

TRABALHO DE CONCLUSÃO DE CURSO
ENGENHARIA FLORESTAL

**REDES NEURAIIS ARTIFICIAIS PARA ESTIMATIVA
DA ALTURA DE ÁRVORES DO CERRADO**

DAISY CELESTINA SOUZA

Montes Claros – MG

2017

Daisy Celestina Souza

**REDES NEURAIIS ARTIFICIAIS PARA ESTIMATIVA DA
ALTURA DE ÁRVORES DO CERRADO**

Monografia apresentada ao Instituto de Ciências Agrárias da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Bacharel em Engenharia Florestal.

Orientador: Prof. Dr. Christian Dias Cabacinha

Montes Claros - MG

2017

Daisy Celestina Souza. **REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DA ALTURA DE ÁRVORES DO CERRADO**

Aprovada pela banca examinadora constituída por:

Carlos Alberto Araújo Júnior
Professor Dr. ICA/UFMG

Eugênio Monteiro da Silva Júnior
Mestre em Modelagem Computacional e Sistemas - UNIMONTES



Professor Dr. Christian Dias Cabacinha
Orientador ICA/UFMG

Montes Claros, 06 de julho de 2017.

Dedico a Deus, meus pais Gedeir e Edilaine, minhas irmãs Joicy e Hellen e a todos os amigos e mestres que me apoiaram nesta caminhada.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus pelas oportunidades, ensinamentos e vitórias que a cada dia me proporciona.

Aos meus pais, Gedeir por ser um exemplo de ética e humildade, e Edilaine por suas palavras de conforto e amor que sempre me levantaram nas dificuldades. Obrigada por nunca medirem esforços em prol da minha educação.

As minhas irmãs Joicy e Hellen, pelo espírito fraterno, cheio de brigas e amores. A minha avó, tios e primos, por serem meu suporte e fonte de alegria.

Aos mestres que dedicaram seu tempo e dedicação ao nosso aprendizado e em especial ao meu orientador Christian pelo imenso conhecimento compartilhado, confiança, respeito, paciência e por ser um exemplo de profissional e ao Professor Carlos pelos conselhos e boa receptividade.

Ao PET Florestal pelas experiências vividas e por todos os ensinamentos nos três pilares da universidade.

Aos ex-petianos por serem meu exemplo de dedicação, aprendi muito com vocês!

A galera do NEMAF, por terem me acolhido ainda precoce e apresentaram-me ao mundo da Engenharia Florestal.

Aos colegas de sala e amigos pelo convívio e por fazerem desta caminhada inesquecível. Agradeço em especial aos Friends (Suzy e Deni) pelos cafezinhos da tarde, Gabriel, Renan e Paulo pelas noites de estudo e diversão, Paty e Cris por sempre me acolherem com carinho, Lisandra, pelas palavras motivadoras, a Sara por sempre ter acreditado em mim e a todos que compartilhei momentos felizes, muito obrigada!

Agradeço também ao #202dopoder por serem praticamente minha segunda casa e pelo amor e carinho concedido, valeu meninas!

*“Palavras são, na minha nada humilde opinião,
nossa inesgotável fonte de magia, capazes de causar
grandes sofrimentos e também remediá-los”.*

(Albus Percival Wulfric Brian Dumbledore)

Resumo

Em florestas nativas, a variável altura deve ser obtida pela mensuração de todas as árvores amostradas para maior precisão do inventário, tornando o trabalho dispendioso e demorado. Esta variável também pode ser estimada com o auxílio de relações hipsométricas, as quais levam em consideração a correlação entre o DAP e a HT da árvore. Em florestas nativas esta correlação tende a ser baixa, o que é explicado pelo variado comportamento dentre as espécies. Uma outra estratégia de estimativa dessa variável é o uso de Redes Neurais Artificiais (RNA). Dessa forma, este trabalho teve como objetivo comparar o desempenho do uso da regressão e de Redes Neurais Artificiais na estimativa da altura de diferentes espécies do cerrado *sensu stricto* em função do diâmetro à altura do peito (DAP) e da adição da variável categórica espécie. A base de dados utilizada é oriunda de um estudo que foi realizado no Instituto de Ciências Agrárias da UFMG, no qual foi realizado a cubagem rigorosa de 580 árvores amostras de diferentes espécies e os dados de altura foram utilizados para este estudo. Para o processamento, os dados foram separados em 70% treinamento e 30% validação tanto para a regressão quanto para a RNA. Na regressão, foi modelada a equação de Curtis, utilizando a variável DAP e a variável categórica espécie como variáveis *dummy*. Da mesma maneira ocorreu com a RNA, utilizando o software *Neuroforest* e o algoritmo de treinamento *Resilient propagation* com oito neurônios na camada intermediária. Para avaliação da qualidade das estimativas utilizou-se o teste F de Gaybill, que tem por objetivo comparar a similaridade entre as estimativas. Após as análises, verificou-se que a estimativa da altura total pode ser realizada utilizando ambas as técnicas, (regressão ou RNA), e todas foram estatisticamente iguais a altura real, embora a não inclusão da variável categórica tenha gerado estimativas mais dispersas. Dessa forma é mais eficiente o uso da RNA com a inclusão da variável categórica espécie, pois a mesma apresenta como vantagem a melhor capacidade de aprendizado e generalização dos dados.

Palavras-chave: Modelagem florestal. Hipsometria. Inteligência artificial. Florestas naturais.

LISTA DE ILUSTRAÇÕES

Figura 1 - Fluxo de processamento do algoritmo <i>backpropagation</i>	18
Figura 2 - Esquema de realização do processamento dos dados.	22
Figura 3 - Gráfico de dispersão de resíduos para os ajustes com e sem a variável categórica espécie.	27
Figura 4 - Histograma de frequência dos resíduos para os ajustes com e sem a variável categórica espécie.....	28
Figura 5 - Gráfico de dispersão de resíduos para o processamento utilizando a RNA, com e sem a variável categórica espécie.....	30
Figura 6 - Histograma de frequência dos resíduos para o processamento utilizando RNAs com e sem a variável categórica espécie.....	30
Figura 7 - Gráficos de correlação entre a altura observada e a estimada nos diferentes métodos para a estimativa da altura total.....	32

LISTA DE TABELAS

Tabela 1 - Parâmetros do modelo, coeficiente de determinação ajustado e erro padrão residual dos ajustes.	26
Tabela 2 - RQME, correlação e bias para os ajustes e validações utilizando a regressão na estimativa da altura total com e sem a presença da variável categórica espécie.....	27
Tabela 3 - RQME, correlação e bias para o treinamento e validação utilizando as RNA na estimativa da altura total com e sem a presença da variável categórica espécie.....	29
Tabela 4 - Parâmetros estimados e estatísticas utilizadas na tomada de decisão no teste F (H_0) de Graybill.....	31

LISTA DE ABREVIATURAS E SIGLAS

CV	-	Coeficiente de variação
DAP	-	Diâmetro a altura de 1,3 metros
Ht	-	Altura total
MLP	-	Multilayer perceptron
$r(y, y)$	-	Coeficiente de correlação de Pearson
R^2 aj.	-	Coeficiente de determinação
RBF	-	<i>Radial basis function</i>
RMSE	-	Raiz quadrada do erro médio
RNA	-	Redes neurais artificiais
Syx (%)	-	Erro padrão residual

SUMÁRIO

1 INTRODUÇÃO	12
2 REFERENCIAL TEÓRICO	14
2.1 Estimativas de altura.....	14
2.2 Relações hipsométricas.....	15
2.3 Regressão com a variável <i>dummy</i>	16
2.4 Redes Neurais Artificiais.....	17
2.5 Aplicações de Redes Neurais artificiais.....	19
3 MATERIAL E MÉTODOS	21
3.1 Área de estudo	21
3.2 Coleta de dados.....	21
3.3 Processamento dos dados.....	21
3.3.1 Ajustes hipsométricos	22
3.3.2 Uso das Redes Neurais Artificiais.....	24
3.3.3 Comparação das metodologias.....	24
4 RESULTADOS E DISCUSSÃO	26
4.1 Ajustes hipsométricos	26
4.2 Uso das Redes Neurais Artificiais	29
4.3 Comparação das metodologias	31
CONCLUSÃO.....	33
REFERÊNCIAS	34

1 INTRODUÇÃO

O Cerrado é um domínio fitofisionômico que compreende três principais formações: florestal, savânica e campestre. Apresenta uma grande diversidade e está presente em 15 estados, além do Distrito Federal, faz fronteira com todos os outros biomas do Brasil e ocupa cerca de um quarto da superfície do nosso território. (MARIMON JUNIOR; HARIDASAN, 2005).

Devido à intensa exploração e redução da diversidade biológica do Cerrado, diversos estudos vêm sendo desenvolvidos como alternativa para auxiliar na definição de áreas para conservação, subsídio para elaboração de planos de manejo sustentável e até mesmo restauração de áreas (CHAVES *et. al.*, 2013).

O inventário florestal da vegetação nativa tem como objetivo conhecer a flora e sua estrutura, além de possibilitar as estimativas confiáveis do potencial da área a fim de otimizar a utilização dos recursos naturais. Por este motivo, estudos que avaliam o comportamento das variáveis dendrométricas e sua distribuição são de grande importância.

Dentre estas variáveis, a altura possui grande importância, pois possibilita a obtenção do volume, peso seco, estoque de carbono além de fornecer informações a respeito do comportamento dos indivíduos. Durante o inventário, para maior precisão, é necessário que sejam medidas as alturas de todos os indivíduos da parcela, o que torna o inventário muito trabalhoso. Também podem ser utilizadas relações hipsométricas, mas estas só são eficientes quando há correlação entre o DAP e a altura, o que por vezes não ocorre em florestas naturais (sobretudo nas fitofisionomina do Cerrado) já que há uma grande diversidade de padrões de crescimento (ANDRADE *et. al.*, 2015).

Atualmente, a utilização de Redes Neurais Artificiais tem sido uma ferramenta muito utilizada nas ciências florestais para resolução de problemas, já que possuem alta capacidade de aprendizado e generalização. Além disso as RNA's possibilitam a inserção de variáveis categóricas ao treinamento, como por exemplo a espécie, idade e clone o que auxilia no reconhecimento de padrões dos dados.

As RNA's são sistemas paralelos, compostos por neurônios artificiais, inspirados no funcionamento do sistema nervoso humano, interligadas por numerosas conexões as quais armazenam o conhecimento (BRAGA; CARVALHO; LUDERMIR, 2014). Em trabalhos encontrados na literatura são destacadas a maior precisão nos resultados encontrados pelas mesmas e a possibilidade de incluir variáveis categóricas na modelagem, tal como a variável

espécie. Por ser um tema recente, estudos que testem a aplicação das RNA's em florestas naturais são ainda escassos embora de grande importância.

Dessa forma este trabalho teve como objetivo comparar o desempenho do uso da regressão e de Redes Neurais Artificiais (RNA) na estimativa da altura de diferentes espécies do cerrado *sensu stricto* em função do diâmetro à altura do peito (DAP) e da adição da variável categórica espécie.

2 REFERENCIAL TEÓRICO

2.1 Estimativas de altura

A variável altura é de grande importância para estimativas do estoque potencial da floresta, como o volume, peso seco e estoque de carbono, classificação de sítios quanto ao crescimento e produção da floresta, dinâmica e sucessão, além da caracterização qualitativa da floresta (DIAMANTOPOULOU; ÖZÇELIK, 2012).

Segundo Souza e Soares (2013), a variável altura possibilita conhecer a estrutura vertical da floresta e dessa forma sua estrutura ecológica, posição sociológica das espécies, estado de regeneração, dentre outros parâmetros que são avaliados juntamente com a análise da estrutura horizontal. Estas informações fornecem subsídios para estudos de viabilidade técnica de exploração da floresta, estado de conservação e fase sucessional da mesma e também avaliar alternativas de gestão dos recursos (ÖZÇELIK *et. al.*, 2013).

Na rotina de inventário florestal, esta variável é obtida de forma indireta, com o auxílio de instrumentos específicos e por ser uma tarefa muito onerosa, em povoamentos equiâneos, são coletados somente a altura de parte das árvores amostradas e posteriormente com o auxílio da regressão são utilizadas equações específicas que estimam a altura das demais árvores do povoamento (CAMPOS; LEITE, 2013).

A obtenção indireta da altura está sujeita a muitas fontes de erro como a habilidade do operador, precisão do aparelho, condições meteorológicas do dia, declividade, distância e posição da árvore, dentre outras, desta forma faz-se necessário o uso de metodologias que garantam estimativas confiáveis das árvores não mensuradas, demandem pouco tempo e que sejam operacionais (SCOLFORO; THIERSCH, 2004).

Segundo Sanqueta *et. al.* (2013), devido à grande variação morfométrica das espécies, a realização de ajustes hipsométricos em florestas nativas é um trabalho desafiador e dessa forma são necessários estudos que abordem as diferentes formas de estratificação dos dados com o objetivo de aumentar a eficiência dos ajustes.

A estratificação dos dados proporciona ganhos na qualidade e precisão dos ajustes, essas estratificações podem ser por parcelas, classes diamétricas e espécie, podendo variar segundo o objetivo do pesquisador (CURTO *et. al.*, 2014).

2.2 Relações hipsométricas

As relações hipsométricas são obtidas pela correlação entre a altura das árvores e o DAP das mesmas e podem também estar relacionadas com as variáveis do povoamento tais como, a idade, altura dominante, qualidade do sítio e área basal. Estas relações devem ser aplicadas onde foram coletados os dados ou em povoamentos com características semelhantes (ANDRADE *et al.*, 2006; TOMÉ; RIBEIRO; FAIAS, 2007).

Para que a relação hipsométrica tenha desempenho satisfatório é necessário que a amostragem seja bem feita, captando as variações do povoamento. Dessa forma várias metodologias são desenvolvidas para definição do número de árvores à serem amostradas em cada parcela, como por exemplo as duas primeiras filas ou 20% do total amostrado (RIBEIRO *et al.*, 2010).

Gonçalves (2015) ao testar diferentes estratégias de ajustes para a obtenção de relações hipsométricas para clones de eucalipto, variando o número de árvores nas primeiras fileiras e das árvores dominantes a serem medidas, observou que ambas as estratégias foram eficientes para os ajustes das relações hipsométricas, cabendo ao inventariador escolher a estratégia que melhor se adequa às suas condições de trabalho.

Em florestas nativas, a correlação entre o DAP e a altura total é fraca, devido à diversidade de espécies e padrões de crescimento, dessa forma os modelos ajustados são fracos e pouco precisos.

Scolforo *et al.* (2015) testou diferentes abordagens para estimativa da altura total de clones de *Eucalyptus* spp. e observou que as estimativas eram mais precisas quando se inseria variáveis que auxiliavam na detecção de padrões de crescimento dos indivíduos, como a qualidade do fuste em seu estudo.

Diversos são os modelos matemáticos que estimam a altura total de árvores, mas para que essa estimativa seja a mais precisa possível, deve-se estabelecer critérios de seleção, analisando o gráfico de resíduos para verificar tendenciosidades em subestimar ou superestimar a variável dependente, o coeficiente de determinação ajustado em porcentagem e o erro padrão da estimativa (SOUSA *et al.*, 2013).

Essas estimativas são tradicionalmente feitas com modelos de regressão linear e não-linear, mas por levarem em consideração somente a variável DAP, nem sempre resulta em estimativas confiáveis. Dessa forma a inserção de outras variáveis como idade, altura dominante e índice de sítio em povoamentos equiâneos são muito importantes. Alguns

modelos foram desenvolvidos para obter o melhor resultado de ajuste, como por exemplo Curtis e Stofel (CAMPOS e LEITE, 2013).

Sanquetta *et. al.* (2013) ao testar diferentes modelos hipsométricos para estimativa da altura de espécies em uma floresta ombrófila mista observou que as estimativas genéricas forneciam baixo grau de ajuste com resíduos elevados e dispersos. Dessa forma, o autor realizou estratificação dos dados por famílias botânicas com o objetivo de melhorar as estimativas e embora tenha encontrado sucesso em algumas famílias, outras apresentaram resultados inferiores ao ajuste genérico, o que segundo o autor, pode ser explicado pela idade, competição, posição sociológica e baixa frequência dos indivíduos.

Nos últimos dez anos, surgiram alternativas que substituem a modelagem das relações obtidas pela regressão, as Redes Neurais Artificiais que têm por objetivo tornar as estimativas mais precisas com um menor tempo de processamento e dessa forma reduzir o custo do inventário florestal (BINOTI; BINOTI; LEITE, 2013).

2.3 Regressão com a variável *dummy*

Uma alternativa que auxilia na melhoria da qualidade dos ajustes é o uso de variáveis qualitativas, essas variáveis podem ser inseridas no modelo de regressão na forma de variáveis binárias e são chamadas *dummy*, o seu uso é simples e muito útil em situações que necessitam de variáveis não numéricas para o ajuste de regressão (SUITS, 1957).

Segundo Missio e Jacobi (2007), as variáveis *dummy* são variáveis binárias, ou seja, adotam valores de 0 ou 1 (presença ou ausência), dessa forma permitem explicar de melhor maneira o comportamento da variável resposta e assim aumentar a precisão e a qualidade dos ajustes. A mesma pode ser aplicada em modelos simples ou em modelos mais complexos, quando são adicionadas várias *dummies* ao modelo que possibilitam a inserção de variáveis qualitativas e dessa forma são de grande importância para a análise dos dados, assim, o uso dessas variáveis ampliam o poder de conhecimento do modelo e são indicadas em diversos tipos de estudo.

Schneider e Tonini (2007) ao estudar o volume de Acácia-negra utilizando variáveis *dummy*, observaram que a mesma mostrou-se eficiente, permitindo ajustar equações que levavam em consideração as diferentes fazendas amostradas.

2.4 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são sistemas paralelos compostos por neurônios (artificiais) simples que possuem capacidade de calcular determinadas funções matemáticas. Estes neurônios são interligados por conexões (sinapses) que estão associadas à pesos que armazenam e ponderam o conhecimento e possuem alta capacidade de generalização e aprendizado (BRAGA; CARVALHO; LUDEMIR, 2014).

Haykin (2001) afirma que as RNA's possuem alta capacidade de aprendizado pois possuem estrutura maciça paralela e devido aos seus pesos que são ajustados segundo a saída desejada, possuem a capacidade de resolver problemas considerados complexos.

O aprendizado das RNAs pode ser supervisionado no qual há a existência de um supervisor responsável por ajustar as entradas para que a saída estimada seja semelhante à observada, ou não-supervisionado, em que não há padrões de saída para a rede. Estes tipos de aprendizado são classificados, segundo o tipo de algoritmo à ser utilizado (BRAGA; CARVALHO; LUDEMIR, 2014).

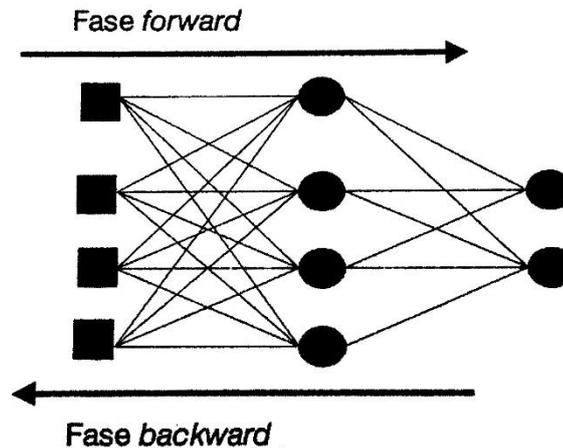
Os tipos mais comuns de RNA são as MLP (*Multilayer perceptron*) e RBF (*Radial basis function*). As Redes do tipo MLP possuem mais de uma camada intermediária e seus neurônios possuem funções de ativação não lineares, o que garante a essa configuração a vantagem de aproximação de qualquer função e conseqüentemente alta capacidade de aprendizado (SILVA JÚNIOR, 2016).

As redes tipo RBF utilizam funções de base radial nos neurônios de camada intermediária, a qual agrupa os dados de entrada em *cluster*, ou seja, em conjunto de padrões e os transformam de conjuntos não linearmente separáveis em linearmente separáveis, dessa forma a camada de saída classifica os padrões recebidos da camada anterior. As funções mais utilizadas são a Gaussiana, Multiquadrática e a Thin plate-spline. Já as redes tipo MLP se diferem por desenvolver as camadas intermediárias que por meio de transformações sucessivas tornam um problema não linear em linearmente separável. Estas transformações irão depender do número de neurônio em cada camada e também da complexidade do problema. As MLPs têm como vantagem a construção de aproximadores globais com maior capacidade de generalização já que a partição dos padrões de dados na RBF é feita por elipsoides (BRAGA; CARVALHO; LUDEMIR, 2014).

Segundo Görgens (2006) o aprendizado no algoritmo *Backpropagation* é supervisionado e ocorre em duas fases (FIGURA 1), a primeira é chamada de *Forward*, ocorre quando a entrada se propaga para frente e fornece uma saída, a segunda fase chamada

Backward, ocorre quando é calculado o erro em função da saída desejada e observada, este erro é propagado para trás, com o objetivo de ajustar os pesos das sinapses e dessa forma minimizar o erro da estimativa.

Figura 1 - Fluxo de processamento do algoritmo *Backpropagation*.



Fonte: Adaptado de: (BRAGA; CARVALHO; LUDEMIR, 2014).

Segundo Braga, Carvalho e Ludemir (2014), o algoritmo *Backpropagation* quando a saída de um neurônio for próxima de 0 ou 1 e o valor desejado for 0 ou 1, a derivada desse valor será próximo de zero fazendo com que os pesos recebam ajuste com valor mínimo ou até mesmo não serem ajustados, o que causa problemas de convergência. O algoritmo *Resilient propagation* tenta eliminar essa influência negativa e resolver o problema utilizando somente o sinal da derivada e não o seu valor, dessa forma indica se o ajuste do peso deverá aumentar ou diminuir.

O uso das RNA's aplicada às Ciências Florestais têm aumentado nos últimos anos e se apresenta como uma alternativa promissora em relação às técnicas de regressão devido a maior precisão dos resultados. Os estudos de Görgens (2006) e Silva *et al.*, (2009) com estimativas de volume de *Eucalyptus* sp., indicam a eficiência do uso das RNAs quando comparadas à modelagem por regressão.

Embora possua diversas aplicações, o uso das RNAs apresenta diferentes configurações que podem variar segundo o objetivo do trabalho, podendo ser modificados o número de neurônios da camada oculta, função de ativação, algoritmo de aprendizado e critérios de parada, dessa forma a melhor configuração pode ser definida de forma empírica (BINOTI, 2012).

2.5 Aplicações de Redes Neurais artificiais

Estudos envolvendo o uso de Redes Neurais Artificiais na estimativa de altura são pouco frequentes, dando destaque aos trabalhos de Binoti, Binoti e Leite (2013) e Vendruscolo *et al.* (2015) que estimaram a altura de povoamentos de *Eucalyptus* sp., Özçelik *et al.* (2013) com diferentes espécies de *Juniperus* sp., Castaño-Santamaría *et al.* (2013) em fragmentos de *Fagus sylvatica* L. e Almeida (2015) com a *Araucaria angustifolia* (Bert.) O. Ktze.

Em florestas nativas, devido a grande variação de espécies e conseqüentemente alta variabilidade, estes estudos são menos frequentes e abrangem outras variáveis além da altura, como no trabalho de Araújo (2015), Miguel *et al.* (2015), Miranda (2015) e Rodrigues *et al.* (2010).

Binoti, Binoti e Leite (2013) testaram a aplicação das RNAs na estimativa de altura de povoamento equiâneos de eucalipto com o objetivo de aumentar a precisão das estimativas e diminuir a necessidade de aferição da altura em campo e conseqüentemente reduzir os custos no inventário florestal. Puderam concluir que a metodologia desenvolvida possibilitou a redução do número de medições nas parcelas.

Em florestas nativas, o uso das RNA's tem sido aplicada em estimativas de diversas variáveis dendrométricas. Rodrigues *et al.* (2010) comparou a eficácia de modelos tradicionais de regressão e o uso das RNA's na estimativa de volume da espécie *Schizolobium amazonicum* (Huber) Ducke e concluiu que as RNA's foram eficazes na resolução do problema.

A utilização de RNA's na estimativa de estoques de biomassa seca e carbono de árvores em uma floresta estacional decidual foi considerada eficiente, apresentando estimativas bastante precisas, embora não tenham se diferido estatisticamente das estimativas geradas pelo modelo de Schumacher e Hall, que também apresentam estimativas razoavelmente precisas para a área (MIRANDA, 2015).

Araújo (2015) ao comparar a eficiência de equações hipsométricas e RNA's para a espécie *Calycophyllum spruceanum* Benth., concluiu que ambas metodologias possuem precisão semelhante, embora as RNA's sejam flexíveis quanto a inserção de variáveis categóricas como a idade. Neste trabalho, o autor utilizou redes do tipo MLP com o algoritmo de aprendizado *Resilient propagation* e o número de neurônios na camada intermediária foi definido de forma empírica.

Almeida (2015) ao estimar a altura de povoamentos de *Araucaria angustifolia*, utilizou redes com função de ativação tangente hiperbólica e critério de parada do treinamento quando o erro médio atingisse 0.0001 ou quando se chegasse a um número de 30.000 épocas. Os autores observaram que as RNAs apresentaram desempenho superior quando comparadas aos modelos hipsométricos, e segundo eles, isso pode ser explicado pela sua estrutura e distribuição paralela, rapidez no aprendizado e na capacidade de generalizar.

3 MATERIAL E MÉTODOS

3.1 Área de estudo

A área do estudo está localizada em uma área de reserva do Instituto de Ciências Agrárias (ICA) da UFMG no município de Montes Claros - MG, com vegetação predominante cerrado *sensu stricto* em bom estado de conservação. De acordo com a classificação Köppen o clima na região é do tipo Aw, tropical com estação seca, a temperatura média é 22,7 °C e o valor da pluviosidade média anual é 1.029 mm.

3.2 Coleta de dados

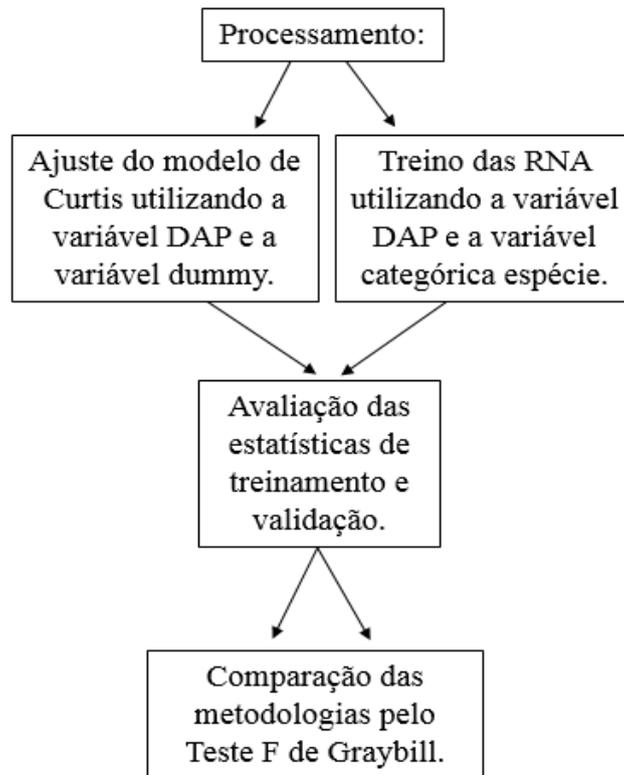
Os dados obtidos para este estudo foram aproveitados de um projeto já finalizado sobre estimativas de volume, peso seco e estoque de carbono realizado no ICA. A área de coleta dos dados foi distribuída em 25 parcelas de 400 m² alocadas de forma sistemática, totalizando uma amostragem de 1 ha. Foram amostrados todos os indivíduos lenhosos com diâmetro à altura de 1,3 metros igual ou superior a 3 cm, com essas árvores foi realizada a cubagem rigorosa pelo método de Huber (SCOLFORO; THIERSCH, 2004) e a medição da altura total foi realizada com o auxílio de uma trena.

3.3 Processamento dos dados

Das 580 árvores medidas, os dados de 406 delas foram destinadas aos ajustes do modelo hipsométrico e treinamento das RNA's. Os dados das outras 174 árvores foram considerados para validação dos ajustes e treinamentos. Tal divisão foi realizada de forma aleatória.

O processamento ocorreu da seguinte maneira (Figura 2):

Figura 2 - Esquema de realização do processamento dos dados.



Fonte: Da autora, 2017.

3.3.1 Ajustes hipsométricos

Os ajustes hipsométricos foram realizados com o auxílio do *software* R Studio® e ocorreram em duas etapas: a primeira foi utilizando apenas o DAP como variável independente e a segunda utilizando o DAP e a variável categórica espécie, na forma de variável *dummy*.

O modelo escolhido para os ajustes foi o de Curtis na sua forma linearizada (Modelo 1), já que é o modelo mais utilizado e apresenta maior sucesso neste tipo de estudo, como no trabalho de Soares *et. al.* (2011).

$$\text{LnHt} = \beta_0 + \beta_1 1/\text{DAP} + \varepsilon \quad (\text{Modelo 1})$$

Onde:

β_0 e β_1 são os parâmetros do modelo;

LnHt = o logaritmo neperiano da altura;

$1/\text{DAP}$ = o inverso do DAP, a variável independente do modelo;

ε = erro da estimativa.

Para os ajustes com a variável *dummy* (Modelo 2), aquelas espécies que possuíam apenas um indivíduo na base de dados e outras com uma pequena frequência que não justificava a sua utilização como uma variável categórica por não haver número de dados suficiente para reconhecimento de padrões importantes para a modelagem, foram agrupadas como uma única espécie, dessa forma o número de espécies caiu de trinta e oito para quinze. Para o processamento, estas foram codificadas na forma binária no banco de dados.

$$\text{LnHt} = \beta_0 + \beta_1 1/\text{DAP} + \beta_2 D_1 + \beta_3 D_2 + \dots + \beta_n D_{n-1} + \varepsilon \quad (\text{Modelo 2})$$

Onde:

β_0, \dots, β_n são os parâmetros do modelo;

LnHt = o logaritmo neperiano da altura;

$1/\text{DAP}$ = o inverso do DAP, a variável independente do modelo;

$(D_1, D_2, \dots, D_{n-1})$ = variável *dummy* para cada espécie presente no estudo.

ε = erro da estimativa.

Para avaliação da eficiência dos ajustes, foram avaliados o gráfico de dispersão de resíduos em porcentagem, o erro padrão da estimativa em porcentagem ($\text{Syx}\%$) (2), o coeficiente de determinação ajustado ($R^2_{aj.}$) (3) e a correlação de Pearson $r(y, \hat{y})$ (4), o bias (5) e a raiz quadrada do erro médio em porcentagem (RQME) (6).

$$\text{Syx} (\%) = \frac{100}{\bar{Y}} * \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n-p}} \quad (2)$$

$$r(y, \hat{y}) = \frac{\sum_{i=1}^n y\hat{y} - \frac{(\sum y)(\sum \hat{y})}{n}}{\sqrt{\sum_{i=1}^n y^2 - \frac{(\sum y)^2}{n}} * \sqrt{\sum_{i=1}^n \hat{y}^2 - \frac{(\sum \hat{y})^2}{n}}} \quad (3)$$

$$R^2_{aj.} = \left(\frac{n-1}{n-p}\right) * \left(\frac{\text{SQresíduos}}{\text{SQtotais}}\right) \quad (4)$$

$$\text{Bias} = \frac{\sum_{i=1}^n (y - \hat{y})}{n} \quad (5)$$

$$\text{RQME (\%)} = \frac{100}{\bar{Y}} * \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n}} \quad (6)$$

Onde:

\bar{Y} = média da altura real;

y = valores da altura reais (observadas);

\hat{y} = valores da altura estimadas;

n = número de observações;

p = número de parâmetros do modelo;

SQ = soma dos quadrados dos resíduos;

SQtotais = soma dos quadrados totais.

n = número total de observações.

3.3.2 Uso das Redes Neurais Artificiais

A configuração das redes seguiu a metodologia adotada por Araújo (2015) em seu trabalho com o Pau-mulato, dessa forma, para o processamento dos dados foi utilizado o software *Neuroforest*[®] (versão 3.3), a arquitetura de rede utilizada foi a Perceptron de Múltiplas Camadas (MLP), com uma camada intermediária de oito neurônios e um neurônio na camada de saída (Ht). Para o treinamento e validação, os dados foram separados em 70% e 30% respectivamente. O algoritmo de aprendizagem utilizado foi o *Resilient Propagation* e a função de ativação do tipo sigmoïdal tanto na camada intermediária quanto na de saída. Como critério de parada do treinamento das redes, utilizou-se o erro médio igual a 0,0001 ou o número total de épocas igual a 3.000, o que primeiro ocorresse. Foram treinadas cinco redes, dessas, a que teve o melhor desempenho foi armazenada e suas estimativas foram avaliadas com base nos coeficientes de correlação de Pearson (3), o bias (5) e a raiz quadrada do erro médio em porcentagem (RQME) (6), além da análise gráfica dos resíduos (BINOTI; BINOTI; LEITE, 2014).

3.3.3 Comparação das metodologias

Para comparar a eficiência das metodologias na estimativa da altura total, utilizou-se o teste F modificado de Graybill (1976), o qual realiza um teste de identidade entre dois vetores,

isto é, faz a comparação de grupos de dados quantitativos, onde um é considerado como método padrão (Y) e o outro como o método alternativo (X) (MIRANDA, 2012). Dessa forma, neste estudo foi ajustado a seguinte equação:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Onde:

Y é a variável dependente, neste caso a altura real;

X é a variável independente, as alturas estimadas em cada metodologia no estudo.

A similaridade das variáveis é observada pelas hipóteses:

$$H_0: \beta \rightarrow = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ e } H_1 = \text{rejeição de } H_0$$

Portanto, a não rejeição de H_0 permite concluir que, a um nível de significância as estimativas são estatisticamente iguais aos valores das alturas reais. Para testar esta hipótese, calculou-se a estatística F de Graybill:

$$F(H_0) = \frac{(\hat{\beta} - \theta)' * (X'X)^{-1} * (\hat{\beta} - \theta)}{2 * QM_{\text{resíduo}}}$$

Onde:

$\hat{\beta}' = [\beta_0 \ \beta_1]$: vetor dos parâmetros da equação de regressão.

$$\theta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

$$(X'X)^{-1} = \begin{bmatrix} n & \sum x \\ \sum x & \sum x^2 \end{bmatrix}^{-1}$$

$$F_{\text{tab.}} = [5\%; 2, n - 2 \text{ gl}].$$

4 RESULTADOS E DISCUSSÃO

4.1 Ajustes hipsométricos

Os parâmetros obtidos no ajuste do modelo de Curtis sem a inclusão da variável categórica foram significativos pelo teste t à 5%, ou seja, são diferentes de zero. Em contrapartida, quando foram adicionadas as variáveis *dummy*, os parâmetros para algumas espécies não foram significativos (TABELA 1). Isso pode ser explicado pela baixa frequência de indivíduos em algumas espécies, comportamento próprio de florestas nativas.

Tabela 1 - Parâmetros do modelo, coeficiente de determinação ajustado e erro padrão residual dos ajustes.

Ajuste sem categórica				Ajuste com categórica			
Parâmetros	R ² aj.	Syx(%)		Parâmetros	Espécies	R ² aj.	Syx(%)
β_0	2,01707*	0,33240	26,7888	1,91017*	-	0,57350	21,3002
β_1	2,50710*	-	-	-2,93628*	-	-	-
β_2	-	-	-	0,01492 ^{ns.}	<i>Sclerolobium</i> sp.	-	-
β_3	-	-	-	0,22321 ^{ns.}	<i>Terminalia argentea</i>	-	-
β_4	-	-	-	0,36725*	<i>Copaifera langsdorffii</i>	-	-
β_5	-	-	-	0,19877 ^{ns.}	<i>Heteropteris byrsonimifolia</i>	-	-
β_6	-	-	-	0,26199*	<i>Tocoyena formosa</i>	-	-
β_7	-	-	-	-0,01508 ^{ns.}	<i>Machaerium opacum</i>	-	-
β_8	-	-	-	-0,14285 ^{ns.}	<i>Curatella americana</i>	-	-
β_9	-	-	-	0,23164 ^{ns.}	<i>Alibertia edulis</i>	-	-
β_{10}	-	-	-	0,13195 ^{ns.}	<i>Byrsonima crassiflora</i>	-	-
β_{11}	-	-	-	0,33364*	Outras	-	-
β_{12}	-	-	-	0,49274*	<i>Sebastiana brasiliensis</i>	-	-
β_{13}	-	-	-	0,45067*	<i>Combretum leprosum</i>	-	-
β_{14}	-	-	-	0,30337*	<i>Jacaranda brasiliensis</i>	-	-
β_{15}	-	-	-	0,21324 ^{ns.}	<i>Magonia pubescens</i>	-	-
β_{16}	-	-	-	0,08157 ^{ns.}	<i>Acosmium dasycarpum</i>	-	-

β = parâmetros do modelo; * significância para o nível de 5% de probabilidade no teste t; ^{ns.} não significativo para o nível de 5% de probabilidade no teste t.

Sanquetta *et. al.* (2013) em seu estudo observou que em florestas nativas as equações de regressão não tiveram bom desempenho, com baixo grau de ajuste (R^2 aj. = 0,41) e erro padrão da estimativa médio (Syx%) com valor mais alto (18,64%) e resíduos elevados e com grande dispersão. Ele afirma que este comportamento é explicado pelos distintos comportamentos morfométricos das espécies e neste caso somente a estratificação dos dados, por família, gênero ou espécies, forneceria ajustes consistentes.

Almeida (2015) em seu estudo com a Araucária encontrou valores de R^2 aj. variando de 68 a 78% e Syx (%) de 13 a 22%. Estes valores são estatisticamente bons e são explicados pelo bom padrão de desenvolvimento da espécie.

O ajuste utilizando a variável categórica (*dummy*) forneceu melhores resultados quando comparado com o método tradicional utilizando somente o DAP como variável independente. Seus valores de correlação foram superiores, dessa forma a equação consegue explicar melhor o comportamento dos dados. O erro médio também foi inferior, tanto nos ajustes quanto na validação (Tabela 2).

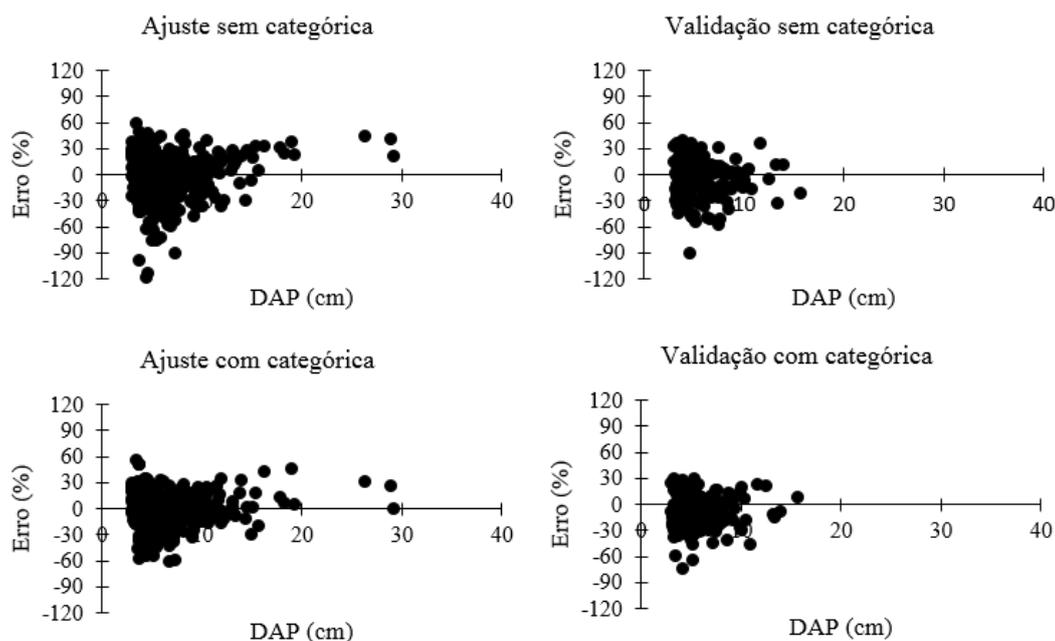
Tabela 2 - RQME, correlação e bias para os ajustes e validações utilizando a regressão na estimativa da altura total com e sem a presença da variável categórica espécie.

Ajuste	Tipo de dados	RQME	$r(y, \hat{y})$	Bias
Sem categórica	Ajuste	26,7227	0,6050	-0,0026
	Validação	20,9109	0,5818	-0,0046
Com categórica	Ajuste	20,8495	0,7798	0,0009
	Validação	18,8237	0,7238	-0,0113

Fonte: Da autora, 2017.

Ao observar a distribuição dos resíduos para os ajustes e a validação sem e com a variável categórica espécie (FIGURA 3), verifica-se que em ambos os ajustes ocorreram uma tendência em subestimar a altura total no intervalo de 20 – 30 cm de DAP.

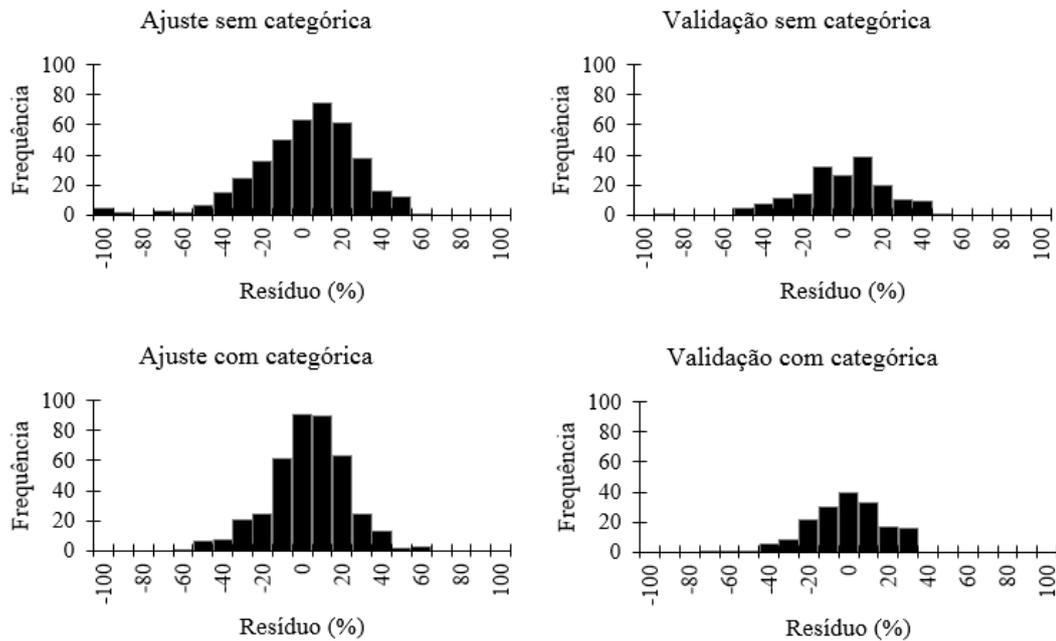
Figura 3 - Gráfico de dispersão de resíduos para os ajustes com e sem a variável categórica espécie.



Fonte: Da autora, 2017.

O histograma de frequências dos resíduos (FIGURA 4) apresentou ocorrência de erros dispersa no intervalo de -100 a +60%, verificando-se ainda uma maior frequência de erros negativos indicando uma tendência geral de superestimativa da altura. Com relação a validação do ajuste, o uso da variável categórica se aproximou mais da distribuição normal do que o ajuste utilizando somente a variável DAP.

Figura 4 - Histograma de frequência dos resíduos para os ajustes com e sem a variável categórica espécie.



Fonte: Da autora, 2017.

Além do uso de variáveis *dummy* no ajuste, a inserção de outras variáveis no modelo como, a altura da copa e o diâmetro médio quadrático das espécies podem garantir melhor qualidade no ajuste (Andrade *et. al.*, 2006). Costa *et al.* (2012), que trabalharam com *Copaíba* em diferentes fisionomias de floresta, observaram que o R^2 aj. para o cerrado foi bem inferior que o obtido em florestas semidecíduais, 0,30 em comparação com 0,80. Segundo o autor, devido à maior competição por luz e profundidade do solo, as plantas de *Copaifera langsdorffii* da floresta semidecidual investem mais no seu crescimento em altura, numa tentativa de dominar o dossel da comunidade, o que não ocorre com a mesma espécie no cerrado e dessa forma é uma explicação para suas fracas relações diâmetro-altura.

Para melhorar essas relações, pode-se realizar uma estratificação dos dados, como no trabalho de Sanquetta *et. al.* (2013), no qual os autores realizaram os ajustes por famílias botânicas e encontraram melhores resultados na relação diâmetro-altura, embora a mesma não tenha sido suficiente para explicar variações internas de cada caso. Segundo os mesmos

autores, outras formas de estratificação podem ser utilizadas, devendo-se atentar ao número mínimo de amostras, e também na forma de obtenção desses dados.

4.2 Uso das Redes Neurais Artificiais

Observando os valores de correlação e RQME (TABELA 3), o uso da RNA foi levemente superior a regressão, embora o bias em ambas tenha demonstrado que os resíduos tenderam a superestimar a altura total das espécies. O Bias indica a existência ou não de tendências entre os resíduos, porém não mede quão afastados do eixo zero estão os valores residuais (ARAÚJO, 2015).

Tabela 3 - RQME, correlação e bias para o treinamento e validação utilizando as RNA na estimativa da altura total com e sem a presença da variável categórica espécie.

RNA	Tipo de dados	RQME	r (y, \hat{y})	Bias
Sem categórica	Treinamento	24,1834	0,6829	-0,0004
	Validação	21,5101	0,5582	-0,1255
Com categórica	Treinamento	17,5173	0,8486	-0,0010
	Validação	19,6571	0,7143	-0,2435

Fonte: Da autora, 2017.

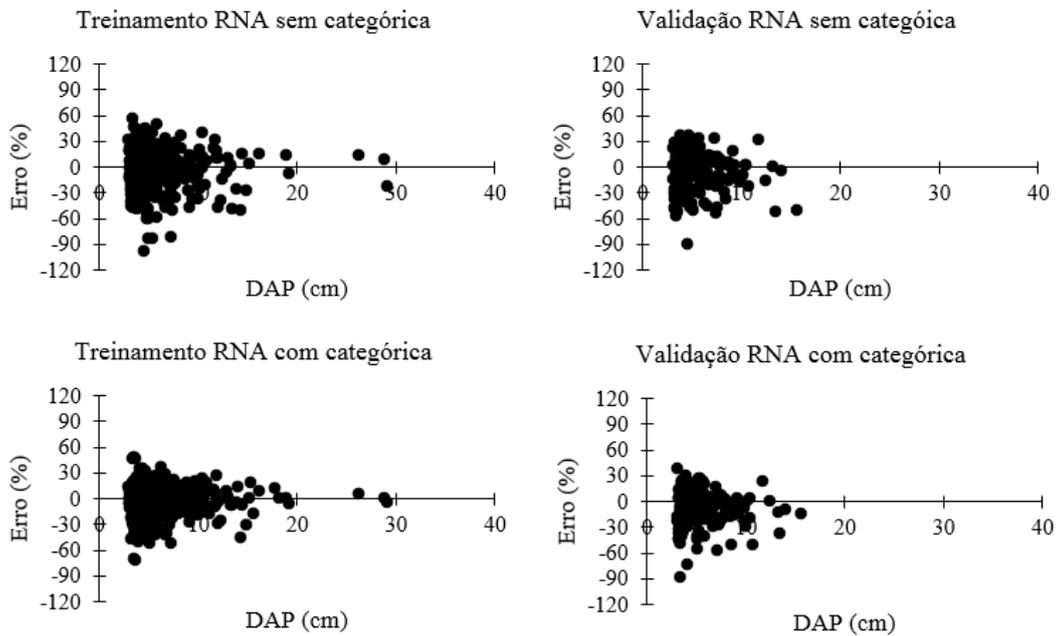
Podemos observar que a inserção da variável categórica espécie contribui de forma positiva no desempenho da rede, reduzindo seu erro e aumentando a correlação entre os dados, de maneira similar ao uso da regressão.

Almeida (2015) ao comparar regressão e o uso de RNA na estimativa da altura total de Araucária, observou que a RNA selecionada apresentou-se superior ao melhor modelo hipsométrico com relação a todas as estatísticas de qualidade de ajuste. O mesmo ainda explica que isso ocorre devido a distribuição paralela e na capacidade de generalizar das RNA's, as mesmas têm a capacidade de resolver problemas mais complexos e modelar uma multiplicidade de variáveis.

A capacidade de aprendizado da rede é uma vantagem nos casos em que é necessário modelar comportamentos complexos, como é o caso da altura total em florestas nativas (ÖZÇELİK *et. al.*, 2013).

Quando comparada com o uso da regressão, observamos que o uso da RNA forneceu melhores valores de RQME e correlação. Além disso, a dispersão dos resíduos não apresentou tendências e se apresentou mais homogênea quanto a sua distribuição (Figura 5).

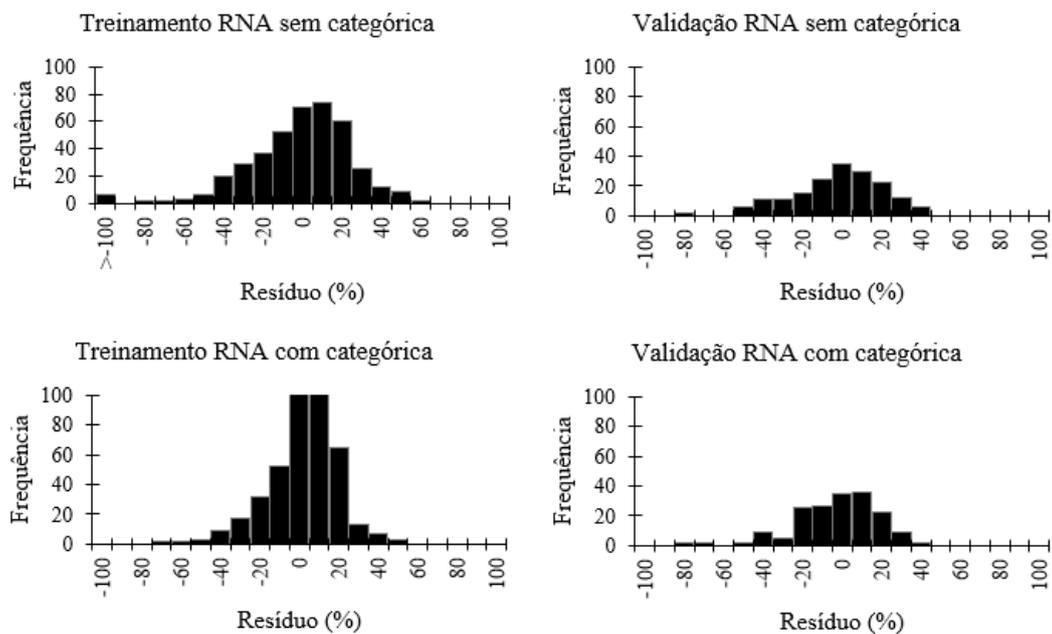
Figura 5 - Gráfico de dispersão de resíduos para o processamento utilizando a RNA, com e sem a variável categórica espécie



Fonte: Da autora, 2017.

Quanto a distribuição da frequência dos resíduos (FIGURA 6), de forma semelhante à regressão ambas técnicas tiveram dispersão concentrada nas classes de 20 a -20%, com padrão normal de distribuição, o que é esperado para este tipo de estudo.

Figura 6 - Histograma de frequência dos resíduos para o processamento utilizando RNAs com e sem a variável categórica espécie.



Fonte: Da autora, 2017.

Araújo (2015) realizou a modelagem da altura total de *Calycophyllum spruceanum* utilizando a regressão e as redes neurais e observou que ambas forneceram estimativas próximas, destacando a vantagem da RNA por ser flexível quanto a inserção de variáveis quantitativas e qualitativas, além de ter uma grande capacidade de aprendizado e generalização.

4.3 Comparação das metodologias

Ao observar o valor do F calculado para o teste de Graybill, verifica-se que todas as metodologias foram estatisticamente iguais à altura real (TABELA 4). Dessa forma, ambas poderiam ser indicadas para estimativas confiáveis da altura de espécies do cerrado.

Tabela 4 - Parâmetros estimados e estatísticas utilizadas na tomada de decisão no teste F (H_0) de Graybill.

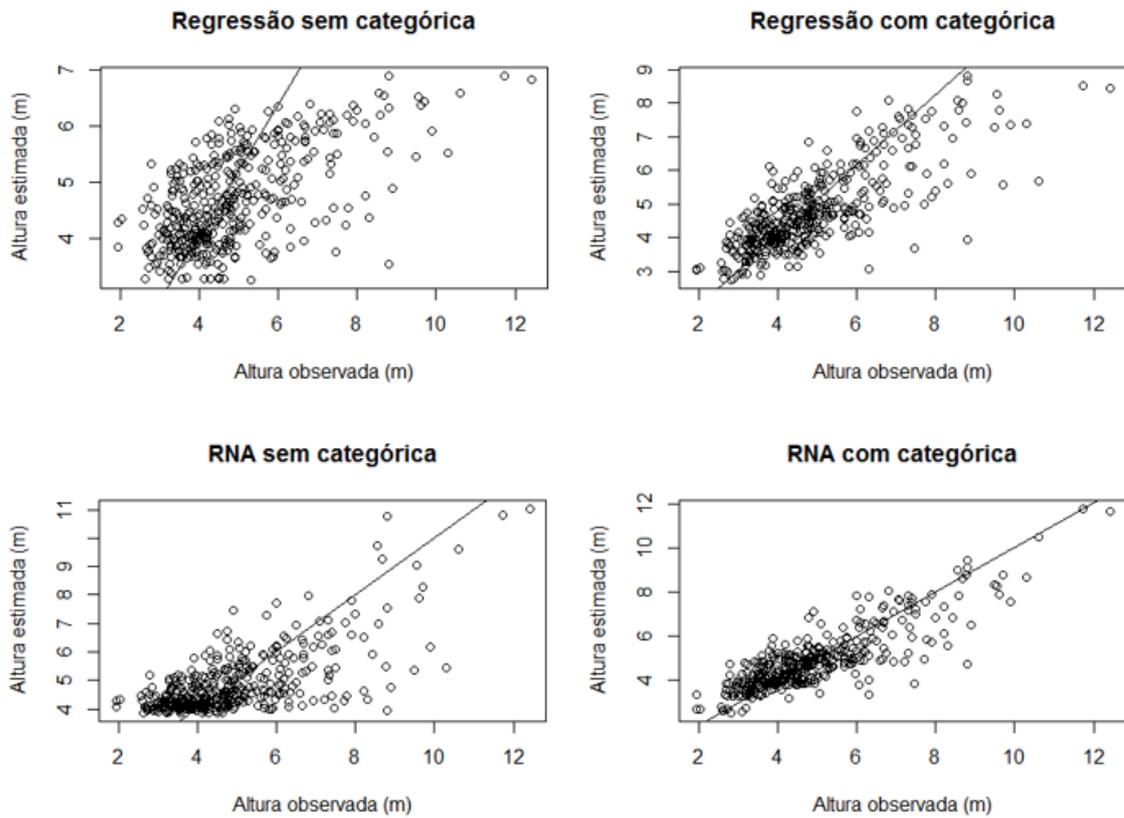
Método	Ajuste	β_0	β_1	F _{calculado}	F _{tabelado}	CV (%)
Regressão	Sem categórica	-0,6138	1,1655	0,010278 ^{ns.}	3,02	33,87
	Com categórica	-0,1458	1,0503	0,000498 ^{ns.}	3,02	20,93
RNA	Sem categórica	-0,0125	1,0025	0,000003 ^{ns.}	3,02	28,51
	Com categórica	-0,0335	1,0067	0,000028 ^{ns.}	3,02	14,95

^{ns.} não significativo a 5% de probabilidade pela distribuição F.

Embora todas as estimativas tenham sido estatisticamente iguais à altura real, quando observa-se o coeficiente de variação (CV) das regressões, verifica-se que aquelas estimativas com o uso de somente o DAP como variável independente apresentam alta variação, no caso da regressão com 33,87% e da RNA com 28,51%. Ao observar o CV das estimativas com o uso da variável categórica, esse número cai quase pela metade, o que indica menor variação do acaso. Dessa forma, o mais indicado é o uso da variável categórica, seja pela modelagem por regressão ou por RNA.

Na Figura 7 observa-se a importância da inserção da variável categórica na correlação dos dados. Os mesmos tiveram a tendência de formar uma reta de 45° de inclinação, comportamento esperado em estudos de modelagem.

Figura 7 - Gráficos de correlação entre a altura observada e a estimada nos diferentes métodos para a estimativa da altura total.



Fonte: Da autora, 2017.

As técnicas utilizadas neste estudo (regressão e RNA) com a presença da variável categórica apresentaram os melhores resultados para RQME e correlação de Pearson. A dispersão dos resíduos foi mais uniforme, com uma leve tendência no ajuste de regressão, além disso, a frequência dos resíduos foi mais concentrada próximo à zero, dessa forma a mesma é indicada com uma adequada metodologia para estimativa da altura total de árvores do cerrado.

CONCLUSÃO

Pode-se concluir que, as estimativas da altura total podem ser realizadas utilizando ambas as técnicas, regressão ou RNA, todas foram estatisticamente iguais à altura real, sendo mais eficiente o uso da variável categórica espécie.

Referências

- ALMEIDA, T. G. **Redes neurais artificiais para estimativa de altura em povoamentos de *Araucaria angustifolia***. 2015, 56 f. Trabalho de conclusão de curso (Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2015.
- ANDRADE, V. C. L.; KROETZ, E. A.; NICOLA, A.; SOUZA, P. B.; NOHAMA, F. K.; LEITE, H. G.; BINOTI, D. H. B.; BINOTI, M. L. M. S. Amostragem e agrupamento de dados de relação hipsométrica em inventários florestais de Cerrado Tocantinense. **Pesquisa florestal brasileira**, v.35, n.83, p.227-238, 2015.
- ANDRADE, V. C. L.; MARTINS, S. V.; CALEGARIO, N.; CHICHORRO, J. F. Influência de três sistemas de amostragem na estimativa da relação hipsométrica e do volume de árvores em um fragmento de Mata Atlântica. **Revista Scientia Forestalis**, São Paulo, v. 70, p. 31-37, 2006.
- ARAÚJO, B. H. P. D. **Modelagem da altura, volume e afilamento do fuste de *Calycophyllum Spruceanum* Benth. empregando regressão e redes neurais artificiais**. 2015. 53 f. Dissertação (mestrado em Ciências de Florestas Tropicais) – Instituto Nacional de Pesquisas Amazônicas, Manaus, 2015.
- 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. 05, n. 01, p. 58-67, 2014.
- BINOTI, M. L. M. S. **Emprego de Redes Neurais Artificiais em Mensuração e Manejo Florestal**. 2012. 138 f. 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**, Viçosa, v. 37, n. 4, p. 639-645, 2013.
- BRAGA, A. D. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. Ed. Rio de Janeiro: LTC, 2014. 226p.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. 4. Ed. Viçosa: UFV, 2013. 605 p.
- CASTAÑO-SANTAMARÍA J, CRECENTE-CAMPO F, FERNÁNDEZ-MARTÁNEZ JL, BARRIO-ANTA M, OBESO JR. Tree height prediction approaches for uneven-aged beech forests in northwestern Spain. **Forest Ecology and Management** v. 307, p. 63-73, 2013.
- CHAVES, A. D. C. G.; SANTOS, R. M. S.; SANTOS, J. O.; FERNANDES, A. A.; MARACAJÁ, P. B. A importância dos levantamentos florístico e fitossociológico para a conservação e preservação das florestas. **Agropecuária Científica no Semiárido**, Campina Grande, v.9, n.2, p. 43-48, 2013.
- COSTA, M. P.; PEREIRA, J. A. A.; BENICIO, M. H. M.; SOUZA, H.; FONTES, M. A. L.; GARCIA, P. O. Alometria e arquitetura de *Copaifera langsdorffii* (desf.) kuntze (fabaceae)

em fitofisionomias neotropicais no sul de Minas Gerais. **Ciência Florestal**, Santa Maria, RS, v. 22, n. 2, p. 223-240, 2012.

CURTO, R. A. J; LOUREIRO, G. H, MÔRA, R; MIRANDA, R. O. V; NETTO, S. P; SILVA, G. F Relações hipsométricas em floresta estacional semidecidual. **Ciências Agrárias**, v.57, n.1, p.57-66, 2014.

DIAMANTOPOULOU, M. J.; ÖZÇELIK, R. Evaluation of different modeling approaches for total tree-height estimation in Mediterranean Region of Turkey. **Forest Systems**, v. 21, n. 3, p. 383-397, 2012.

GONÇALVES, A. F. A. **Relação hipsométrica em povoamentos de eucalipto: avaliação de diferentes estratégias de ajustes**. 2015. 194 f. Trabalho de conclusão de curso (Engenharia Florestal) – Instituto de Ciências Agrárias, Universidade Federal de Minas Gerais, Montes Claros, 2015.

GÖRGENS, E. B. **Estimação do volume de árvores utilizando redes neurais artificiais**. 2006. 94 f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2006.

GRAYBILL, F.A. **Theory and application of linear model**. Belmont: Duxbury Press, 1976. 413p.

HAYKIN, S. S. **Redes neurais artificiais: princípio e prática**. 2. Ed. São Paulo: Bookman, 2001. 908 p.

MARIMOND JUNIOR, B. H.; HARIDASAN, M. Comparação da vegetação arbórea e características edáficas de um cerradão e um cerrado sensu stricto em áreas adjacentes sobre solo distrófico no leste de Mato Grosso, Brasil. **Acta Botanica Brasilica**, v. 19, n. 4, p. 913-926, 2005.

MIGUEL, E. P.; REZENDE, A. V.; LEAL, F. A.; MATRICARDI, E. A. T.; VALE A. T.; PEREIRA, R. S. Redes neurais artificiais para a modelagem do volume de madeira e biomassa do cerradão com dados de satélite. **Revista Pesquisa Agropecuária Brasileira**, v. 50, n. 9, p. 829-839, 2015.

MIRANDA, J. F. N. D. **Modelos de regressão e de redes neurais artificiais na quantificação de carbono e biomassa lenhosa em floresta estacional decidual no Brasil Central**. 2015. 88f. Dissertação (mestrado em Ciências Florestais) – Faculdade de Tecnologia, Universidade de Brasília, Brasília, 2015.

MIRANDA, R. O. V. **Modelagem em nível de povoamento para Eucalyptus sp. com estratificação em classes de precipitação pluviométrica**. 78 f. Dissertação (Mestrado em Ciências Florestais) – Universidade Estadual do Centro-Oeste, Irati, 2012.

MISSIO, F.; JACOBI, L. Variáveis dummy: especificações de modelos com parâmetros variáveis. **Ciência e Natura**, v. 29, n. 1, p. 111-135, 2007.

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

RIBEIRO, A.; FERRAZ FILHO, A. C.; MELLO, J. M.; FERREIRA, M. Z. Estratégias e metodologias de ajuste de modelos hipsométricos. **Revista Cerne**, Lavras, v. 16, n. 1, p. 22-31, 2010.

RODRIGUES, E. F.; OLIVEIRA, T. F.; MADRUGA, M. R.; SILVEIRA, A. M. Um Método para determinar o volume comercial do *Schizolobium amazonicum* (Huber) Ducke utilizando redes neurais artificiais. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 1, p. 16-23, 2010.

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

SCHNEIDER, P. R.; TONINI, H. Utilização de variáveis dummy em equações de volume para *Acacia mearnsii* De Wild. **Ciência Florestal**, v. 13, n. 2, 2003.

SCOLFORO, H. F.; RAIMUNDO, R.; SCOLFORO, R. S.; MELLO, J. M.; BATISTA, A. P. B.; & BULLOCK, B. Hypsometric approaches to Eucalyptus experiments in Brazil. **African Journal of Agricultural Research**, v. 45, n. 10, p. 4176-4184, 2015.

SCOLFORO, J. R. S.; THIERSCH, C. R. **Biometria florestal: medição, volumetria e gravimetria**. 1. Ed. Lavras: UFLA/FAEPE, 2004. 285 p.

SILVA JÚNIOR, E. M. **Estimativa de volume de madeira de árvores de eucalipto utilizando redes neurais artificiais do tipo RBF**. 2016. 75f. Dissertação (mestrado em Modelagem Computacional e Sistemas) – Universidade Estadual de Montes Claros, Montes Claros, 2016.

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

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

SOARES, C. P. B.; MARTINS, F. B.; LEITE-JR, H. U.; SILVA, G. F.; FIGUEIREDO, L. T. M. Equações hipsométricas, volumétricas e de taper para onze espécies nativas. **Revista Árvore**, Viçosa, MG, v. 35, n. 5, p. 1039-1051, 2011

SOUSA, G. T. O.; AZEVEDO, G. B.; BARRETO, P. A. B.; JÚNIOR, V. C. Relações hipsométricas para *Eucalyptus urophylla* conduzidos sob regime de alto fuste e talhadia no Sudoeste da Bahia. **Revista Scientia Plena**, São Cristóvão, v. 9, n. 4, p. 1-7, 2013.

SOUZA, A. L.; SOARES, C. P. B. **Florestas Nativas: estrutura, dinâmica e manejo**. 1. Ed. Viçosa: UFV, 2013. 322 p.

SUITS, D. B. Use of dummy variables in regression equations. **Journal of the American Statistical Association**, v. 52, n. 280, p. 548-551, 1957.

TOMÉ, M; RIBEIRO, F.; FAIS, S. Relação hipsométrica geral para *Eucalyptus globulus* Labill. em Portugal. Lisboa – Portugal, **Revista Silva Lusitana**, Oeiras, v. 15, n. 1, p. 41 - 55, 2007.

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