,6, Dg = 21,21; PARCELA 3: IDADE = 13, h_{dom} = 19,4, hd_1 = 19,9, hd_3 = 19,3, Dg = 24,82; PARCELA 4: IDADE = 17, h_{dom} = 21,8, hd_1 = 21,2, hd_3 = 21,9, Dg = 29,90; PARCELA 5: IDADE = 21, h_{dom} = 27,6, hd_1 = 29,0, hd_3 = 27,8, Dg = 34,37).

Os resultados para os indicadores estatísticos e análises gráficas apontam que tanto o modelo de Gompertz modificado quanto os algoritmos de aprendizado de máquina apresentaram robustez suficiente para descrever com precisão satisfatória a relação hipsométrica existente em um conjunto de dados desbalanceados e de elevada heterogeneidade.

2.3.5 ANÁLISE DE VARIÂNCIA E TESTE DE COMPARAÇÃO DE MÉDIAS

Por meio da análise de variância foi detectado que ao menos um tratamento diferiu dos demais (Tabela 2.5). Desse modo, procedeu-se o teste de comparação de médias de Dunnett, que indicou que as três técnicas de regressão aplicadas ao conjunto que considera a medição de 5 árvores por parcela produziram erros médios absolutos que não diferem significativamente do método tradicional, a 5% de probabilidade (Tabela 2.6). As médias obtidas considerando a medição de três árvores por unidade amostral diferiram significativamente do método tradicional, atestando a precisão inferior obtida pela aplicação das técnicas ao conjunto A3.

TABELA 2.5 - ANÁLISE DE VARIÂNCIA PARA A DIFERENÇA ABSOLUTA ENTRE AS ALTURAS OBSERVADAS E ESTIMADAS.

Fonte de Variação	GL	SQ	QM	F	F crítico	Valor-p
Blocos	2.375	13.117	5,523	27,3889	1,053	< 0,001
Tratamentos	5	42	8,356	41,436	2,215	< 0,001
Erro	11.875	2.395	0,202			
Total	14.255	15.553				

FONTE: O autor (2019).

TABELA 2.6 - TESTE DE DUNNETT PARA COMPARAÇÃO DAS MÉDIAS DOS TRATAMENTOS COM O CONTROLE.

Hipótese	Média	Valor t	Valor-p
Gompertz (A3) - Tradicional = 0	0,14211	4,626	< 0,001**
RNA (A3) - Tradicional = 0	0,12722	4,141	< 0,001**
MVS (A3) - Tradicional = 0	0,12729	4,144	< 0,001**
Gompertz (A5) - Tradicional = 0	0,02138	0,696	0,959 ^{ns}
RNA (A5) - Tradicional = 0	0,01857	0,604	0,979 ^{ns}
MVS (A5) - Tradicional = 0	0,01324	0,431	0,996 ^{ns}

LEGENDA: ** = Tratamentos que diferem do método tradicional a 5% de probabilidade pelo teste de Dunnett; ns = Tratamentos que não diferem do método tradicional a 5% de probabilidade pelo teste de Dunnett.

FONTE: O autor (2019).

2.3.6 CONSIDERAÇÕES SOBRE A REDUÇÃO DO NÚMERO DE MEDIÇÕES DE ALTURA EM PARCELAS DE INVENTÁRIO FLORESTAL

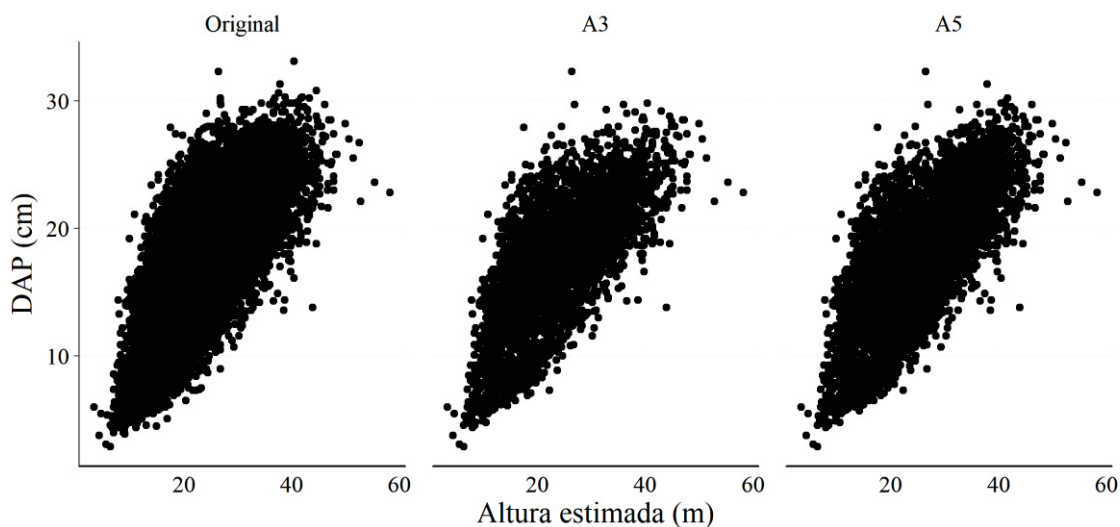
O presente estudo analisou a aplicação de técnicas de regressão com covariantes e AM em conjuntos de dados que simulavam a mensuração de três e de cinco árvores por parcela. A quantidade de árvores medidas em cada unidade amostral demonstrou ser mais decisiva na qualidade dos modelos, uma vez que os diferentes técnicas de regressão apresentaram resultados similares quando aplicadas ao mesmo conjunto de dados, e maiores diferenças foram observadas para uma mesma técnica aplicada a diferentes conjuntos.

Binoti *et al.* (2012) concluíram não haver redução significativa no desempenho da estimativa de alturas por meio de RNA quando se utilizou no treinamento apenas 10% das unidades amostrais de um povoamento de eucaliptos submetidos à medição da altura de 7 árvores/parcela. Os resultados aqui apresentados demonstram que outras técnicas como MVS e modelos não lineares podem também ser utilizados quando se objetiva a redução do número de medições para representação da relação hipsométrica.

Foi demonstrado neste estudo que a mensuração de cinco pares de diâmetro-altura por unidade amostral é suficiente para aplicação de um único modelo para modelagem da relação hipsométrica de povoamentos com idades, espaçamentos e qualidades de sítio diferentes. As três técnicas de regressão testadas são capazes de substituir, sem perda significativa de precisão, 146 modelos lineares selecionados dentre 438 ajustados, uma vez que apresentaram desempenhos que não diferem estatisticamente.

A medição das árvores de maior diâmetro e de uma dentre as de menor na unidade amostral, visa cobrir a amplitude da distribuição de valores de DAP e altura nos povoamentos, enquanto a medição da primeira árvore visa formar a nuvem de observações aleatórias entre os extremos. Apesar da redução no número de medições de altura por parcela, as nuvens de observações dos conjuntos A3 e A5 permaneceram densas, especialmente no conjunto A5, em que a variação presente nos dados permaneceu suficientemente representada, permitindo a geração de modelos de performance satisfatória para representação de todos os povoamentos (Figura 2.9).

FIGURA 2.9 - DISPERSÃO DE OBSERVAÇÕES DOS CONJUNTOS DE AJUSTE/TREINAMENTO UTILIZADOS NO PRESENTE ESTUDO.

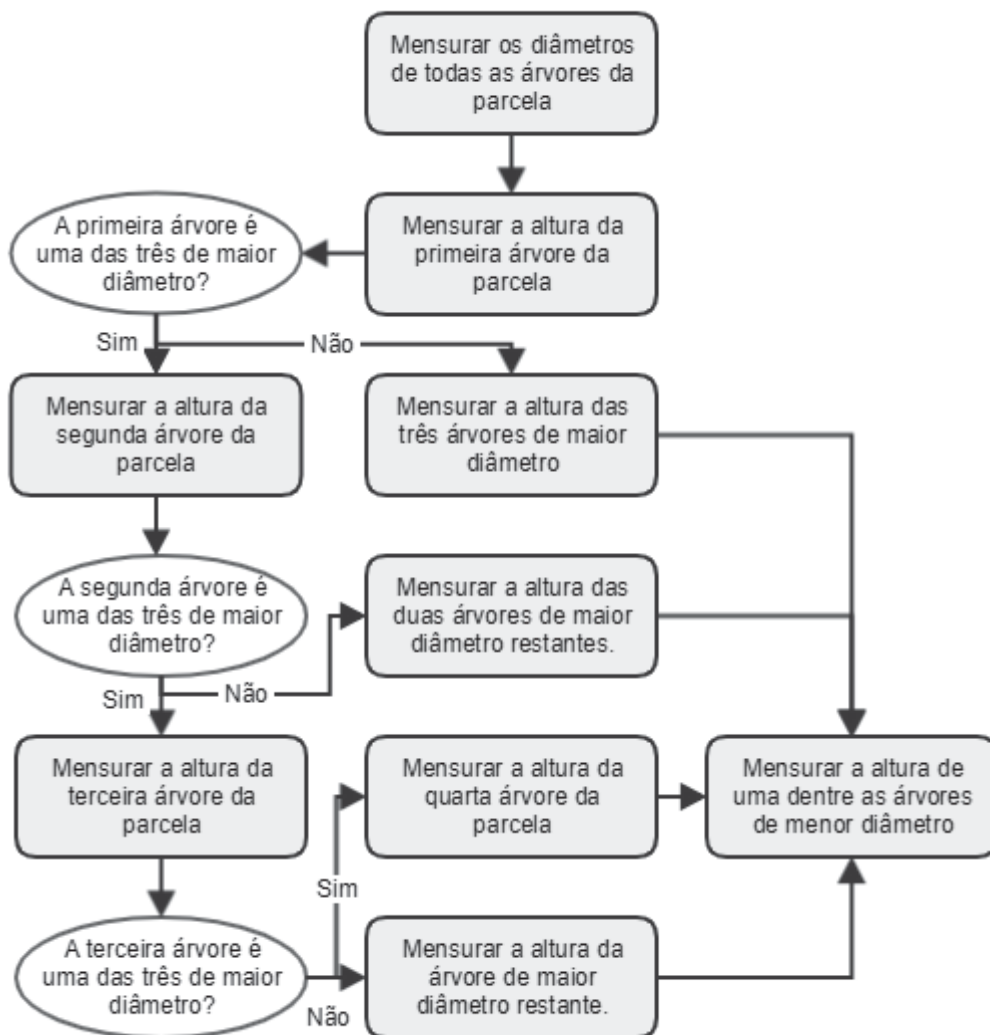


FONTE: O autor (2019).

É importante ressaltar que o método de inventário utilizado pela empresa não contempla a mensuração da árvore de menor diâmetro de cada parcela e que, portanto, o método de simulação da redução no número de árvores coletadas por unidade amostral aplicado neste estudo contemplou a permanência árvore de menor diâmetro dentre as que tiveram a altura mensurada, ou seja, uma dentre as árvores de menor diâmetro da parcela, e não necessariamente a menor.

A medição de uma dentre as árvores de menor diâmetro em vez da árvore de menor diâmetro flexibiliza o método e facilita a operacionalização desta estratégia em campo. A partir disso, a Figura 2.10 ilustra a sequência de procedimentos a serem realizados para execução em campo do método proposto.

FIGURA 2.10 - SEQUÊNCIA DE PROCEDIMENTOS PARA MEDIÇÃO DE ALTURAS DE CINCO ÁRVORES POR UNIDADE AMOSTRAL PELO MÉTODO PROPOSTO NESTE ESTUDO.



FONTE: O autor (2019).

2.4 CONCLUSÕES

- A relação hipsométrica existente em povoamentos de idades, espaçamentos e qualidades de sítio diferentes pode ser representada por uma única equação não linear, Rede Neural Artificial ou Máquina de Vetor de Suporte.
- Redes Neurais Artificiais, Máquinas de Vetores de Suporte e modelos não lineares com adição de covariantes são capazes de substituir uma grande quantidade de equações que representam a relação hipsométrica existente em povoamentos de diferentes características.
- Modelos não lineares com adição de covariantes apresentam maior replicabilidade que Redes Neurais Artificiais e Máquinas de Vetores de Suporte, por serem representados matematicamente por uma única equação.
- A quantidade de alturas a serem mensuradas em inventários de *Pinus taeda* para construção de modelos hipsométricos genéricos pode ser reduzida para cinco árvores por parcela, sem perda de precisão significativa.

2.5 REFERÊNCIAS

BARROS, D. A.; MACHADO, S. A.; ACERBI JÚNIOR, F. W.; SCOLFORO, J. R. S. Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. **Boletim de Pesquisa Florestal**, Colombo, n. 45, p. 03-28, 2002.

BINOTI, D. H. B., BINOTI, M L. M. da S.; LEITE, H. G.; SILVA, A. Redução dos custos em inventário de povoamentos equiâneos. **Revista Brasileira de Ciências Agrárias**, Recife, v.8, n.1, p.125-129, 2013a.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v. 37, n. 4, p. 639-645. 2013b.

CAMPOS, J. C. C; RIBEIRO, J. O.; PAULA NETO, F. **Inventário florestal nacional, reflorestamento**: Minas Gerais. Brasília, DF: IBDF, 1984. 126 p.

COUTO, H. T. Z.; BASTOS, N. L. M. Modelos de equações de volume e relações hipsométricas para plantações de *Eucalyptus* no Estado de São Paulo. **Scientia Forestalis**, n. 60, p.149-163, 1986.

HUSCH, B.; MILLER, C. I.; BEERS, T. W. **Forest Mensuration**. 3 ed. New York: John Wiley & Sons, 1982. 613 p.

LEITE; H. L.; ANDRADE, V. C. L. de. Importância das variáveis altura dominante e altura total em equações hipsométricas e volumétricas. **Revista Árvore**, Viçosa, v. 27, n. 3, p. 301-310, 2003.

SENA, A. L. M; SILVA NETO, A. J. da; OLIVEIRA G. M. V.; CALEGARIO, N. Modelos não lineares com uso de covariantes para relação hipsométrica de duas espécies de *Pinus* topicais. **Ciência Florestal**, Santa Maria, v. 25, n. 4, p. 969-980, 2015.

SHALEV-SHWARTZ, S.; BEM-DAVID, S. **Understanding Machine Learning: From Theory to Algorithms**. New York: Cambridge University Press, 1. ed., 2014. 449 p.

CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

Nesta pesquisa foi confirmada uma tendência que tem sido observada ao comparar métodos de aprendizado de máquina, como Redes Neurais Artificiais e Máquinas de Vetores de Suporte, com modelos estatísticos de regressão. Os métodos baseados em inteligência artificial tendem a apresentar indicadores de qualidade de ajuste levemente mais favoráveis que as técnicas tradicionais. Em contrapartida, a aplicação destes métodos exige maior atenção por parte do analista, uma vez que são aproximadores de funções e não possuem uma forma pré-definida, como ocorre nos modelos lineares e biológicos, o que torna os métodos de aprendizado de máquina mais susceptíveis a erros e mais sensíveis a inconsistências na base de dados.

Isso significa que a utilização de métodos de aprendizado de máquina na mensuração florestal exige maior esforço de avaliação, de modo a identificar comportamentos improváveis nas curvas de predição, uma vez que, diferente dos métodos estatísticos, há grande dificuldade em interpretar a influência das variáveis preditoras e dos parâmetros ajustados.

O emprego de modelos não lineares com inserção de covariantes se mostrou promissor, uma vez que apresentou resultados muito similares às RNA e as MVS, que foram os métodos que apresentaram os melhores resultados no Capítulo 1. A técnica foi capaz de reunir a robustez dos métodos de aprendizado de máquina, com a simplicidade operacional das técnicas de regressão tradicional. Apesar de demandar grande esforço computacional, o ajuste dos modelos não lineares com inserção de covariantes é menos exigente que o treinamento de RNA ou MVS, uma vez que o número de parâmetros a serem otimizados nos modelos estatísticos tende a ser inferior.

A replicação da equação gerada a partir do modelo não linear também é menos custosa, uma vez que uma única equação, representada pela inserção dos valores estimados para os coeficientes no modelo, é utilizada para a realização de novas predições em qualquer *software* ou computador. Por outro lado, a replicação de RNA e MVS é mais custosa, pois exige cálculos de padronização das variáveis de entrada, operação entre matrizes e vetores que armazenam os pesos sinápticos, *bias* e valores de entrada/saída dos neurônios, e despadronização da saída.

Nas RNA quanto maior a quantidade de neurônios, e nas MVS quanto maior a quantidade de vetores de suporte, maiores serão as matrizes e os vetores a serem

reproduzidos. Estes cálculos podem ser realizados de maneira rápida e simples, por meio de poucos procedimentos no *software* R ou em outro ambiente de desenvolvimento como o SAS. No entanto, a reprodução em MS Excel ou mesmo a representação matemática são de grande dificuldade, tornando mais custosa a replicação e representação dos modelos treinados.

A redução do número de medições de altura por parcela se mostrou viável independente das técnicas de regressão aplicadas. O número médio de mensurações por parcela utilizadas para modelagem da relação hipsométrica foi reduzido de 17 para 5, sem perda de qualidade nas predições. A redução da quantidade de árvores a terem a altura mensurada permite um aumento no rendimento de equipes de medição em inventário florestais, especialmente em dias de vento, fator que dificulta a medição das alturas, reduzindo o rendimento da atividade.

Para estudos futuros, recomenda-se a investigação do impacto da redução do número de alturas a serem medidas no tempo de medição de parcelas e nos custos de inventários florestais.

REFERÊNCIAS GERAIS

- ABDOLLAHNEJAD, A. et al. Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. *Forests*, [s.l.], v. 8, n. 2, p.42-60, 2017.
- ALPAYDIN, E. Introduction to machine learning. 3. Ed. Boston: MIT press, 2009. 616p.
- ALVES, J. de A. et al. Equações hipsométricas para *Toona ciliata* com inclusão de covariantes. *Ciência Florestal*, Santa Maria, v. 27, n. 2, p. 581-595, 2017.
- ANDRADE, V. C. L. D.; LEITE, H. G. Hipsometric relationship modeling using data sampled in tree scaling and inventory plots. *Revista Árvore*, v. 35, n. 1, p. 157-164. 2011.
- ARAÚJO, B. H. P. D. Modelagem da altura, volume e afilamento do fuste de *Calycophyllum Spruceanum* Benth. empregando regressão e redes neurais artificiais. Dissertação (Mestrado em Recursos Florestais e Engenharia Florestal) – Instituto Nacional de Pesquisas da Amazônia – INPA, Manaus, 2015.
- BARROS, D. A. et al. Comportamento de modelos hipsométricos tradicionais e genéricos para plantações de *Pinus oocarpa* em diferentes tratamentos. *Boletim de Pesquisa Florestal*, Colombo, n. 45, p. 03-28, 2002.
- BARTOSZECK, A. C. P. S.; et al. Dinâmica da relação hipsométrica em função da idade, do sítio e da densidade inicial de povoamentos de bracatinga da região metropolitana de Curitiba, PR. *Revista Árvore*, v. 28, n. 4, p. 517-533, 2004.
- BATES, D. M.; WATTS, D. G. Nonlinear regression: iterative estimation and linear approximations. In:_. *Nonlinear Regression Analysis and Its Applications*. New York: Wiley, 1988. cap. 2, p. 32-66.
- BEAUDOIN, A. et al. Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. *Canadian Journal of Forest Research*, v. 44, n. 5, p. 521-532, 2014.
- BECKER, P.; MEINZER, F. C.; WULLSCHLEGER, S. D. Hydraulic limitation of tree height: a critique. *Functional Ecology*, v. 14, n. 1, p. 4-11. 2000.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *The Journal of Machine Learning Research*, Montréal, v. 13, n. 1, p.281-305, 2012.
- BHATIA, N. Survey of nearest neighbor techniques. *International Journal of Computer Science and Information Security*, v. 8, n. 2, p. 302-305, 2010.
- BINOTI, D. H. B., BINOTI, M L. M. da S.; LEITE, H. G.; SILVA, A. Redução dos custos em inventário de povoamentos equiâneos. *Revista Brasileira de Ciências Agrárias*, v. 8, n.1, p.125-129, 2013a.

BINOTI, D. H. B.; BINOTI, M. L. M. S.; LEITE, H. G. Configuração de Redes Neurais Artificiais para Estimção do Volume de Árvores. *Revista Ciência da Madeira, Pelotas*, v. 5, n. 1, p.58-67, 2014.

BINOTI, M. Emprego de redes neurais artificiais em mensuracção e manejo florestal. Tese (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa, 2012.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimacção da altura de povoamentos equiâneos de eucalipto. *Revista Árvore*, v. 37, n. 4, p. 639-645. 2013b.

BREIMAN, L. Random Forests. *Machine Learning*, [s.l.], v. 45, n. 1, p.05-32, 2001.

BREIMAN, L. Stacked regressions. *Machine learning*, v. 24, n. 1, p. 49-64, 1996.

BREUSCH, T.; PAGAN, A. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, Menasha, v.47, p.1287–1294, 1979.

CALEGARIO, N. et al. Melhoria da qualidade de ajuste de modelos biométricos florestais pelo emprego da teoria dos modelos não-lineares generalizados. *Scientia Forestalis*, v. 69, p. 38-50, dez., 2005.

CAMPOS, B. P. F. et al. Predicção da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. *Pesquisa Florestal Brasileira*, v. 36, n. 88, p. 375-385, 2016.

CAMPOS, J. C. C.; LEITE, H. G. Mensuracção florestal: perguntas e respostas. 3. ed. Viçosa, MG: UFV, 2009. 548 p.

CAMPOS, J. C. C; RIBEIRO, J. O.; PAULA NETO, F. Inventário florestal nacional, reflorestamento: Minas Gerais. Brasília, DF: IBDF, 1984. 126 p.

CANDEL, A. et al. Deep Learning with H2O. Disponível em: <<http://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/DeepLearningBooklet.pdf>>. Acesso em: 02 jul. 2017.

CHEN, X.; ISHWARAN, H. Random forests for genomic data analysis. *Genomics*, v. 99, n. 6, 323-329, 2012.

CHIRICI, G. et al. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. *Remote Sensing of Environment*, v. 176, p. 282-294, 2016.

CHOJACZYK, A. A.; TEIXEIRA, A. P.; NEVES, L. C.; CARDOSO, J. B.; SOARES, C. G. Review and application of artificial neural networks models in reliability analysis of steel structures. *Structural Safety*, v. 52, p. 78-89, 2015.

- CORDEIRO, M. A., et al. Estimativa do volume de *Acacia mangium* utilizando técnicas de redes neurais artificiais e máquinas vetor de suporte. Pesquisa Florestal Brasileira, Colombo, v. 35, n. 83, 255-261, 2015.
- CORTES, C; VAPNIK, V. N. Support-Vector Networks. Machine Learning. 2 ed. New York: Springer, 1995, 314 p.
- COUTO, H. T. Z.; BASTOS, N. L. M. Modelos de equações de volume e relações hipsométricas para plantações de *Eucalyptus* no Estado de São Paulo. Scientia Forestalis, n. 60, p.149-163, 1986.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. An introduction to support vector machines and other kernel-based learning methods. New York: Cambridge University Press, 2000, 189 p.
- CUTLER, D. R.; EDWARDS JR, T. C.; BEARD, K. H.; CUTLER, A.; HESS, K. T.; GIBSON, J.; LAWLER, J. J. Random forests for classification in ecology. Ecology, v. 88, n. 11, p. 2783-2792, 2007.
- DIAMANTOPOULOU, M. J.; MILIOS, E. Modelling total volume of dominant pine trees in reforestations via multivariate analysis and artificial neural network models. Biosystems engineering, v. 105, n. 3, p. 306-315, 2010.
- DIETTERICH, T. G. Ensemble methods in machine learning. In: International workshop on multiple classifier systems, 2010. Multiple Classifier Systems. Berlin: Springer, p. 1-15, 2010.
- EISFELD, R.L.; SANQUETTA, C. R.; ARCE, J. E.; MAESTRI, R.; WEBER, K. S. Modelagem do crescimento e da produção de *Pinus taeda* L. por meio de função probabilística. Floresta, v. 35, n. 2, p. 167-177. 2005.
- FACELI, K. et al. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC, 2011. 378p.
- FIGUEIREDO FILHO, A.; DIAS, A. N.; KOHLER, S. V.; VERUSSA, A. A.; CHIQUETTO, A. L. Evolution of the hypsometric relationship in *Araucaria angustifolia* plantations in the mid-south region of Paraná state. Cerne, v. 16, n. 3, p. 347-357. 2010.
- GARDNER, M. W.; DORLING, S. R. Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. Atmospheric environment, v. 32, n. 14-15, p. 2627-2636, 1998.
- GLEASON, C. J.; IM, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. Remote Sensing of Environment, v. 125, p. 80-91, 2012.
- GORGENS, E. B. et al. Influência da arquitetura na estimativa de volume de árvores individuais por meio de redes neurais artificiais. Revista Árvore, Viçosa, v. 38, n. 2, p.289-295, abr. 2014.

GORGENS, E. B.; MONTAGHI, A.; RODRIGUEZ, L. C. E. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. *Computers and Electronics in Agriculture*, v. 116, p. 221-227, 2015.

GUJARATI, D. *Econometria básica*. São Paulo: Elsevier; 2006.

HAARA, A.; MALTAMO, M.; TOKOLA, T. The K-nearest-neighbour method for estimating basal-area diameter distribution. *Scandinavian Journal of Forest Research*, v. 12, n. 2, p. 200-208, 1997.

HANSEN, L. K.; SALAMON, P. Neural network ensembles. *Transactions on pattern analysis and machine intelligence*, v. 12 n. 10, p. 993-1001, 1990.

HEARST, M. A. et al. Support vector machines. *IEEE Intelligent Systems*, IEEE Computer Society, v. 13, n. 4, p. 18–28, 1998.

HECHENBICHLER, K; SCHLIEP, K. Weighted k-Nearest-Neighbor Techniques and Ordinal Classification. Discussion paper 399, Ludwig-Maximilians University Munich, Munich, 2004. Disponível em: < https://epub.ub.uni-muenchen.de/1769/1/paper_399.pdf>. Acesso em: 20 jun. 2017.

HENRY, H. A. L.; AARSSSEN, L. W. The interpretation of stem diameter–height allometry in trees: biomechanical constraints, neighbour effects, or biased regressions? *Ecology letters*, v. 2, n. 2, p. 89-97. 1999.

HSU, C. W; CHANG, C. C; LIN, C. J. A Practical Guide to Support Vector Classification. Technical report, Department of Computer Science, National Taiwan University, 2010.

HUANG, C. L.; WANG, C. J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with applications*, v. 31, n. 2, p. 231-240, 2006.

HUSCH, B.; MILLER, C. I.; BEERS, T. W. *Forest Mensuration*. 3 ed. New York: John Wiley & Sons, 1982. 613p.

KUDUS, K. A.; AHMAD, M. I.; LAPONGAN, J. Nonlinear regression approach to estimating Johnson SB parameters for diameter data. *Canadian Journal of Forest Research*, v. 29, n. 3, p. 310-314. 1999.

LARRANAGA, P. et al. Machine learning in bioinformatics. *Briefings in bioinformatics*, v. 7, n.1, p. 86-112, 2006.

LATIFI, H.; NOTHDURFT, A.; KOCH, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry*, v. 83 n. 4, p. 395-407, 2010.

LEITE, H. G.; ANDRADE, V. D. Importância das variáveis altura dominante e altura total em equações hipsométricas e volumétricas. *Revista Árvore*, v. 27, n. 3, p. 301-310. 2003.

LORENA, A. C.; CARVALHO, A. C. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43-67, 2007.

MACHADO, S. A.; DOS SANTOS, A. A. P.; DO NASCIMENTO, R. G. M.; AUGUSTYNCZIK, A. L. D.; ZAMIN, N. T. Modelagem da distribuição diamétrica de quatro espécies de Lauraceae em um fragmento de Floresta Ombrófila Mista. *RECEN-Revista Ciências Exatas e Naturais*, v. 12, n. 1, p. 91-105, 2010.

MACHADO, S. A.; NASCIMENTO, R. G. M.; AUGUSTYNCZIK, A. L. D.; DA SILVA SILVA, L. C. R.; FIGURA, M. A.; PEREIRA, E. M.; TÉO, S. J. Comportamento da relação hipsométrica de *Araucaria angustifolia* no capão da Engenharia Florestal da UFPR. *Pesquisa Florestal Brasileira*, n. 56, p. 5, 2008.

MARCON, F.; TÉO, S. J.; SCHNEIDER, C. R.; DOS SANTOS, F. B.; CHIARELLO, K. M. A.; FIORENTIN, L. D. Modelagem da distribuição diamétrica de um fragmento de Floresta Ombrófila Mista em Lebon Régis, SC. *Floresta*, v. 45, n. 2, p. 337-348. 2014.

MARTINS, E. R. et al. Configuração de redes neurais artificiais para estimação da altura total de árvores de eucalipto. *Revista Brasileira de Ciências Agrárias (Agrária)*, Recife, v. 11, n. 2, p. 117-123, 2016.

MAZUCHELI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. *Acta Scientiarum*, v. 24, n. 6, p. 1761-1770. 2002.

MCCULLAGH, A.; BLACK, K.; NIEUWENHUIS, M. Evaluation of tree and stand-level growth models using national forest inventory data. *European journal of forest research*, v. 136, n. 2, p. 251-258, 2017.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, v. 5, n. 4, p. 115-133, 1943.

MENDONÇA, A. R.; DA SILVA, J. C.; AOZAI, T. S.; DA SILVA, E. R.; SANTOS, J. S.; BINOTI, D. H. B.; DA SILVA, G. F. Estimação da altura total de árvores de ipê felpudo utilizando modelos de regressão e redes neurais. *Revista Brasileira de Biometria*, v. 36, n. 1, p. 128-139. 2018.

MENDONÇA, A. R.; DE PÁDUA CHAVES, S.; CARVALHO, N. C. Modelos hipsométricos generalizados mistos na predição da altura de *Eucalyptus* sp. *Cerne*, v. 21, n. 1, p. 107-115. 2015.

MEZA, A. M. A.; SANTACOLOMA, G. D.; MEDINA, C. D. A.; DOMÍNGUEZ, G. C. Parameter selection in least squares-support vector machines regression oriented, using generalized cross-validation. *Dyna*, v. 79, n. 171, p. 23-30, 2012.

MIGUEL, E. P.; REZENDE, A. V.; LEAL, F. A.; MATRICARDI, E. A. T.; VALE, A. T. D.; PEREIRA, R. S. Artificial neural networks for modeling wood volume and aboveground biomass of tall Cerrado using satellite data. *Pesquisa Agropecuária Brasileira*, v. 50, n. 9, p. 829-839, 2015.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*. Editora Manole Ltda. v. 1, n. 1, p. 32, 2003.

MONTAÑO, R. A. N. R. Aplicação de técnicas de aprendizado de máquina na mensuração florestal. Tese (doutorado). Universidade Federal do Paraná, 2016.

NIKLAS, K. J. The scaling of plant height: a comparison among major plant clades and anatomical grades. *Annals of botany*, v. 72, n. 2, p. 165-172, 1993.

ÖZÇELİK, R. et al. Estimating Crimean juniper tree height using nonlinear regression and artificial neural network models. *Forest Ecology And Management*, [s.l.], v. 306, p.52-60, out. 2013.

PALAHÍ, M.; PUKKALA, T.; BLASCO, E.; TRASOBARES, A. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). *European Journal of Forest Research*, v. 126, n. 4, p. 563-571. 2007.

PAYANDEH, B. Choosing regression models for biomass prediction equations. *The Forestry Chronicle*, v. 57, n. 5, p. 229-232. 1981.

PAYANDEH, B. Some applications of nonlinear regression models in forestry research. *The Forestry Chronicle*, v. 59, n. 5, p. 244-248. 1983.

PEERBHAY, K. Y.; MUTANGA, O.; ISMAIL, R. Random forests unsupervised classification: The detection and mapping of *Solanum mauritianum* infestations in plantation forestry using hyperspectral data. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 8 n. 6, p. 3107-3122, 2015.

RAMOS, E. G.; MARTÍNEZ, F. V. A Review of Artificial Neural Networks: How Well Do They Perform in Forecasting Time Series. *Analítica: revista de análisis estadístico*, v. 6, p. 7-18, 2013.

REESE, H.; GRANQVIST-PAHLÉN, T.; EGBERTH, M.; NILSSON, M.; OLSSON, H. Automated estimation of forest parameters for Sweden using Landsat data and the kNN algorithm. In: 31st International Symposium on Remote Sensing of Environment, 2005.

REIS, L. P. et al. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. *Ecological Engineering*, v. 112, p. 140-147, 2018.

RETSLAFF, F. A. S.; FIGUEIREDO FILHO, A.; DIAS, A. N.; BERNETT, L. G.; FIGURA, M. A. Curvas de sítio e relações hipsométricas para *Eucalyptus grandis* na região dos Campos Gerais, Paraná. *Cerne*, v. 21, n. 2, p. 219-225. 2015.

REZENDE, S. O. Sistemas inteligentes: fundamentos e aplicações. Editora Manole Ltda, 2003.

RIBEIRO, A., FERRAZ FILHO, A. C., MELLO, J. M. D., ZÉLIA FERREIRA, M., MARTINS LISBOA, P. M., & SOARES SCOLFORO, J. R. Estratégias e metodologias de ajuste de modelos hipsométricos em plantios de *Eucalyptus* sp. *Cerne*, v. 16, n. 1, p. 22-31. 2010.

RICHARDS, F. J. A flexible growth function for empirical use. *Journal of experimental Botany*, v. 10, n. 2, p. 290-301. 1959.

RICKEN, P.; MATTOS, P. P.; BRAZ, E.; KALIL FILHO, A. N. Crescimento diamétrico de povoamento de *Toona ciliata* var. *australis* em Adrianópolis, PR. Paraná: Embrapa Florestas, 2011. Relatório técnico.

ROBINSON, A. P.; LANE, S. E.; THÉRIEN, G. Fitting forestry models using generalized additive models: a taper model example. *Canadian Journal of Forest Research*, [s.l.], v. 41, n. 10, p.1909-1916, out. 2011.

RUMELHART, D.E.; HINTON, G.E.; WILLIAMS, R.J. Learning internal representations by backpropagating errors. *Nature*, v. 323, p. 533–536, 1986.

SANQUETTA, C. R. et al. On the use of data mining for estimating carbon storage in the trees. *Carbon Balance And Management*, [s.l.], v. 8, n. 1, p.6-14, 2013.

SANQUETTA, C. R.; DALLA CORTE, A. P.; ROGL, A.; PIMENTEL, A. Relações diâmetro-altura para espécies lenhosas em um fragmento de floresta Ombrófila Mista no Sul do Paraná. *Série Botânica.*, v. 68, n. 1, p. 103-114. 2013.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural networks*, v. 61, p. 85-117, 2015.

SCHNEIDER, P. R. Manejo florestal I – Tópicos para o planejamento da produção florestal. Santa Maria: Universidade Federal de Santa Maria, 1986. 289p.

SCOLFORO, J. R. S. Biometria florestal: parte I: modelos de regressão linear e não linear: parte II: modelos para relação hipsométrica, volume, afilamento e peso da matéria seca. Lavras: UFLA/FAEPE, 2005. 352 p.

SENA, A. L. M. et al. Modelos não lineares com uso de covariantes para relação hipsométrica de duas espécies de *Pinus* topicais. *Ciência Florestal*, Santa Maria, v. 25, n. 4, p. 969-980, 2015.

SHALEV-SHWARTZ, S.; BEM-DAVID, S. Understanding Machine Learning: From Theory to Algorithms. New York: Cambridge University Press, 1. ed., 2014. 449 p.

- SILVA, C. A. et al. Estimating Stand Height and Tree Density in *Pinus taeda* plantations using in-situ data, airborne LiDAR and k-Nearest Neighbor Imputation. *Anais da Academia Brasileira de Ciências*, v. 90, n. 1, p. 295-309, 2018.
- SILVA, G. F. da.; XAVIER, A. C.; RODRIGUES, F. L.; PETERNELLI, L. A. Análise da influencia de diferentes tamanhos e composições de amostras no ajuste de uma relação hipsométrica. *Revista Árvore*, v.31, n.4, p.685-694, 2007.
- SILVA, M. L. M.; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Adjustment of the Schumacher and Hall model and application of artificial neural networks to estimate volume of eucalypt trees. *Revista Árvore*, v. 33, n. 6, 1133-1139, 2009.
- SMOLA, A. Regression Estimation with Support Vector Learning Machines. 1996. 78 f. Master's Thesis – Physics, Technische Universität München, München, 1996.
- SOUZA, C. R. Kernel Functions for Machine Learning Applications. César Souza. 2010. Disponível em: <<http://crsouza.com/2010/03/17/kernel-functions-for-machine-learning-applications/>>. Acesso em: 7 nov. 2018.
- SOUZA, G. S. Introdução aos modelos de regressão linear e não-linear. Brasília: Embrapa-SPI/Embrapa-SEA, 1998.
- TOMMOLA, M., et al. Estimating the characteristics of a marked stand using k-nearest-neighbour regression. *Journal of Forest Engineering*, [s.l.], v. 10, p.75-81, 1999.
- TONINI, H.; COSTA, M. C. G.; SCWENGBER, L. A. M. Crescimento da teca (*Tectona grandis*) em reflorestamento na Amazônia Setentrional. *Pesquisa Florestal Brasileira*, n. 59, p. 05, 2009.
- VAPNIK, V N. The nature of statistical learning theory. 2. ed. New York: Springer-Verlag, 2000. 314 p.
- VENDRUSCOLO, D. G. S. et al. Estimativa da altura de árvores de *Tectona grandis* Lf utilizando regressão e redes neurais artificiais. *Nativa*, Sinop, v. 5, n. 1, p. 52-58. 2017.
- VENDRUSCOLO, D. G. S. et al. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. *Revista Brasileira de Biometria*, v. 33, n. 4, p. 556-569. 2015.
- VENDRUSCOLO, D. G. S. et al. Height prediction of *Tectona grandis* trees by mixed effects modelling and artificial neural networks. *International Journal of Current Research*, [s.l.], v. 8, n. 12, p.43189-43195, 2016.
- WOJCIECHOWSKI, J. JCarbon – Software na Web com Data Mining para estimativas de volume, biomassa e carbono em florestas. Tese (Doutorado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba, 2015.

YU, X.; HYYPPÄ, J.; VASTARANTA, M.; HOLOPAINEN, M.; VIITALA, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *Journal of Photogrammetry and remote sensing*, v. 66, n. 1, p. 28-37, 2011.

ZHU, J. J.; FAN, Z. P.; ZENG, D. H.; JIANG, F. Q.; MATSUZAKI, T. Comparison of stand structure and growth between artificial and natural forests of *Pinus sylvestris* var. *mongolica* on sandy land. *Journal of Forestry Research*, v. 14, n. 2, p. 103-111. 2003.

APÊNDICE 1 – COEFICIENTES DOS MODELOS SELECIONADOS NO AJUSTE PELO MÉTODO TRADICIONAL (POR ESTRATO).

# Estrato	Modelo*	Coeficientes ajustados				Syx	Syx%
		b0	b1	b2	b3		
9	2	-2,3103	0,0960	0,9461	0	0,61	2,72
10	2	-4,0664	0,0681	1,0488	0	0,85	3,39
11	3	2,1246	0,2518	0	0	0,07	0,34
12	1	-22,7870	1,4584	-0,0186	0,7641	1,31	6,33
13	1	1,6214	0,3814	-0,0035	0,5201	1,19	5,81
15	1	-10,1211	0,6072	-0,0076	0,9290	1,16	5,18
16	1	-11,4573	0,4998	-0,0055	1,0275	0,81	3,80
17	2	5,7766	0,0863	0,5796	0	0,79	3,71
19	2	18,8893	0,1808	0	0	1,17	4,77
20	1	-7,4922	0,5668	-0,0073	0,8432	0,84	4,33
21	1	-4,8774	0,3695	-0,0040	0,8605	1,21	5,50
22	3	0,5828	0,6699	0	0	0,15	0,98
23	1	-7,4641	0,5858	-0,0064	0,7712	0,97	5,33
24	1	33,8248	-1,3340	0,0270	0	1,07	6,03
26	1	-6,5870	0,6093	-0,0087	0,8070	0,77	4,38
27	2	-0,7028	0,1930	0,7425	0	1,02	5,50
28	1	-8,8776	0,6279	-0,0084	0,8835	0,79	4,26
30	1	-10,2636	0,5768	-0,0072	1,0089	1,08	6,49
31	2	-6,1624	0,2077	1,0219	0	0,78	4,79
32	1	23,5292	-0,1834	0,0112	-0,4196	0,91	5,14
33	2	-0,9781	0,2226	0,6936	0	0,77	4,65
34	1	-4,0579	0,6031	-0,0087	0,6163	1,03	6,88
35	2	-2,0656	0,2313	0,7452	0	0,54	3,66
36	2	-5,2642	0,1667	1,0474	0	1,17	8,62
37	3	1,9304	0,2944	0	0	0,04	0,25
38	3	1,8961	0,2432	0	0	0,04	0,32
39	1	-5,0991	0,4945	-0,0072	0,8104	0,78	6,04
40	1	-5,4411	0,4521	-0,0055	0,8317	0,55	4,31
41	2	-3,8816	0,1908	0,9464	0	0,82	6,37
42	2	-2,1144	0,2405	0,7358	0	0,65	5,42
43	1	-1,9541	0,6309	-0,0096	0,4224	0,68	5,60
44	1	-22,2525	0,6845	-0,0139	2,3180	0,42	4,18
45	1	-3,3477	0,5760	-0,0091	0,5769	0,62	5,99
46	1	-2,4315	0,3361	-0,0032	0,6949	0,51	5,77
47	1	-2,6168	0,4680	-0,0080	0,6559	0,51	5,53
49	1	-32,2619	2,3460	-0,0369	0,7736	0,63	3,09
50	2	-2,8939	0,2482	0,7831	0	0,86	4,35
51	1	-2,5369	0,4678	-0,0079	0,6186	0,46	6,10
54	2	7,8889	0,1211	0,5267	0	1,05	4,15
55	1	-7,3893	0,6570	-0,0089	0,7799	0,99	5,70
56	3	2,3561	0,1688	0	0	0,05	0,27
57	2	5,7436	0,0980	0,5744	0	0,52	2,43
59	1	-6,1870	0,6418	-0,0076	0,6814	0,85	3,97
60	1	-4,9101	0,5232	-0,0067	0,7632	0,67	3,14
62	1	-8,7363	0,6113	-0,0086	0,9024	1,14	6,09
63	1	-15,4382	0,9927	-0,0139	0,8873	0,66	3,34
64	2	-2,5906	0,2201	0,7699	0	0,64	5,00
65	2	-1,0007	0,1771	0,7831	0	0,63	3,31
66	1	-4,0935	0,3623	-0,0033	0,7741	0,64	3,48

# Estrato	Modelo*	Coeficientes ajustados				Syx	Syx%
		b0	b1	b2	b3		
67	3	2,4156	0,1705	0	0	0,05	0,23
68	1	-8,0266	0,5253	-0,0067	0,9087	0,81	4,52
69	1	-4,5602	0,4576	-0,0056	0,7847	0,77	4,34
70	1	-12,7282	0,7954	-0,0114	0,9661	0,58	3,57
71	1	-9,8513	0,7350	-0,0120	0,9270	0,69	4,33
72	1	-9,7813	0,6188	-0,0100	1,0191	0,60	3,87
73	2	-5,5738	0,2339	0,9325	0	0,52	3,58
74	1	-3,8520	0,4728	-0,0060	0,7310	0,81	5,76
75	1	-4,5951	0,7609	-0,0126	0,5302	0,64	4,69
76	2	0,4780	0,2383	0,4891	0	0,72	6,36
77	2	-2,3161	0,2559	0,6994	0	0,50	6,17
78	1	-3,9285	0,7838	-0,0185	0,5121	0,45	6,16
83	1	-2,0131	1,4759	-0,0207	0	0,83	3,53
84	2	-15,3228	0,1509	1,4199	0	0,99	4,44
85	1	-5,2219	0,6903	-0,0099	0,6872	1,08	5,07
86	1	-4,9170	0,7467	-0,0109	0,6867	1,14	4,92
87	1	-8,6341	0,6201	-0,0078	0,8619	1,11	5,51
88	1	-9,8159	0,5187	-0,0055	0,9695	0,99	5,05
89	1	-5,1256	0,4680	-0,0059	0,8069	0,83	4,97
90	1	-4,6742	0,3944	-0,0032	0,7841	1,05	6,84
91	2	-2,2646	0,2121	0,7892	0	0,90	6,14
92	1	-7,0316	0,5855	-0,0101	0,8902	0,55	5,69
93	2	-0,9553	0,2306	0,6592	0	0,44	6,24
96	2	-3,1066	0,2175	0,8352	0	1,08	4,96
97	1	-5,9639	0,6222	-0,0098	0,7943	0,68	3,91
98	2	-1,0290	0,2029	0,6911	0	1,29	9,38
101	1	6,1063	0,6074	-0,0070	0,2529	1,41	5,72
103	1	-7,7064	0,7538	-0,0107	0,7597	0,88	3,94
104	1	-7,4924	0,3516	-0,0035	0,9714	1,19	5,73
105	1	-56,2247	1,7558	-0,0240	1,9692	0,80	3,35
106	1	-10,1456	0,6289	-0,0077	0,8950	1,18	5,45
107	1	-6,6290	0,3854	-0,0034	0,8893	1,36	6,21
108	2	10,7949	-0,0808	0,6560	0	1,29	5,59
109	1	-8,2525	0,4314	-0,0044	0,9556	1,32	6,14
110	1	-4,4694	0,3424	-0,0037	0,8598	1,13	5,66
111	2	0,1851	0,0966	0,8385	0	0,68	3,43
112	2	-3,4213	0,1722	0,8772	0	0,84	3,93
113	3	2,1951	0,2323	0	0	0,04	0,21
114	1	3,9161	1,0187	-0,0164	0	0,42	2,16
115	1	5,8703	-0,1727	0,0071	0,6479	0,93	4,91
116	2	-3,3031	0,1415	0,9438	0	1,10	6,38
117	2	-22,1458	0,2597	1,8268	0	0,95	5,93
118	2	-10,3976	0,2440	1,2680	0	0,84	5,42
119	2	129,5321	0,2392	-9,5935	0	0,76	6,34
120	1	-8,1043	0,3365	-0,0040	1,1408	0,58	4,24
121	1	-5,7986	0,6158	-0,0102	0,7573	0,75	6,44
122	1	-1,4340	0,6780	-0,0127	0,3114	0,48	5,06
123	1	-2,8762	0,5387	-0,0093	0,5850	0,65	7,68
124	2	-6,8818	0,1059	1,1021	0	0,96	4,01
125	1	-5,7018	0,3997	-0,0054	0,9190	1,02	5,34
133	1	1,4018	0,5045	-0,0055	0,5486	1,20	4,32
135	1	-11,1082	0,6143	-0,0091	1,0347	0,96	3,85

# Estrato	Modelo*	Coeficientes ajustados				Syx	Syx%
		b0	b1	b2	b3		
136	1	-5,2061	0,4687	-0,0051	0,7825	1,02	4,41
137	1	-9,7185	0,6199	-0,0086	0,9429	1,02	4,47
138	1	-9,9216	0,4016	-0,0048	1,0855	0,86	3,91
139	1	-9,1047	0,6169	-0,0081	0,9022	0,87	3,73
140	1	-4,8092	0,5749	-0,0071	0,7380	1,02	4,20
141	1	-4,1559	0,3478	-0,0035	0,8332	0,86	3,97
143	2	-5,7456	0,2194	0,9238	0	1,04	4,69
144	2	-6,4167	0,3097	0,8801	0	1,05	4,72
145	2	1,1897	0,0998	0,8011	0	0,89	3,84
146	1	-15,7661	0,6200	-0,0078	1,1737	0,88	3,96
147	1	-13,7828	0,5844	-0,0078	1,1456	0,76	3,63
148	1	-1,3970	0,7559	-0,0075	0,3281	2,21	11,21
149	1	-11,6113	0,9391	-0,0145	0,8233	0,63	3,48
150	1	-6,6198	0,5273	-0,0068	0,8372	0,83	4,69
151	1	-6,3242	0,7569	-0,0111	0,6773	0,93	5,41
152	1	-11,2306	0,5069	-0,0072	1,1685	0,62	3,99
153	1	-9,6017	0,7483	-0,0120	0,8929	0,70	5,08
154	1	-10,5638	0,7799	-0,0138	0,9785	0,64	4,29
155	3	1,5151	0,3446	0	0	0,07	0,56
156	1	-1,9556	-0,0773	0,0051	0,9832	1,02	5,37
157	1	1,7435	0,6192	-0,0094	0,3454	0,91	5,43
158	1	-137,8650	1,0820	-0,0167	6,1716	2,18	10,40
160	1	35,0616	-0,4875	0,0150	-0,9037	1,09	6,56
162	2	-7,0316	0,1780	1,0422	0	1,54	9,20
163	1	-15,2708	0,6856	-0,0088	1,1374	1,13	6,77
164	1	3,4901	-0,1719	0,0100	0,6025	1,10	8,77
165	2	-1,0502	0,1707	0,7706	0	0,91	7,16
166	1	56,5378	0,3873	-0,0057	-4,9213	0,75	8,10
167	2	-2,2517	0,2501	0,7670	0	0,79	8,70
170	1	-12,2716	0,6896	-0,0079	0,9043	1,10	4,24
171	1	-8,7438	0,5994	-0,0078	0,8988	1,04	4,20
172	1	-2,1064	0,5950	-0,0080	0,6344	1,03	4,45
173	1	-7,5240	0,4588	-0,0048	0,8743	1,30	6,86
174	1	-13,0292	0,9913	-0,0154	0,8689	0,84	4,22
175	1	-1,7849	0,3196	-0,0038	0,7702	1,09	6,06
176	2	-3,8157	0,1813	0,9147	0	0,92	5,67
177	2	6,4591	0,2619	0,2043	0	1,27	7,73
178	1	-10,5147	1,1487	-0,0198	0,6384	1,08	6,72
179	2	-4,0367	0,1817	0,9500	0	1,13	6,81
180	2	2,7712	0,2566	0,4333	0	1,31	7,39
181	2	-19,2091	0,1233	1,7850	0	1,44	7,49
182	2	2,8743	0,1177	0,6679	0	0,71	4,06
183	1	-3,6750	0,5343	-0,0066	0,6488	0,93	5,72
184	3	1,4076	0,4334	0	0	0,05	0,31
185	1	5,0241	-0,1406	0,0095	0,5003	0,99	6,81

*Modelo 1: $Ht = \beta_0 + \beta_1 DAP + \beta_2 DAP^2 + \beta_3 Hdom + \varepsilon$

Modelo 2: $Ht = \beta_0 + \beta_1 DAP + \beta_2 Hdom + \varepsilon$

Modelo 3: $\ln(Ht) = \beta_0 + \beta_1 \ln(DAP) + \varepsilon$

APÊNDICE 2 – CÓDIGOS DE PROGRAMAÇÃO EM R PARA TREINAMENTO DOS ALGORITMOS DE APRENDIZADO DE MÁQUINA E AJUSTE DE MODELOS NÃO LINARES

```
#####
#AJUSTE DO MODELO NÃO LINEAR DE GOMPERTZ COM INSERÇÃO DE COVARIANTES#
#####

install.packages('minpack.lm')
library(minpack.lm)

aj_gomp <- nlsLM(altura ~ (a+a1*hdom+a3*dg)*(exp(-(b+b2*idade)*exp(-(c+c1*hdom)*dap))),
  data = dataset,
  start = list(a=-6,a1=0.1,a3=0.01, b=1.8,b2=0.001,c=0.1,c1=0.001),
  control = nls.control(maxiter = 500))

summary(aj_gomp)
pred_gomp <- predict(aj_gomp, data = dataset)

#####
#TREINAMENTO DO ALGORITMO K-VMP#
#####

install.packages('kkn')
library(kkn)

tr_knn <- train.kkn(altura ~ dap+hdom+idade+dg,
  data=dataset,
  distance = 1, #MÉTRICA DE DISTÂNCIA (1 = MANHATTAN, 2 = EUCLIDIANA)
  kernel=c("gaussian",'optimal','inv','triangular', 'cos',
  'rectangular','epanechnikov', 'biweight','triweight'),
  ks = 1:100, #QUANTIDADE DE VIZINHOS MAIS PRÓXIMOS A SEREM TESTADOS
  cross = 3)#QUANTIDADE DE SUBCONJUNTOS PARA VALIDAÇÃO CRUZADA K-FOLD

pred_knn <- predict(tr_knn, data = dataset)

#####
#TREINAMENTO DO ALGORITMO FLORESTA ALEATÓRIA#
#####

install.packages('e1071')
install.packages('randomForest')
library(e1071)
library(randomForest)

rf_tune <- tune(randomForest, altura ~ dap+hdom+idade+dg, data = tns,
  ranges=list(mtry=c(2,3,4), #QUANTIDADE DE VARIÁVEIS POR NÓ
  ntree=c(seq(50,500,by=50))), #QUANTIDADE DE ÁRVORES
  tunecontrol = tune.control(sampling = 'cross', cross = 3))

rf_model <- FA_tune$best.model #SELECIONAR O MELHOR MODELO
pred_rf <- predict(rf_model, dataset)
```

```

#####
#TREINAMENTO DO ALGORITMO MÁQUINA DE VETORES DE SUPORTE#
#####

svm_tune <- tune.svm(altura~idade+dap+hdom+dg, data = dataset,
  type = 'nu-regression', # default: eps-regression
  gamma = 2^(-8:-1), # valores para gamma
  cost = 2^(1:8), #valores para C
  tunecontrol = tune.control(cross = 3)) #validação cruzada 3-fold

svm_model <- svm_tune$best.model
pred_svm <- predict(svm_model, dataset)

#####
#REPLICAÇÃO DO MODELO TREINADO DE MVS#
#####

vs <- as.matrix(model$SV) #MATRIZ DE VETORES DE SUPORTE
w <- as.matrix(model$coefs) #MATRIZ DE COEFICIENTES
g <- model$gamma #VALOR DO PARÂMETRO GAMMA
b <- model$rho #BIAS

#-----#
#PADRONIZAÇÃO DOS DADOS#
#-----#

med <- unlist(svm_model$x.scale[1]) #VETOR DE MÉDIAS DAS VARIÁVEIS PREDITORAS
sd <- unlist(svm_model$x.scale[2]) #VETOR DE DESVIOS PADRÃO DAS VARIÁVEIS PREDITORAS
xs <- matrix(nrow=nrow(dataset), ncol=4) #MATRIZ DE DADOS PADRONIZADOS
xs[,1] <- (training_set$dap-med[1])/sd[1] #PADRONIZAR A VARIÁVEL 1
xs[,2] <- (training_set$hdom-med[2])/sd[2] #PADRONIZAR A VARIÁVEL 2
xs[,3] <- (training_set$id-med[3])/sd[3] #PADRONIZAR A VARIÁVEL 3
xs[,4] <- (training_set$id-med[3])/sd[4] #PADRONIZAR A VARIÁVEL 4

#-----#
#KERNELIZAÇÃO DAS VARIÁVEIS PREDITORAS (FUNÇÃO DE BASE RADIAL)#
#-----#

kernelxs <- matrix(ncol = nrow(vs), nrow=nrow(xs)) #MATRIZ DE VARIÁVEIS KERNELIZADAS
for (i in 1:nrow(xs)) {
  for (j in 1:nrow(v)) {
    xi <- xs[i,]
    vj <- vs[j,]
    ximenosvj <- xi - vj
    ximenosvj2 <- t(as.matrix(ximenosvj)) %*% (as.matrix(ximenosvj))
    kernelxs[i,j] <- exp(-g*ximenosvj2)
  }
}

#-----#
#PREDIÇÃO#
#-----#

pred <- kernelxs %*% a-r #CALCULAR A SAÍDA PADRONIZADA
pred_svm <- pred*unlist(model$y.scale[2])+unlist(model$y.scale[1]) #DESPADRONIZAR AS
PREDIÇÕES

```

```

#####
#TREINAMENTO DO ALGORITMO REDES NEURAIS ARTIFICIAIS#
#####

install.package('h2o')
library(h2o)

h2o.init(nthreads = -1) # CONECTA AO SERVIDOR H2O

hidden1 = list(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15) #NÚMERO DE NEURÔNIOS NA CAMADA
OCULTA A SEREM TESTADOS
af = list('Tanh') #FUNÇÕES DE ATIVAÇÃO A SEREM TESTADAS
hyperp = list(hidden = hidden1, activation = af)

ann_tune = h2o.grid("deeplearning",
                    hyper_params = hyperp,
                    y = 'altura',
                    x = c('dap', 'idade', 'hdom', 'dg'),
                    training_frame = as.h2o(dataset),
                    nfolds = 3, #QUANTIDADE DE SUBCONJUNTOS DA VALIDAÇÃO CRUZADA K-FOLD
                    epochs = 3000,
                    stopping_rounds = 50,
                    reproducible = TRUE,
                    seed = 1234,
                    export_weights_and_biases = TRUE,
                    stopping_tolerance = 0)

ann_mod = h2o.getModel(ann_tune@model_ids[[1]]) #SELEÇÃO DO MELHOR MODELO
pred_ann <- as.data.frame(h2o.predict(ann_mod,as.h2o(dataset)))

#####
#REPLICAÇÃO DO MODELO DE RNA#
#####

ann_w1 <- as.matrix(h2o.weights(ann_mod, matrix_id=1)) #PESOS DA CAMADA DE ENTRADA
ann_w2 <- as.matrix(h2o.weights(ann_mod, matrix_id=2)) #PESOS DA CAMADA INTERMEDIÁRIA
ann_b1 <- as.matrix(h2o.biases(ann_mod, vector_id=1)) #BIAS DA CAMADA DE ENTRADA
ann_b2 <- as.matrix(h2o.biases(ann_mod, vector_id=2)) #BIAS DA CAMADA INTERMEDIÁRIA
ann_b12 <- do.call(rbind, rep(list(t(ann_b1)), nrow(dataset)))
ann_b22 <- do.call(rbind, rep(list(t(ann_b2)), nrow(dataset)))

xs <- apply(dataset, 2, scale) #PADRONIZAÇÃO DAS ENTRADAS

x_layer1 <- xs %*% t(ann_w1) + ann_b12
x_layer1_tanh <- tanh(x_layer1) #SAÍDAS DA CAMADA INTERMEDIÁRIA
pred_ann <- x_layer1_tanh %*% t(ann_w2) + ann_b2 #VALORES PREDITOS

```